

Настоящее изобретение относится к последовательностям ДНК, расположенным в 3'-5'-направлении, и их использованию для управления экспрессией генов в развивающихся семенах растений и применению последних.

Предпосылки изобретения

Изучение экспрессии генов растений позволило сделать ряд общих выводов относительно элементов, которые контролируют экспрессию. Растения, как и другие организмы, как прокариоты, так и эукариоты, содержат консервативные или консенсусные последовательности в 5'-направлении от сайта инициации транскрипции гена, который, видимо, способен регулировать скорости транскрипции. В эукариотах эти последовательности включают фрагмент, обычно составляющий 25 п.н., в 5'-направлении от участка инициации транскрипции, который имеет последовательность ТАТАА/ТААТ, так что его называют ТАТА-блок.

Роль этого ТАТА-блока, видимо, заключается в определении начала транскрипции для РНК-полимеразы II. Вторую последовательность в обратном направлении называют СААТ-блок. Он обычно содержит около 75 п.н. в обратном направлении от участка инициации транскрипции и связан с регулированием частоты инициации транскрипций. У растений консенсусные последовательности могут иметь вид ССААТ или иногда АГГА. Однако, не обязательно любая из этих альтернативных консенсусных последовательностей должна присутствовать во всех генах растений. Указанные фрагменты последовательностей и связанный с ними участок ДНК, содержащий 7-90 нуклеотидов в обратном направлении от участка инициации транскрипции, часто называют промотором. В общем случае в 5'-направлении от участка промотора, а наиболее часто в пределах приблизительно 2000 нуклеотидов от него, располагаются цис-активные элементы, которые варьируют свойства промоторов и способны модулировать транскрипционную активность как конститутивным, так и неконститутивным образом. Указанные цис-активные последовательности можно называть усилителями (если они отвечают за ускорение транскрипции) или глушителями (если они отвечают за подавление транскрипции). Усилители и глушители часто являются сайтами, к которым присоединяются или с которыми взаимодействуют белки ядер. Модулирующие белки ядер называют транс-активными факторами. Они, как считается, являются очень важными для неконститутивных или регулируемых экспрессии, и могут оказаться решающим фактором активности гена в конкретных тканях или органах или же они могут откликаться на внешнее стимулирующее воздействие. Связь между присоединением указанных белков и элементом усилитель/глушитель может определять транскрипционную активность. Выделение генов, которые активируются теплом, светом или действием химических соединений, таких как эндогенные гормоны, или активируются в специфических органах, таких как семена, листья или цветки, позволяет провести анализ факторов, которые помогают выявить, каким образом регулируется экспрессия. Во многих, но не во всех, случаях было показано, что конструирование химерных генов, которые содержат промоторы и необязательно цис-элементы, из данного регулируемого гена и кодирующей последовательности репортерного белка, не связанного, как правило, с указанным промотором, приводит к регулируемой экспрессии репортерного гена. Были предприняты лишь отдельные попытки использовать промоторы из семяспецифичных генов при экспрессии последовательностей генов в семени, которые либо обычно не экспрессируются семяспецифичным образом либо требуют применения измененной схемы экспрессии. Во всех случаях в химерных генах, разработанных для семяспецифичной экспрессии, используют регуляторные сигналы и про-моторы для запасных (Storage) белков семян. Однако, из работ по экспрессии генов запасных белков стало очевидно, что экспрессия начинается на сравнительно поздней стадии эмбрионального развития, а именно как только эмбрион достигает (для двудольных растений) классической формы жгутика. Таким образом, хотя запасные белки экспрессируются на высоких уровнях и их регулирование часто является транскрипционным, время и уровень экспрессии могут не оказаться идеальными для применений, зависящих от типа семян. По этой причине представляет интерес выявление других семяспецифичных промоторов и усилителей, обладающих временной или клеточной специфичностью, отличной от запасных белков семян, таких как белки олеозина.

Аналоги

Далее приводятся органо- или тканеспецифичные регуляторные последовательности, которые используют для органо- или тканеспецифичной экспрессии в трансформированных растениях. В настоящее время известно несколько ставших "классическими" примеров регулируемой экспрессии гена белка хлоралфемиколацетилтрансферазы, которая может быть экспрессирована зависимым от света и органспецифичным образом в трансгенных растениях, если кодирующую последовательность репортерного белка сливают с про-мотором и предшествующей последовательностью гена гороха, кодирующего рибулоза-бифос-фат-карбоксилазу (Fluhr, Science, 1986, Vol. 232, pp. 1106-1112) Sengupta-Gopalan et al., Proc. Natl. Acad. Sci. USA, 1985, Vol. 82, pp. 3320-3324 сообщают об экспрессии основного запасного белка французского гороха, называемого бета-фазаеолином, в растениях табака. Ген экспрессируется именно в семенах и лишь незначительно в других частях растения. Однако, конструкции, которую применили. Sengupta-Gopalan et al., не были химерными. Был использован полный ген бета-фазаеолина, включающий нативную 5'-фланкирующую последовательность. Последующие эксперименты с другими образцами (Radke et al., Theor. App. Genet., 1988. Vol. 75, pp. 685-694) или другими генами (L. Perez-Grau, R.B. Goldberd, Plant Cell, Vol. 1, pp. 1095-1109) показали, что экспрессия является семяспецифичной как в случае Arabidosis, так и Brassica. Radke et al. (1988, см. выше) использовали "меченный" ген, т.е. полный ген напина плюс не транскрибируемую "метку".

В случае орган- и тканеспецифичной экспрессии можно привести ряд примеров, которые показывают, что последовательности, расположенные в обратном направлении от участка инициации транскрипции,

можно использовать для придания орган/тканеспецифичности гену, введенному в растения методами генной инженерии. Примеры включают регулирование методами генной инженерии семяспецифичных генов (Radke et al., 1988. см. выше); Bustos et al., *Plant Cell*, 1989, Vol. 1, pp. 839-853). В обоих примерах последовательности в 3'-5'-направлении от кодирующих последовательностей белков семян присоединяют к репортерной метке (либо в виде РНК либо в виде белка) и семяспецифичность подтверждают экспрессией репортерного гена. Во всех случаях это были гены запасных белков, отличные от олеозина.

Фрагменты ДНК, которые могут приводить к семяспецифичной экспрессии, являются в настоящее время объектом многих исследований. Marcotte et al. (W.R. Marcotte, I.S. Russel, R.S. Quantrano, *Plant Cell*, 1989, Vol. 1, pp. 969-976) изучали Em-ген пшеницы и предложили два фрагмента, названные "Em-блоками", которые могут играть роль консенсусных последовательностей для семяспецифичной экспрессии. Интересно отметить, что один из указанных блоков, получивший название EM-2, аналогичен блоку, найденному в других генах однодольных (тритицин из пшеницы) и даже двудольных (бета-конглианин из сои). Hatzopoulos et al. *Plant Cell*, 1990, Vol. 2, p. 457-467) исследовали последовательности, управляющие эмбрионспецифичной экспрессией гена белка липидных тел моркови. Был выявлен ряд фрагментов, содержащих большое число звеньев АТ, которые защищены от расщепления во время обработки ДНКазой преимущественно с помощью транс-активных белков. Однако, не было показано, что выделенные фрагменты являются консенсусными фрагментами для других семяспецифичных генов.

"Dellercq et al, *Plant Physiol*, 1990. Vol. 94, pp. 970-979 использовали промотор 2S альбумина растения *Arabidopsis* и комбинированные кодирующие последовательности 2S альбуминов *Arabidopsis* и бразильского ореха. Слияние осуществляли по участкам с низкой консервативностью. Трансформации как табака, так и *Brassica napus* приводит к семяспецифичной экспрессии и точной аккумуляции модифицированных запасных белков. Уровни экспрессии составляли от 0,05% до 0,3% от общего клеточного белка.

Другим примером этой формы семяспецифичной экспрессии чужеродной последовательности является экспрессия лей-энкефалинов в семенах. Для получения семяспецифичной экспрессии химерную последовательность ДНК, кодирующую 2S альбумин, и короткий олигонуклеотид, кодирующий лей-энкефалин (пентапептид), вводят в кодирующую альбумин последовательность между 6-м и 7-м цистеинами нативного белка (Vanderker-hove et al., *Bio/Technology*, 1989, Vol. 7, pp. 929-932). Указанный ген вновь экспрессирует семяспецифичным образом, позволяя накапливать до 50 пмол лей-энкефалина на каждый грамм семян.

Геномные клоны, кодирующие белки масличных тел, со связанными с ними участками, расположенными в обратном направлении, описаны для двух образцов, кукурузы (*Zea mays*, Bowman-Vance and Huang, *J.Biol.Chem.*, 1987, Vol. 262, pp. 11275-11279; и Qu and Huang, *J. Biol. Chem.* 1990, Vol. 265, pp. 2238-2243) и моркови (Hatzopoulos et al., *Plant Cell*, 1990, Vol. 2, pp. 357-467). КДНК и геномные клоны также приводятся для одного культивируемого масличного растения, *Brassica napus* (Murphy et al., *Biochem. Biophys. Acta*, 1991, Vol. 1088, pp. 86-94, и Lee and Huang, *Plant Physiol*, 1991, Vol. 96, pp. 1395-1397). Появились и другие сообщения об экспрессии этих генов белков масличных тел в развивающихся семенах. В случае *Zea mays* транскрипция генов, кодирующих изоформы белков масличных тел, начинается уже на самых ранних стадиях развития семян и легко обнаруживается на 18-й день после опыления. В нездоспелых семенах, таких как семена двудольного растения *Brassica napus* (Canoi) экспрессия генов белков масличных тел, видимо, начинается на гораздо более поздних, чем у кукурузы, стадиях развития семян (Murphy et al., *Bio-chem. J*, 1989, Vol. 258, pp. 285-293).

Краткое описание изобретения

Описываются способы и композиции для использования регуляторной последовательности транскрипции белков масличных тел и необязательно соседствующей с ней в 5'-направлении не транслируемой лидерной последовательности для семяспецифичной экспрессии гетерологичного гена. Способ включает стадии трансформации клетки растения с помощью ДНК конструкции, включающей регуляторную последовательность и ДНК последовательность, отличную от нативной открытой рамки считывания, по отношению к регуляторной последовательности, генерацию растения из трансформированной клетки и выращивание его в условиях, при которых возможно образование семян и осуществляется экспрессия последовательности ДНК под транскрипционным контролем регуляторного участка. Указанные последовательности представляют ценность в тех случаях, когда необходимо модифицировать, усилить или подавить экспрессию продукта семени. Они также могут использоваться для получения модифицированных семян, содержащих чужеродные белки, с целью усиления качества семян.

Краткое описание фигур.

На фиг. 1 приведена схематичная диаграмма вектора pPAW4, включающего регуляторную последовательность олеозина, иницирующий кодон, чужеродную ДНК, которая должна экспрессироваться, терминаторную последовательность олеозина и ген устойчивости к ампициллину.

На фиг. 2А показана последовательность ДНК геномного клона *Arabidopsis*, кодирующего 18КДа белок масличного тела. Открытая рамка считывания прерывается коротким интроном (помечен), а два экзона транскрипруются и обозначаются однобуквенными символами, принятыми Международным союзом теоретической и прикладной химии (ИЮПАК) для аминокислот.

На фиг. 2В приведен рестрикционный фрагмент из геномной библиотеки *Agrobacterium* EMBL3, который включает кодирующую последовательность 18 кДа белка масличного тела. Выделено приблизительно местоположение кодирующего участка.

На фиг. 3 показаны воздействия 19 мкМ абсцизовой кислоты (ABA) на эволюционную экспрессию мРНК олеозина методом Нозерн-блот анализа всей РНК. А) (70 мкг на дорожку) с использованием в качестве зонда 50 нг ОВ990, меченного с помощью 32 р дЦТФ (с удельной активностью ДНК 10^9 распадов в минуту). Зародыши на стадии ядра (Н) (13-ый день), жгута (Т) (17-ый день) и семядоли (С) (21-25 день), выращенные из микроспор и обработанные (+) и необработанные (-) 10 мкМ абсцизовой кислоты в течение 48 час. Пятно экспонируют на пленку Kodak XAR5 при 70°С в течение 20 минут. Кажущиеся различия в размере мРНК на разных дорожках вызваны влиянием различного количества крахмала в разных препаратах мРНК. На все дорожки наносят одинаковое количество соединений, что определяют по измерению содержания ОД260 и по окрашиванию с помощью бромид этидия. В) Экспонирование фиг. 3-А в течение 4,5 час. С) Относительная интенсивность накопления мРНК, определяемая методом сканирующей денситометрии.

На фиг. 4 показана тканевая специфичность олеозина. 50 мкг поли(А)+РНК из корней (R), каллуса (Ca), семядоли (Co), листьев (L) и зиготных эмбрионов (E) через 24 дня после цветения исследуют действием 50 нг ОВ990, меченного 32 РгЦТФ (с удельной активностью ДНК 10^8 распадов в минуту).

На фиг. 5 показана эволюционная чувствительность синтеза белков масличных тел от применения абсцизовой кислоты. Парные контрольные образцы (дорожки А, С, Е, G) и образцы, обработанные с помощью абсцизовой кислоты (дорожки В, D, F, H) помещают в ячейку с приблизительной удельной активностью 10000 распадов в минуту на ячейку. Все образцы обрабатывают в течение 2 дней абсцизовой кислотой, а затем помещают в течение 4 час с помощью [32 S]-мети-онина с активностью 1,85 МБк/мл. Дорожки А и В - 10-дневные культуры, просеянные на сите с размером ячеек 62 мкм, с целью выделения глобулярных эмбрионов. Дорожки С и D - 13-дневные культуры, просеянные на сите с размером ячеек 125 мкм, с целью выделения эмбрионов на стадии образования ядра. Дорожки Е и F - 17-дневные культуры, просеянные на сите с размером ячеек 250 мкм, с целью выделения эмбрионов от стадии образования жгута до стадии образования семядоли. Дорожки G и H - 25-дневные культуры, просеянные на сите с размером ячеек 500 мкм, с целью выделения эмбриона на стадии образования семядоли.

Краткое описание конкретных вариантов осуществления изобретения.

В соответствии с настоящим изобретением, заявляются ДНК конструкции, которые позволяют осуществить модуляцию фенотипа растения в семенах, в частности во время ранних стадий эмбрионального развития. Заявляются ДНК конструкции для регулирования транскрипции в семенах с использованием не транслированных последовательностей генов в 5'-направлении, которые активны в течение развития от поздней глобулярной стадии до созревания семян (стадия семядоли). В прямом направлении под участком инициации транскрипции гена белка масличных тел должны располагаться нужная ДНК последовательность, которая будет получена и которая позволяет интегрировать транскрипционную кассету в геном клетки растения. Таким образом, в прямом направлении от участка инициации семяспецифичной транскрипции может быть включен множественный клонирующий сайт, так что для разнообразных ДНК последовательностей можно эффективно использовать интеграционную конструкцию.

Особый интерес представляет регуляторная последовательность гена белков масличных тел, преимущественно ген белка масличных тел, экспрессируемый в масличных семенах двудольных растений. Сообщается, что белки масличных тел, накапливаются значительно позднее чем другие масла (триглицериды) или запасные белки. Эта поздняя экспрессия ограничивает ценность любых промоторов, связанных с указанными генами для семяспецифической экспрессии, поскольку их нельзя использовать модификации экспрессии генов на ранних стадиях эмбриогенеза. Однако, неожиданно было обнаружено, что экспрессия указанных генов в семенах двудольных масличных культур наблюдается значительно ранее, чем считалось до настоящего времени. Таким образом, промоторы и элементы указанных генов, расположенные в обратном направлении, представляют ценность для широкого применения, включая модификацию метаболизма на различных фазах эмбриогенеза, которые предшествуют накоплению запасных белков.

Белки масличных тел идентифицированы у широкого круга таксономически различных видов (см., например Moreau et al., *Plant Physiol.*, 1980, Vol. 65, pp. 1176-1180, Qu et al., *Biochem J.*, 1986, Vol. 235, pp. 57-65). Указанные белки локализованы лишь в масличных телах и не обнаруживаются в органеллах тканей овощей. В *Brassica napus* (рапс) имеются, по крайней мере, три полипептида, связанные с масличными телами развивающихся семян (Taylor et al., 1990, *Planta*, Vol. 181, pp. 18-26). Количества и размеры белков, связанных с масличными телами, могут изменяться от образца к образцу. В кукурузе, например, имеются четыре иммунологически различных полипептида, которые встречаются в масличных телах (Bowman-Vance and Huang, *J. Biol. Chem.*, 1988, Vol. 263, pp. 1476-1481). Было показано, что олеозины включают участки с чередующейся гидрофильностью, гидрофобностью, гидрофильностью (Bowman-Vance and Huang, *J. Biol. Chem.*, 1987, Vol. 262, pp. 11275-11279). Были получены последовательности аминокислот олеозинов, выделенных из кукурузы, рапса и моркови. См., соответственно, Qu and Huang, *J. Biol. Chem.*, 1990, Vol. 265, pp. 2238-2243, Hatzopoulos et al., *Plant Cell*, 1990, Vol. 2, pp. 457-467. В масличных растениях, таких как рапс, содержание олеозина может составлять от 8% (Taylor et al., *Planta*, 1990, Vol. 181, pp. 18-26) до 20% (Murphy et al., *Biochem. J.*, 1989, Vol. 258, pp. 285-293) от общего количества белка в семенах. Этот уровень сравним с количествами, найденными для других запасных белков.

Особый интерес представляет участок инициации транскрипции, связанный с ранними стадиями эмбриогенеза, в частности с периодом, предшествующим экспрессии запасных белков, так что на ранних стадиях развития семян он обеспечивает необходимый уровень транскрипции представляющей интерес ДНК последовательности. Обычный процесс эмбриогенеза растения проходит ряд определенных фаз. Для

двудольных семян процесс эмбрионального развития включает следующие фазы: глобулярную фазу, фазу ядра, фазу жгутика и фазу семядоли. В настоящем описании используются определения указанных терминов, взятые из книги Ray, Steves and Fultz, Botany (Saunders College Publishing), Глава 17, стр. 294. Обычно участок инициации транскрипции можно получить из гена, который экспрессируется на ранних стадиях образования семени. Участок инициации транскрипции предпочтительно сохраняет свою активность от поздней глобулярной стадии вплоть до стадии образования семядоли, а еще более предпочтительно сохраняет активность от глобулярной стадии через стадии образования ядра, жгутика и семядоли, процесса эмбрионального развития. Под этим участком инициации транскрипции понимают участок инициации транскрипции, имеющий последовательность нуклеотидов, в значительной степени схожую с последовательностью участка инициации транскрипции гена белка природного масличного тела, который был бы способен обеспечивать транскрипцию на ранних стадиях формирования семени. Эта последовательность может быть природной, синтетической или частично синтетической.

Участок инициации транскрипции белка масличного тела обычно входит в кассету, которая включает в 5'-3'-направлении транскрипции участок инициации, транскрипции, представляющую интерес ДНК последовательность и участок терминации транскрипции, при этом регуляторные участки транскрипции присоединяются операбельно и способны функционировать в клетках растений. Может присутствовать также один или большее количество интронов. После каждой манипуляции ДНК, которая будет использоваться в конечной конструкции, может быть ограничена и операбельно присоединена к другой ДНК, которая будет использоваться в конечной конструкции, где каждая из частичных конструкций может быть клонирована на той же или разных плаزمиде. В предпочтительном способе осуществления изобретения кодирующая последовательность с совместимым сайтом рестрикции может быть лигирована в положение, соответствующее кодону I гена белка масличного тела. Схематичная диаграмма такого замещения представлена на фиг. 1. Рекombинантная кодирующая последовательность может быть встроена таким образом, что она полностью замещает кодирующую последовательность гена белка масличного тела и, таким образом, фланкирована на ее 3'-конце терминатором гена белка масличного тела и сигналом полиаденилирования. По другому способу можно осуществить амплификацию с помощью цепной полимеразной реакции (ЦРП), с целью получения ДНК фрагментов, содержащих участок инициации транскрипции, который удобно соседствует с сайтами рестрикции. Скопированные фрагменты могут быть присоединены к последовательности кодирующей нужный полипептид путем транскрипционного или трансляционного "слияния", например, с целью получения химерного гена, в котором кодирующая последовательность нужного полипептида транскрибируется под контролем участка инициации транскрипции на фрагменте, амплифицированном с помощью ЦРП.

Участок инициации транскрипции может быть нативным или гомологичным по отношению к клетке-хозяину или чужеродным или гетерологичным по отношению к клетке-хозяину. Под чужеродным понимают, что участок инициации транскрипции не встречается в клетке-хозяине дикого типа, в которую помещают конструкцию, содержащую участок инициации транскрипции. Как правило, регуляторная последовательность включает ДНК размером до 15000 нуклеотидов в 5'-направлении от сайта инициации транскрипции гена белка масличного тела. Эту последовательность можно модифицировать в положении, соответствующем первому кодону представляющего интерес белка, путем сайт-направленного мутагенеза (T.A. Kukul, Proc. Natl. Acad. Sci. USA, 1985, Vol. 82, pp. 488-492) или введения удобного линкерного олигонуклеотида путем лигирования, если рядом с П-концевым кодоном находится подходящий сайт рестрикции.

В некоторых случаях желательно осуществить экспрессию нужной ДНК последовательности в виде белка слияния, в частности, в виде белка слияния с белком масличного тела. Нужную ДНК последовательность можно ввести по обычной методике в последовательность, кодирующую белок масличного тела (в рамке с последовательностью, кодирующей белок масличного тела), так что транскрипция химерного гена приведет к получению белка слияния. Белок слияния преимущественно содержит кодирующий участок для аминокислот с номерами от 44 до 122 в белке масличного тела *Arabidopsis*, как показано на фиг. 2А, или эквивалентный участок из белка масличного тела видов, отличных от *Arabidopsis* с целью обеспечить транспорт масличного тела тогда, когда это желательно.

Чтобы отделить последовательности, кодирующие белки масличного тела от других видов, можно пойти по крайней мере двумя путями. Первый заключается в использовании клона *Arabidopsis*, описанного в примерах, в качестве зонда в геномной библиотеке других растительных видов. Этот клон легко гибридизуется с клонами олеозина из близко родственных видов, в частности, практически всех крестоцветных растений. Для образцов, которые эволюционно отличны от *Arabidopsis*, например, пасленовых, бобовых и всех однодольных, альтернативный метод заключается в использовании антитела, вырабатываемого против генного продукта клона олеозина, такого как клон *Arabidopsis*. Это антитело можно применить для скрининга произведенной из семян кДНК – экспрессирующей библиотеки, например, используя лямбда gt11; Huynh et al., cDNA Cloning, Vol. 1, A Practical Approach (Ed. Grover), 1985, IRL Press, pp. 49-78. Этим путем можно получить клон кДНК олеозина для новых видов, который можно использовать для выделения геномного клона из геномной библиотеки указанных видов стандартными методами гибридизации ДНК.

Представляющие интерес последовательности ДНК могут являться любой открытой рамкой считывания, кодирующей нужный пептид, например, фермент, или последовательностью, комплементарной последовательности генома, при этом геномная последовательность может являться одной из открытых рамок считывания интроном, некодирующей лидерной последовательностью или любой другой последовательностью, в которой комплементарная последовательность подавляет транскрипцию,

процессинг матричной РНК, например, сплайсинг или трансляцию. Представляющая интерес последовательность ДНК может быть синтетической, выделяемой из природного источника или их комбинацией. В зависимости от природы представляющей интерес последовательности ДНК может быть желательно синтезировать последовательность с предпочтительными кодонами растения. Предпочтительные кодоны растения можно определить по кодонам с наибольшей частотой в белках, экспрессированных в наибольшем количестве в конкретных видах растений, представляющих интерес.

Нужная последовательность ДНК может кодировать любое разнообразие рекомбинантных белков. Примеры рекомбинантных белков, которые могут экспрессироваться указанным способом, включают антикоагулянты, такие как гирудин, лимфокины, такие как лимфокины семейства интерлейкинов, пептидные гормоны, такие как гонадотропин-выделяющие гормоны, иммунологические реагенты, такие как одноцепочечные антитела и множество имеющих промышленное значение ферментов, таких как протеазы, липазы и полиглюкангидролазы.

Используемым участком терминации является в первую очередь наиболее удобный, поскольку участки терминации являются относительно взаимозаменяемыми. Участок терминации может быть нативным по отношению к нужной последовательности ДНК или же может быть получен из другого источника. Удобные участки терминации доступны и включают 3'-конец терминатора гена белка масляного тела и сигнала полиаденилирования того же самого гена, из которого получен 5'-регуляторный участок. По другому способу с тем же успехом можно использовать отличный терминатор и сигнал полиаденилирования, например, терминатор гена нопалинсинтетазы *Agrobacterium*.

Кассета экспрессии может дополнительно включать средства для идентификации трансформированных клеток и/или селекции трансформированных клеток. Например, рекомбинантный ген может быть связан с конститутивно экспрессируемым селективным маркером, таким как ген, придающий свойство биолюминесценции или окрашиваемости трансформированным клеткам.

Нужная последовательность ДНК, фланкированная по ее 5'-концу промотором белка масляного тела и регуляторной последовательностью, и фланкированная терминатором по ее 3'-концу, может быть введена в подходящий вектор трансформации, включая Ti- или бинарную плазмиду *Agrobacterium* или простую плазмиду для клонирования (в частности, pC19, pBP322) для прямого введения ДНК в клетки растений методами микроинъекции, электропорации, опосредованного полиэтиленгликолем поглощения или баллистического метода. Эти способы хорошо известны специалистам в области трансформации растений. См., например, Horsch et al., *Science*, 1985, Vol. 227, pp. 1229-1231, Newhaus and Spangenberg, *Physiol. Plant*, 1990, Vol. 79, pp. 213-217, и Sandford et al., *Physiol. Plant*, 1990, Vol. 79, pp. 206-209.

Трансформированные растения можно получить из трансформированных клеток, используя стандартные методики выращивания (см., например, Moloney et al., *Plant Cell Rep.*, 1989, Vol. 8, pp. 238-242), совместимые с методами трансформации.

Кассета экспрессии, которую конструируют как указано ранее, в основном предпочтительно экспрессируется в развивающихся семенах. Поэтому клетки растения, которые трансформировались с помощью подходящего белка слияния, выращивают в виде растения в соответствии с известными способами и дают им возможность сформировать семена. См., например, McCormick et al., *Plant Cell Rep.*, 1986, Vol. 5, pp. 81-84. Можно вырастить два или больше поколений и опылить их тем же трансформированным штаммом или другими штаммами, определяя полученный гибрид, обладающий нужными фенотипическими характеристиками, с тем, чтобы убедиться, что изучаемые фенотипические характеристики стабильно сохраняются и наследуются, а затем собирают семена для выделения нужных белков или для использования в качестве посевного материала с новыми фенотипическими свойствами. Затем регенерированные растения культивируют в идентичных условиях с нерекомбинантными растениями в ростовых камерах, парниках или на поле и выявляют семяспецифичную экспрессию рекомбинантного гена на уровне мРНК, а часто на уровне полипептида или белка.

Полипептид/белок может представлять ценность сам по себе и его можно экстрагировать, и если потребуется, подвергнуть дальнейшей очистке. Иначе полипептид/белок или же саму мРНК можно использовать для введения нового биохимического фенотипа в развивающиеся семена. Новые фенотипы могут включать такие модификации, как изменение состава белка или состава масел семян, увеличение воспроизводства уже существующих нужных веществ или пептидов и ослабление и даже подавление образования нежелательных генных продуктов с использованием десенсибилизации, рибозимной технологии или совместной супрессии (Tzant and Weintraub, *Cell*, 1984, Vol. 36, pp. 1007-1015, десенсибилизация, Hazelhoff and Gerlach, *Nature*, 1988, Vol. 334, pp. 585-591, рибозимная технология, Napoli et al., *Plant Cell*, 1990, Vol. 2, pp. 279-289, совместная супрессия).

Если трансформацию проводят с целью получения новых белков или пептидов семян, для выделения которых необходима экстракция, то ее осуществляют, используя водные растворы, содержащие небольшие количества поверхностно-активных веществ, таких как не приводящие к денатурации количества додецилсульфата натрия, Triton X-100, Tween 20, MEGA-8 или любого другого поверхностно-активного вещества, которое не приводит к необратимой дезактивации нужного белка. Для того, чтобы провести экстракцию белка или полипептида сухие семена измельчают вручную или в механическом гомогенизаторе для получения водной суспензии или взвеси. Полученную смесь можно разделить на три фазы (отдельные частицы, растворимая в воде фаза, гидрофобная фаза) путем центрифугирования, например, с ускорением 50000 x g. В зависимости от типа продукта его можно затем очистить в составе каждой из указанных фаз и селективно осадить после солюбилизации действием сульфата аммония или очистить с помощью колоночной хроматографии, например, используя ионный обмен, гель-фильтрацию или аффинную матрицу.

Хотя идеальным хозяином описываемой здесь регуляторной последовательности могло бы быть растение семейства крестоцветных, можно использовать указанные промоторы для широкого круга образцов растений с относительно высокой степенью консервации олеозина в генах. Главным препятствием на пути использования указанных промоторов являются однодольные и двудольные растения. Для трансформаций, включающих указанную специфическую экспрессию в однодольных, необходимо использовать регуляторную последовательность олеозина однодольных. Для семейства двудольных необходимо использовать регуляторную последовательность олеозина двудольных. Рассматриваемая последовательность может использоваться для широкого круга двудольных растений, в том числе всех членов вида *Brassica* и крестоцветных в целом. Пасленовые, такие как табак и помидоры, также распознают эту последовательность и демонстрируют корректное регулирование экспрессии в развивающихся семенах.

Можно ожидать, что нужные белки будут экспрессироваться во всех тканях эмбриона, несмотря на то, что в различных тканях жилки эмбриона и семядолях можно обнаружить различные клеточные экспрессии. Настоящее изобретение имеет многочисленные применения, которые включают повышение качества семян растений за счет накопления ими измененных полипептидов или новых рекомбинантных полипептидов или путем включения или исключения стадий метаболизма. В наиболее простом варианте его осуществления использование настоящего изобретения может привести к повышению качества белка (например, к повышенной концентрации ванных или редких аминокислот), улучшенному качеству жидкости за счет модификации состава жирных кислот, или улучшенному составу углеводов или повышению их количества. Примеры включают экспрессию богатых серой белков, таких как белки, содержащиеся в люпинах или бразильских орехах, в семенах, дефицитных по серусодержащим остаткам аминокислот. Аналогично может экспрессироваться жирный ацил кофермент А - фермент семейства трансфераз, способный модифицировать относительные содержания жирных кислот в триглицеридах (запасных липидах). Если рекомбинантному белку позволяют накапливаться в семенах, белок также может представлять собой пептид, имеющий ценность для фармацевтического и технического применения или для использования в пище. В этом случае пептид можно экстрагировать из семян и использовать в сырой или очищенной форме в зависимости от предполагаемого применения. Пептид может быть действительно чужеродным пептидом растительному миру, например, животным гормоном, ферментом, лимфокином, антикоагулянтом и т.п., который способен экспрессироваться в семенах. Гетерологичные белки затем можно экстрагировать из семян и после частичной или полной очистки использовать в исследовательских целях, в пищевой или фармацевтической промышленности.

Следующие примеры приводятся для иллюстрации, но не ограничивают настоящее изобретение.

Примеры

Пример 1

Ген белка масличного тела из *Arabidopsis* изолируют для получения вставки, размером 15 тыс. нуклеотидов, присутствующей в клоне из геномной библиотеки. *Arabidopsis thaliana* v. Columbia в фаге EMBL3A путем гибридизации с клоном олеозина *B. napus*. Фрагмент, имеющий 1800 нуклеотидов и содержащий приблизительно 868 пар нуклеотидов в 3'-5'-направлении от участка инициации трансляции, субклонируют в плазмидном векторе. Ген олеозина *Arabidopsis* размером 18 кДа удобно клонировать в виде фрагмента величиной 1803 п.н., фланкированного сайтами Pco1 и Kpn1 в векторе, называемом pPAW4 (см. фиг. 1). Для превращения фрагмента в кассету экспрессии, с целью использования для широкого круга чужеродных/альтернативных генов, необходимо осуществить две модификации. Вначале, применяя методику сайт-специфичного мутагенеза Kunkel, см. ранее) вводят мутации в положениях -2, -1 и +4 с использованием неправильно спаренных олигонуклеотидов. Требуемые мутации заключаются в замене А на Т (-2), А на С (-1) и Г на А (+4). Результатом указанных мутаций является образование сайта Bsp H1 в положении от -2 до +4. Сайт Bsp H1 (T/CATGA) включает кодон инициации ATG и образует рецессивный конец, совместимый с Pco1. Вторая модификация включает расщепление с помощью EcoRV и Msc1, в результате которого выделяется фрагмент размером 658 п.н., содержащий большую часть кодирующей последовательности нативного олеозина. После указанной операции остаются тупые края на сайтах разрыва, которые при отделении вектора и вспомогательной последовательности из фрагмента EcoRV – Msc1 обеспечивают рециркуляцию комбинации вектор-промотор-терминатор. Эту рециркуляцию осуществляют в присутствии олигонуклеотидного линкера, содержащего сайт рестрикции, который отсутствует в оригинальном фрагменте размером 1803 тыс. нуклеотидов.

В процессе рециркуляции получают плазмиду, содержащую все предшествующие последовательности гена олеозина, сайт инициации транскрипции и иницирующий кодон, встроенный в сайт BspH1. Его фрагмент в прямом направлении, содержащий 31 основание, представляет собой короткий полилинкер, содержащий одно или несколько уникальных сайтов рестрикции. Для введения какой-либо последовательности ДНК в эту кассету, чужеродная последовательность должна иметь или должна быть соответствующим образом модифицирована, чтобы содержать сайт Bsp H1 или Msc1 в исходной ATG позиции. Если последовательность будет экспрессироваться в виде белка, это обеспечивает сохранение расстояния между "cap"-сайтом и иницирующим кодоном.

Последовательность ДНК, которую необходимо ввести, должна заканчиваться когезивным концом сайта рестрикции, который отсутствует в плазмиде. Этот сайт следует учитывать при выборе полилинкера, который вводят в кассету экспрессии. Расщепление плазмиды с помощью BspH1 и подходящего фермента рестрикции для 3'-конца чужеродной последовательности обеспечивает направленное клонирование нужного фрагмента ДНК. Используя подходящие условия лигирования, выдерживают совместно кассету

экспрессии плазмиды с Bsp HI и сайт, совместимый с нужным фрагментом ДНК, и получают лигированный продукт, как это показано на фиг. 1.

Законченную конструкцию из Pco1-Krp1 затем отсекают и вводят в подходящий вектор трансформации растения, такой как плазида *Agrobacterium*. Для введения конструкции в обычные плазмиды *Agrobacterium*, такие как Bin 19 (Bevan, Nucl. Acid Reseach, 1984, Vol. 12, pp. 8711-8721) может потребоваться использование дополнительного сайта рестрикции в плазмиде pPAW4. По одной из методик плазмиду можно разрезать с помощью Sma1 и Krp1. Полученный очищенный фрагмент лигируют к Krp1 олигонуклеотидному и расщепляют с помощью Krp1. В результате получают ненаправленный Krp1 фрагмент для введения в Bin 19. Иным способом конструкцию можно вырезать с помощью Krp1 и Bam HI и датировать непосредственно в Bin 19, предварительно разрезанную теми же самыми ферментами рестрикции. Полученную бинарную плазмиду *Agrobacterium* мобилизуют в штамм *Agrobacterium* путем трехстороннего спаривания (Ditta et al., PNAS, 1980, Vol. 77, pp. 7347-7351) или ДНК трансформацией компетентных *Agrobacterium*. (An, Plant Mo 1. Biology Manual, 1988, Vol. A3, pp. 1-19, Kluwer Academic, Dordrecht, Netherlands).

Agrobacterium, которая несет рекомбинантную Bin 19, используют для трансформации любого способного воспринимать изменения растения, в частности, *Brassica* sp., путем стандартного культивирования вне растения (Horsch et al., см. ранее). Трансформированные клетки подвергают селекции в культуре с помощью канамицина, используя введенный совместно ген устойчивости (неомицин фосфотрансфераза), который также содержится между границами T-ДНК pBin. Указанные трансформированные клетки вносят для выращивания целого растения, используя стандартные процедуры (в частности, масличного растения, такого как рапс, см. Moloney et al., Plant Cell Rep., 1989, Vol. 8, pp. 238-242). Регенерированным растениям дают зацвести и подвергают самоопылению (или перекрестному опылению). В тех случаях, когда чужеродная ДНК в конструкции кодирует способный транслироваться продукт, его можно выделить из водных экстрактов созревших семян с последующим фракционированием суспензии на центрифуге (30 минут с ускорением 100000 x g). В зависимости от желаемого продукта он может распределиться в одной из трех полученных фаз. Он может находиться в таблетке, растворимой в воде фазе и липидной пленке на поверхности подвергнутого центрифугированию образца.

Воли целью экспрессии является изменение метаболизма в семени и изменение, таким образом фенотипа семян (в частности, путем изменения размера или цвета семян, изменения соотношения между остатками жирных кислот в семенах или запрещения определенной стадии метаболизма, которая, как считается, приводит к менее ценным семенам) экстракцию продукта проводить не обязательно. Указанные стадии обмена веществ могут включать производство непригодных для употребления в пищу вторичных продуктов, присутствие которых снижает ценность семян или являются нежелательными. В таких случаях семена просто собирают и используют в соответствии с принятой практикой.

Пример 2

Методом ПЦР получают ряд конструкций, содержащих различные количества последовательностей ДНК в 3'-5'-направлении от участка инициации транскрипции гена олеозина *Arabidopsis*, операбельно присоединенных к кодирующему участку бета-глюкуронидазы (GUS). Конструкции обозначают в соответствии с величиной 5'-участка олеозина, так, например, конструкция 2500 содержит приблизительно 2500 п.н. 5'-участка олеозина. Конструкции вводят в *Brassica napus* и табак и оценивают экспрессию гена бета-глюкуронидазы, как это подробно описано далее. Результаты экспрессии GUS пяти конструкций - конструкции 2500, 1200, 800, 600 и 200 - в трансформированных растениях *Brassica napus* представлены в табл. 1. Приведены также данные для негативного контрольного образца (нетрансформированного растения). Результаты экспрессии GUS двух конструкций - конструкции 2500 и конструкции 800 - в трансформированных растениях табака представлены в табл. 2. В табл. 3 приведены данные об развитии экспрессии промотора олеозина в трансгенных эмбрионах.

Конструкции получают, используя стандартные методики молекулярной биологии, включая расщепление ферментом рестрикции, лигирование и полимеразную цепную реакцию (ПЦР). Для пояснения использованной методики подробно описывается конструкция 800.

Для получения фрагмента ДНК, содержащего приблизительно 800 п.н. в 5'-направлении от участка инициации транскрипции гена олеозина *Arabidopsis* с конфигурацией, удобной для связывания с кодирующей последовательностью GUS, используют методику, основанную на ПЦР. Она включает применение полимеразной цепной реакции для амплификации точных последовательностей, которые необходимы для анализа экспрессии. Для проведения необходимой ПЦР амплификации синтезируют два олигонуклеотидных праймера (Milligen-Bioscience, Cyclon DNA synthesizer), содержащих следующие последовательности:

Pst. 1 последовательность олеозина:

5'-праймер: 5' CAGTGCAGGAACCTCTGGTAA 3' (GVR 10).

Основания, помеченные шрифтом "италик", соответствуют нуклеотидам в позициях от -833 до -817 в последовательности, приведенной на фиг. 2A. Дополнительные нуклеотиды в 5'-направлении этой последовательности не идентичны в праймере гену олеозина, однако их включают с целью размещения сайта Pst 1 У 5'-конца продукта амплификации. Сайт Pst не выделен.

Синтезируют второй (3') праймер, который имеет следующую последовательность:

3' праймер (APL 1)

Bam HI последовательность олеозина:

5-CTACCCGGGATCCTGTTTACTAGAGAGAATC-3

Sma1.

Этот праймер содержит точный комплемент (выделен шрифтом "италик") последовательности, приведенной на фиг. 2А от оснований -13 до -30. Далее, он включает еще 13 оснований у 5'-конца. Указанная последовательность не является комплементарной гену олеозина, однако введена для получения двух (перекрывающихся) сайтов рестрикции, Sma1 и Bam H1, у 3'-конца продукта амплификации, с целью облегчения клонирования ПЦР фрагмента.

Эти два праймера используют в ПЦР реакции амплификации для получения фрагмента ДНК, содержащего последовательность между нуклеотидами -833 и -13 гена олеозина с сайтом Pst1 на 5'-конце и сайтами Sma1 и Bam H1 на 3'-конце. Амплификацию с помощью ПЦР осуществляют, используя фермент Таг-полимеразу (Perkin-Elmer-Cetus) в условиях, рекомендованных производителем фермента по следующей температурной программе: 92°C (денатурация) 1 минута, 55°C (восстановление двойной спирали) 1 минута и 72°C (элонгация) 1 минута. В качестве матрицы используют геномный клон олеозина, приведенный в верхней части фиг. 2В, который в оригинальной библиотеке веществ, выделенных из природного сырья, содержал приблизительно 15000 нуклеотидов ДНК из *Arabidopsis*.

Продукт амплификации (OLEO p800) очищают в геле 0,7% агарозы, выделяют, используя метод стеклянных шариков (Vogelstein and Gillespie, Pre-parative and analytical purification of DNA from agarose – Proc. Natl. Acad. Sci. USA, 1979, Vol. 76, pp. 615-619) и расщепляют Pst1. Продукт расщепления очищают на геле и дополняют конец, используя фрагмент Кленова ДНК-полимеразы, а затем разрезают Sma1, получая фрагмент с тупым краем. Его клонируют на Sma1 сайте pUC19 и получают плазмиду pUC OLEOp800. Применяя асимметричное положение сайта Acc1 во вставке (в положении, соответствующем -649 гена олеозина, приведенного на фиг. 2В), можно селективировать обе ориентации вставки в векторе pUC. Клон имеющий вставку, ориентированную таким образом, что крайний 5'-конец амплифицированного фрагмента (в направлении транскрипции) расположен ближе к единственному сайту Hind III в векторе клонирования pUC19, а крайний 3'-конец амплифицированного фрагмента расположен ближе к единственному сайту Eco R1 в векторе pUC19.

Полученную плазмиду затем разрезают с помощью Bam H1 и выделяют фрагмент OLEOp800, фланкированный сайтами Bam H1. Указанный фрагмент, Bam H1-OLEOp800, клонируют в сайты Bam H1 расщепленной с помощью Bam H1 плазмиды, которую обозначают Hsp GUS 1559. Hsp GUS 1559 является плазмидой, которая используется в качестве бинарного вектора в *Agrobacterium* и образуется из вектора pCGN 1559 (MacBride and Summerfeldt, Plant Molecular Biology, 1990, No 14, pp. 269-276) путем введения вставки, содержащей промотор генов теплового шока (фланкированный сайтами Bam H1), открытую рамку считывания бета-глюкуронидазы и терминатор нопалинсинтетазы (получаемый из pB1221, R.A. Jefferson, в книге "Cloning Vectors" 1988, Eds. P. Pouwels., B.E. Enger-Valk, W.J. Brammer, Elsevier Science Pub BV, Amsterdam section VII, A111). Расщепление HspGUS 1559 с помощью Bam-H1 приводит к высвобождению промотора генов теплового шока и позволяет осуществить инсерцию любого другого фрагмента Bam-H1. В это место лигируют фрагмент. Bam-H1-OLEOp800 и получают *Agrobacterium* pOLEOp800 CUS 1559. Указанную плазмиду используют для трансформации *E.coli* и амплифицированную плазмиду вводят *Agrobacterium* (штамм EHA101) методом электропорации, как описано ранее (Rogers et al., Plant Molecular Biology Manual, 1988, Vol. A2, pp. 1-12, Eds. S. Gelvin and R. Schilperoort, Kluwer Academic, Dordrecht, Netherlands).

Полученный штамм *Agrobacterium* (EHA101x pOLEOp800 GUS 1559) используют для трансформации растений *Brassica napus* по методу Молони и др. (M.M. Moloney, J.M. Walker, K.K. Sharma, Plant Cell Reports, 1989, Vol. 8, pp. 238-242) или растений табака по методу Хорша и др. (Horsch et al., Science, 1985, Vol. 227, pp. 1299-1302). Образовавшимся трансгенным растениям дают возможность образовать семена и проводят анализ экспрессии GUS (R.A. Jefferson, Plant Mol. Biol. Rep., 1987, No 5, pp. 387-405) в развивающихся семенах, а также непродуктивных частях растений, используемых в качестве контрольных образцов. Приведенные величины экспрессии являются средними значениями, которые получают в среднем из пяти семян от каждого из приблизительно пяти различных трансгенных растений.

Другие конструкции получают аналогичным методом ПЦР, указанным ранее, используя подходящие праймеры для амплификации -2500 фрагмента, -1200 фрагмента, -600 фрагмента или -200 фрагмента. Результаты для *Brassica napus*, выраженные в виде удельной активности фермента GUS, представлены в табл. 1. данные для табака приведены в табл. 2.

Полученные результаты показывают, что олеозинный фрагмент от -833 до -813, который используют в конструкции 800, содержит достаточно информации для прямой семяспецифичной экспрессии репортерного гена в трансгенном эмбрионе *Brassica napus*, начиная уже со стадии ядра, а промотор олеозина *Arabidopsis* способен направлять транскрипцию в растениях, отличных от *Arabidopsis*.

Эти эксперименты также показывают, что последовательности, содержащиеся в конструкции промотора, содержат цис-элементы, необходимые для усиления транскрипции в ответ на добавление абсцизовой кислоты, что является характерным для промотора нативного олеозина.

Следует отметить, что семяспецифичная экспрессия, которая продемонстрирована в настоящем описании, не зависит от взаимодействий с нативными терминаторами у 3'-конца гена олеозина. В данном примере 3'-концевой терминатор олеозина заменен терминатором, выделенным из гена нопалинсинтетазы *Agrobacterium*. Таким образом, конструкция 800 достаточна для осуществления необходимого профиля экспрессии независимо от вспомогательных последовательностей.

Семяспецифичная экспрессия в *Brassica napus*
Активность GUS (в пмол продукта/мг белка)

Промотор/GUS конструкция	Семя (стадия жгутика)		Корень	Лист	Стебель	Семя (поздняя стадия семядоли)
	+ABA*	-ABA				
2500	10,185	7,709	444	46,9	88,2	11,607
1200	18,298	1,795				8,980
800	2,250	475	285	277	650	7,130
600	1,506	144				1,365
200	18,1	64,8	260	5,9	26	11
Негативный контроль – не-трансформированное растение	18,4	13,9	300	6,1	30	14

*ABA обозначает обработку в течение 24 ч. 10^{-5} М абсцизовой кислоты перед проведением измерений активности GUS.

Таблица 2

Семяспецифичная экспрессия в *Тобacco*
Активность GUS (в пмол продукта/мг белка)

Промотор/ конструкция	Зрелые семена
2500	11,300
800	10,970

Таблица 3

Эволюционная экспрессия в *Brassica napus*
Активность GUS (в пмол продукта/мг белка)

Промотор/GUS конструкция	Стадия ядра	Стадия жгутик	Ранняя стадия семядоли	Средняя стадия семядоли	Поздняя стадия семядоли
2500	272	1207	2541	1819	11,607
1200	124	262	388	5094	8,980
800	149	260	962	2617	7,128
600	59	41	29	38	1,365

Продолжение табл. 3

Эволюционная экспрессия в *Brassica napus*
Активность GUS (в пмол продукта/мг белка)

Промотор/GUS конструкция	Стадия ядра	Стадия жгутик	Ранняя стадия семядоли	Средняя стадия семядоли	Поздняя стадия семядоли
200	30	25	15	20	II
Негативный контроль – отрицательный образец			II	14	14

Все публикации и патентные заявки, на которые даются ссылки в настоящем описании, приведены здесь для справки в той мере, в какой каждая индивидуальная публикация или патентная заявка была специально и индивидуально указана в качестве ссылки.

После того, как изобретение полностью описано, для специалиста становится очевидным, что в него могут быть внесены многие изменения и модификации, которые не приводят к изменению сущности и изобретения и объема притязаний, изложенного в прилагаемой формуле изобретения.

Описание последовательностей

(2) Информация для последовательности с номером идентификации: I

(i) Характеристики последовательности:

(A) Длина: 1803

(B) Тип: нуклеиновая кислота

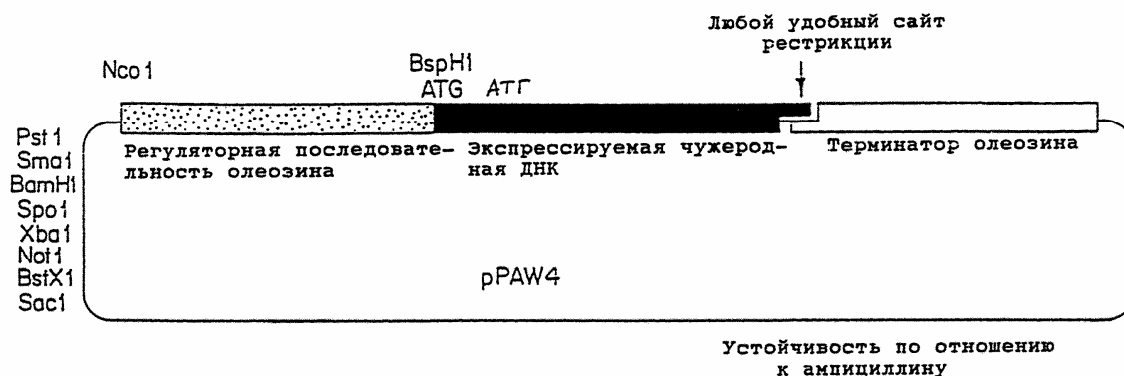
(C) Вид цепи: двойная

(D) Топология: линейная

(xi) Описание последовательности:

последовательность с номером идентификации I:

(x1) SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
CCATGGCTAT ACCCAACCTC GGTCTTGGTC ACACCAGGAA CTCTCTGGTA AGCTAGCTCC	60
ACTCCCCAGA AACAAACGGC GCCAAATTGC CGGAATTGCT GACCTGAAGA CGAACATCA	120
TCGTGGGGTC CTTGGGGGAT TGCGGCGGAA GATGGGTCAG CTTGGGCTTG AGGACGAGAC	180
CCGAATCGAG TCTGTTGAAA GGTGTTTCAT TGGGATTGTG ATACGGAGAT TGGTCGTCGA	240
GAGGTTTGAG GGAAGGACA AATGGGTTTG GCTCTGGAGA AAGAGACTGC GGCTTTAGAG	300
AGAGAATTGA GAGGTTTGA GAGAGATGCG CCGGCGATGA CGGGAGGAGA GACGACGAGG	360
ACCTGCATTA TCAAAGCAGT GACGTGGTGA AATTGGAAC TTITAAGAGG CAGATAGATT	420
TATTATTGTG ATCCATTTC TTCAATTGTC TAGAATGTCG CGGAACAAAT TTTAAACTA	480
AATCCTAAAT TTTCTAATT TTGTTCCAA TAGTGGATAT GTGGGCCGTA TACAAGGAAT	540
CTATTGAAGG CCCAAACCCA TACTGACGAG CCCAAAGGTT CGTTTTGCGT TTTATGTTTC	600
GGTTTCGATGC CAACGCCACA TTCTGAGCTA GGCAAAAAAC AACGTGTCT TTGAATAGAC	660
TCCTCTCGTT AACACATGCA GCGGCTGCAT GGTGACGCCA TTAACACGTG GCCTACAATT	720
GCATGATGTC TCCATTGACA CGTGACTTCT CGTCTCCTTT CTTAATATAT CTAACAAACA	780
CTCCTACCTC TTCCAAAATA TATACACATC TTTTGATCA ATCTCTCATT CAAAATCTCA	840
TTCTCTCTAG TAAACAAGAA CAAAAA ATG GCG GAT ACA GCT AGA GGA ACC CAT	894
Met Ala Asp Thr Ala Arg Gly Thr His	
1 5	
CAC GAT ATC ATC GGC AGA GAC CAG TAC CCG ATG ATG GGC CGA GAC CGA	942
His Asp Ile Ile Gly Arg Asp Gln Tyr Pro Met Met Gly Arg Asp Arg	
10 15 20 25	
GAC CAG TAC CAG ATG TCC GGA CGA GGA TCT GAC TAC TCC AAG TCT AGG	990
Asp Gln Tyr Gln Met Ser Gly Arg Gly Ser Asp Tyr Ser Lys Ser Arg	
30 35 40	
CAG ATT GCT AAA GCT GCA ACT GCT GTC ACA GCT GGT GGT TCC CTC CTT	1038
Gln Ile Ala Lys Ala Ala Thr Ala Val Thr Ala Gly Gly Ser Lys Lys	
45 50 55	
GTT CTC TCC AGC CTT ACC CTT GTT GGA ACT GTC ATA GCT TTG ACT GTT	1086
Val Lys Ser Ser Lys Thr Lys Val Gly Thr Val Ile Ala Lys Thr Val	
60 65 70	
GCA ACA CCT CTG CTC GTT ATC TTC AGC CCA ATC CTT CTC CCG GCT CTC	1134
Ala Thr Pro Lys Lys Val Ile Phe Ser Pro Ile Lys Val Pro Ala Lys	
75 80 85	
ATC ACA GTT GCA CTC CTC ATC ACC GGT TTT CTT TCC TCT GGA GGG TTT	1182
Ile Thr Val Ala Lys Lys Ile Thr Gly Phe Lys Ser Ser Gly Gly Phe	
90 95 100 105	
GGC ATT GCC GCT ATA ACC GTT TTC TCT TGG ATT TAC AAG TAAGCACACA	1231
Gly Ile Ala Ala Ile Thr Val Phe Ser Trp Ile Tyr Lys	
110 115	
TTTATCATCT TACTTCATAA TTTTGTGCAA TATGTGCATG CATGTGTTGA CCCAGTAGCT	1291
TTGGATCAAT TTTTTTGGTC GAATAACAAA TGTAACAATA AGAAATTGCA AATTCTAGGG	1351
AACATTGGT TAACTAAATA CGAAATTGGA CCTAGCTAGC TTGAATGTGT CTGTGTATAT	1411
CATCTATATA GGTAAATGC TTGGTATGAT ACCTATTGAT TGTGAATAGG TAC GCA ACG	1470
Tyr Ala Thr	
1	
GGA GAG CAC CCA CAG GGA TCA GAC AAG TTG GAC AGT GCA AGG ATG AAG	1518
Gly Glu His Pro Gln Gly Ser Asp Lys Leu Asp Ser Ala Arg Met Lys	
5 10 15	
TTG GGA AGC AAA GCT CAG GAT CTG AAA GAC AGA GCT CAG TAC TAC GGA	1566
Leu Gly Ser Lys Ala Gln Asp Leu Lys Asp Arg Ala Gln Tyr Tyr Gly	
20 25 30 35	
CAG CAA CAT ACT GGT TGG CAA CAT GAC CGT GAC CGT ACT CGT GGT GGC	1614
Gln Gln His Tyr Gln Gln Glu His Asp Arg Asp Arg Thr Arg Gly Gly	
40 45 50	
CAG CAC ACT ACT TAAGTTACCC CACTGATGTC ATCGTCATAG TCCAATAACT	1666
Gln His Tyr Tyr	
55	
CCAATGTCGG GGAGTTAGTT TATGAGGAAT AAGTGTTTA GAATTTGATC AGGGGGAGAT	1726
AATAAAGCC GAGTTTGAAT CTTTTTGTTA TAAGTAATGT TTATGTGTGT TTCTATATGT	1786
TGTCAAATGG TACCGAG	1803



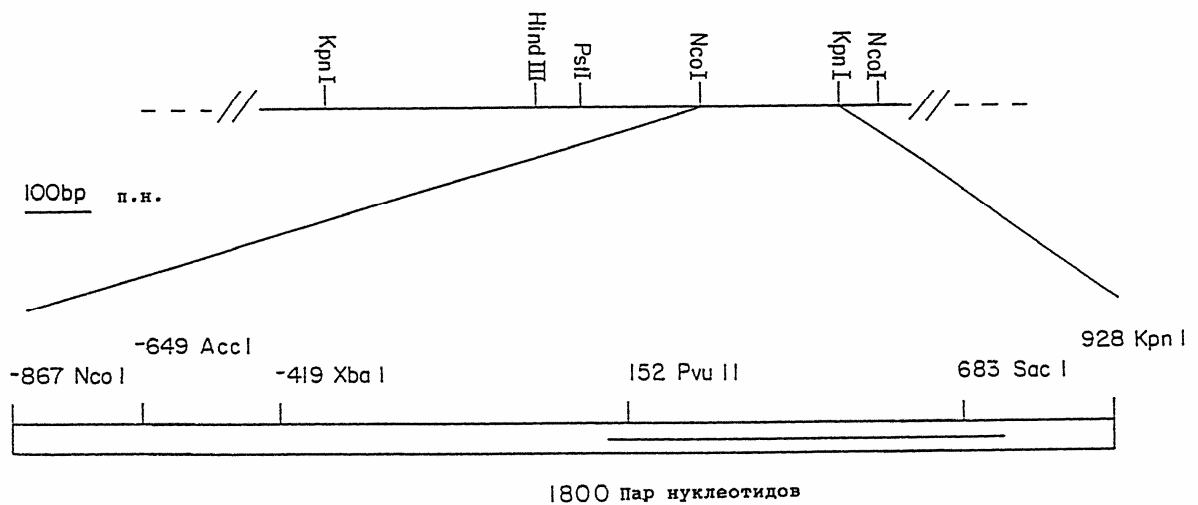
Фиг. 1

-867 ^{NcoI} CCATGGCTATACCCAACCTCGGTCTTGGTCACACCAGGAAC³¹CTCTGGTAAGCTAGCTCCACTGCCAGAAACAACCGGCGCCAAATTGC
 -777 CGGAATTGCTGACCTGAAGACGGAACATCATCGTCGGGTCCTTGGGCGATTGCGGCGGAAGATGGGTCAGCTTGGGCTTGAGGACGAGAC
 -687 CCGAATCGAGTCTGTTGAAAGGTTGTTCAATTGGGATTTGTATACGGAGATTGGTCGTCGAGAGGTTTGAGGGAAAGGACAAATGGGTTTG
 -597 GCTCTGGAGAAAGAGAGTGC^{R2}GGCTTTAGAGAGAGAATTGAGAGGTTTAGAGAGAGATGCGGCGGCGATGACGGGAGGAGAGACGACGAGG
 -507 ACCTGCATTATCAAAGCAGTGACGTGGTGAAATTTGGAAC^{R1}TTTTAAGAGGCAGATAGATTTATTATTTGTATCCATTTTCTTCATTGTTT
 -417 TAGAATGTCGCGGAACAAATTTTAAACTAAATCCTAAATTTTCTAATTTTGTGCGCAATAGTGGATATGTGGGCCGTATAGAAGGAAT
 -327 CTATTGAAGGCCCAAACCCATACTGACGAGCCCAAGGTTCTGTTTGCCTTTTATGTTTCGGTTCGATGCCAACGCCACATTCTGAGCTA
 -237 GGCAAAAAACAACGTGTCTTTGAATAGACTCCTCTCGTTA^{R2}ACACATG^{R1}CAGCGGCTGCATGGTGACGCCATTAACACGTGGCCTACAATT
 -147 GCATGATGTCTCCATTGACACGTGACTTCTCGTCTCCTTTCTTAATATATCTAACAACACTCCTACCTCTTCCAAAATATATACACATC
 -57 TTTTGTATCAATCTCTCATTCAAATCTCATTCTCTCTAGTAAACAAGAACAATAATGGCGGATACAGCTAGAGGAACCCATCAGAT
 34 I I G R D Q Y P M M G R D R D Q Y Q M S G R G S D Y S K S R
 ATCATCGGCAGAGACCACTACCCGATGATGGGCCGAGACCGAGACCACTACAGATGTCCGACGAGGATCTGACTACTCCAAGTCTAGG

Фиг. 2А-1

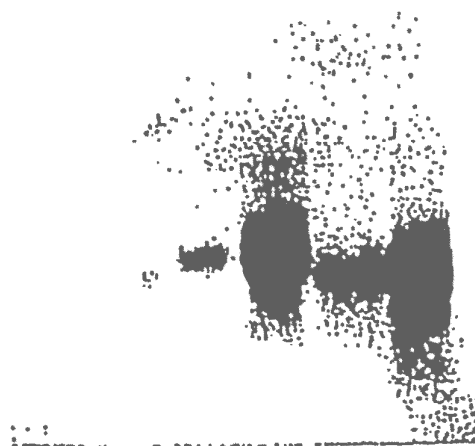
Q I A K A A T A V T A G G S L L V L S S L T L V G T V I A L
 124 CAGATTGCTAAAGCTGCAACTGCTGTCACAGCTGGTGGTTCCTCTCTCCAGCCTTACCCTTGTGGAACTGTCATAGCTTTG
 T V A T P L L V I F S P I L V P A L I T V A L L I T G F L S
 214 ACTGTTGCAACACCTCTGCTCGTTATCTTCAGCCCAATCCTTGTCCCGGCTCTCATCAGTTGCACTCCTCATCACCGGTTTCTTTCC
 S G G F G I A A I T V F S W I Y K
 304 TCTGGAGGGTTTGGCATTGCCGCTATAACCGTTTTCTTGGATTTACAAGtaagcācācātttatcatcttacttcataatttgtgca
 394 atatgtgātgātgātggttgagccagtagctttggatcaattttttggtcgaataacaaatgtaacaataagaaattgcaaattctagg
 484 gaacatttggttaactaaatagcaaatggaccagctagcttgatgtgtctgtgtatatcatctatataggtaaaaatgcttggtatga
 Y A T G E H P Q G S D K L D S A R M K L G S K
 574 tacctattgattgtgaatagGTACGCAACGGGAGAGCACCCACAGGGATCAGACAAGTTGGACAGTGCAAGGATGAAGTTGGGAAGCAAA
 A Q D L K D R A Q Y Y G Q Q H T G G E H D R D R T R G G Q H
 664 GCTCAGGATCTGAAAGACAGAGCTCAGTACTACGGACAGCAACATACTGGTTGGGAACATGACCGTGACCGTACTCGTGGTGGCCAGCAC
 T T *
 754 ACTACTTAAGTTACCCCACTGATGTCATCGTCATAGTCCAATAACTCCAATGTCGGGGAGTTAGTTTATGAGGAATAAAAGTGTTTAGAAT
 844 TTGATCAGGGGGAGATAAATAAGCCGAGTTTGAATCTTTTTGTTATAAGTAATGTTTATGTGTGTTTCTATATGTTGTCAAATGGTACC
 KpnI

Фиг. 2А-2



Фиг. 2В

H		T		C	
-	+	-	+	-	+

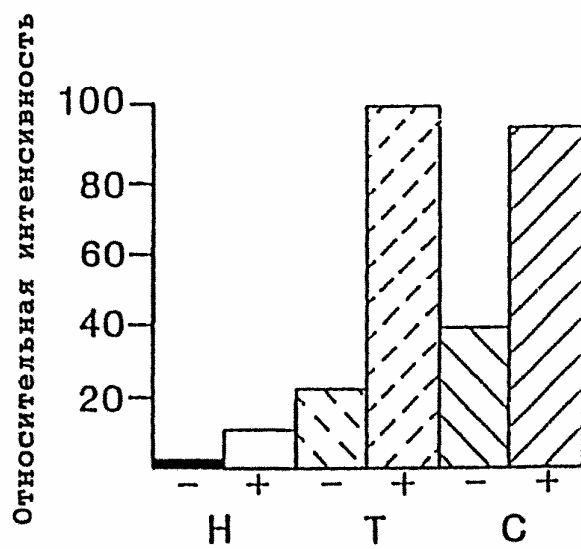


Фиг. 3А

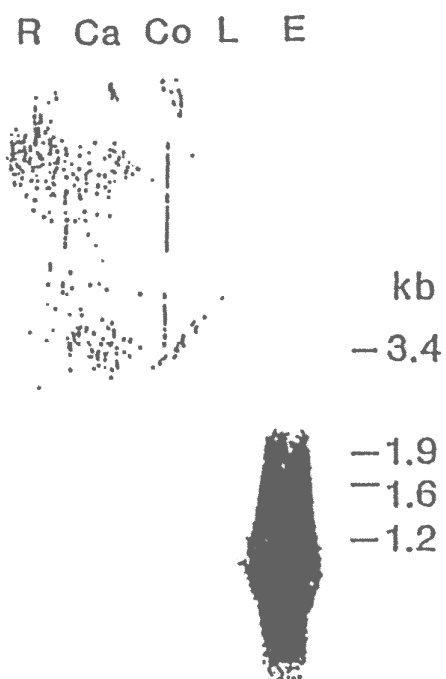
H		T		C	
-	+	-	+	-	+



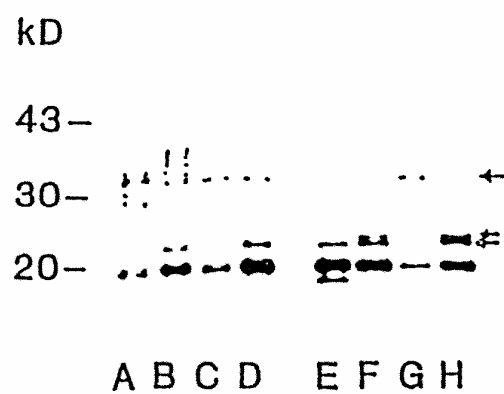
Фиг. 3В



Фиг. 3С



Фиг. 4



Фиг. 5