



(19) **RU**⁽¹¹⁾ **2 201 449**⁽¹³⁾ **C2**
(51) МПК⁷ **C 12 N 15/29, 15/62, 15/82, C**
07 K 19/00

РОССИЙСКОЕ АГЕНТСТВО
ПО ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ РОССИЙСКОЙ ФЕДЕРАЦИИ

(21), (22) Заявка: 94046181/13, 15.04.1992
(24) Дата начала действия патента: 15.04.1992
(43) Дата публикации заявки: 10.11.1996
(46) Дата публикации: 27.03.2003
(56) Ссылки: EMBL Nucleic Acid Sequence Database Accession №X 62353, 20.02.1992. VAN ROOIJEN G.J.H., Plant molecular Biology, v.18, №6, 1992.
(85) Дата перевода заявки РСТ на национальную фазу: 14.10.1994
(86) Заявка РСТ: CA 92/00161 (15.04.1992)
(87) Публикация РСТ: WO 93/21320 (28.10.1993)
(98) Адрес для переписки: 129010, Москва, ул. Б. Спасская, 25, стр.3, ООО "Юридическая фирма Городисский и Партнеры", Н.Г. Лебедевой, рег.№ 112

(71) Заявитель: СЕМБАЙОЗИС ДЖЕНЕТИКС ИНК. (CA)
(72) Изобретатель: МОЛОНИ Морис (GB)
(73) Патентообладатель: СЕМБАЙОЗИС ДЖЕНЕТИКС ИНК. (CA)
(74) Патентный поверенный: Лебедева Наталья Георгиевна

(54) СЛИТЫЙ ПОЛИПЕПТИД, СПОСОБНЫЙ К ЦЕЛЕНАПРАВЛЕННОМУ ПЕРЕНОСУ К МАСЛЯНОМУ Телу, ХИМЕРНАЯ ДНК-КОНСТРУКЦИЯ, ЭКСПРЕССИРУЮЩАЯ КАССЕТА

(57) Изобретение относится к генетической инженерии и может быть использовано для получения представляющего интерес белка в семенах растения. Предложена химерная ДНК-конструкция, включающая последовательность, кодирующую часть белка масляного тела - олеозина, которая определена как достаточная для обеспечения переноса к масляному телу, и последовательность, кодирующую предназначенный для синтеза в растении чужеродный белок, между которыми дополнительно может быть встроена последовательность, кодирующая сайт

протеазного расщепления. Путем помещения указанной ДНК-конструкции под контроль промотора гена, в норме экспрессирующегося в семени растения, получена кассета экспрессии, которая использована для трансформации растительных клеток. Семена выразившего трансгенного растения использованы для получения представляющего интерес белка, синтезированного в виде белка слияния с олеозином. Накопление экспрессированного белка во фракции масляных тел существенно облегчает его выделение и очистку. 3 с. и 12 з.п.ф-лы, 7 ил., 2 табл.

RU 2 201 449 C2

RU 2 201 449 C2



(19) **RU** ⁽¹¹⁾ **2 201 449** ⁽¹³⁾ **C2**
 (51) Int. Cl.⁷ **C 12 N 15/29, 15/62, 15/82, C**
07 K 19/00

RUSSIAN AGENCY
 FOR PATENTS AND TRADEMARKS

(12) **ABSTRACT OF INVENTION**

(21), (22) Application: **94046181/13**, 15.04.1992
 (24) Effective date for property rights: 15.04.1992
 (43) Application published: 10.11.1996
 (46) Date of publication: 27.03.2003
 (85) Commencement of national phase: 14.10.1994
 (86) PCT application:
CA 92/00161 (15.04.1992)
 (87) PCT publication:
WO 93/21320 (28.10.1993)
 (98) Mail address:
129010, Moskva, ul. B. Spasskaja, 25, str.3,
OOO "Juridicheskaja firma Gorodisskij i
Partnery", N.G. Lebedevoj, reg.No 112

(71) Applicant:
SEMBAJOZIS DZhENETIKS INK. (CA)
 (72) Inventor: **MOLONI Moris (GB)**
 (73) Proprietor:
SEMBAJOZIS DZhENETIKS INK. (CA)
 (74) Representative:
Lebedeva Natal'ja Georgievna

(54) **FUSED POLYPEPTIDE ABLE TO PURPOSEFUL TRANSFER TO OILY BODY, CHIMERIC DNA-CONSTRUCTION, EXPRESSING CASSETTE**

(57) Abstract:
 FIELD: genetic engineering. SUBSTANCE:
 invention proposes chimerical
 DNA-construction comprising sequence
 encoding part of oily body protein oleosine
 that is determined as sufficient for
 providing transfer to oily body and sequence
 encoding foreign protein designated for
 synthesis in plant. Between these components
 the sequence encoding site of protease
 cleavage can be fused. By placing indicated
 DNA-construction under control of gene
 promoter expressing in plant seed under

normal conditions the expression cassette is
 obtained that is used for transformation of
 vegetable cells. Seeds of grown transgenic
 plant are used for preparing protein
 presenting the interest synthesized as
 fusion protein with oleosine. Accumulation
 of the expressed protein in fraction of oily
 bodies alleviates its isolation and
 purification significantly. Invention can be
 used for preparing protein in plant seeds
 presenting the interest. EFFECT: valuable
 biological properties of polypeptide. 15 cl,
 8 dwg, 2 tbl

RU 2 201 449 C2

RU 2 201 449 C2

Область изобретения

Настоящее изобретение относится к рекомбинантному способу получения представляющих интерес протеинов, которые легко выделять из компонентов клеток хозяев. Способ иллюстрируется экспрессией представляющего интерес протеина в растениях, особенно в семенах в форме химерного пептида, содержащего протеин масляного тела и представляющий интерес протеин.

Предпосылки изобретения

В растениях были экспрессированы различные протеины. Однако хотя и была продемонстрирована общая доступность осуществления экспрессии чужеродных протеинов в растениях, получение очищенных протеинов из таких источников имеет некоторые ограничения. Эти ограничения включают стадию очистки, необходимую для получения чистого протеина, практически не содержащего растительных материалов и продуктов разложения, которые могут образовываться в экстрактах, полученных в процессе очистки, когда полученные рекомбинантные протеины контактируют с водными буферами.

Растения, имеющие маслячные семена, такие как соя, рапс, подсолнечник и ряд других видов растений, таких как кукуруза, морковь и т.д., хранят в своих семенах триглицериды. В растениях эти триглицериды выполняют роль источника энергии для прорастания семян и последующего роста сеянца. Триглицериды широко используют в качестве растительных масел в пище и пищевых продуктах, а также в некоторых областях промышленности.

Триглицериды не смешиваются с водой и выделяются за счет флотации на поверхности водных растворов или образуют небольшие глобулы или липосомы в виде суспензии в водной фазе. Такие глобулы естественно сливаются, если они не стабилизированы модифицированным поверхностным слоем. Такое слияние может привести к образованию суспензии из глобул случайного размера. В семенах же при сохранении триглицеридов масляные глобулы на деле представляют собой капсулированные липидные или масляные тела обычно одинакового размера. С поверхностью этих масляных тел ассоциированы полуфрагменты (half unit) мембран, усыпанные несколькими протеинами, которые обычно называют протеинами масляного тела.

По крайней мере, один класс протеинов масляного тела имеет несколько характеристик, которые существенно консервативны от вида к виду. Этот класс протеинов масляного тела называют "олеозинами". Гидрофильные N- и C- концы этих протеинов, по-видимому, совершенно различны, тогда как липофильные внутренние участки (центральное ядро), по-видимому, строго консервативны для различных видов. Олеозины прочно связаны с масляными телами, эта прочная связь с масляными телами может быть, в основном, связана с липофильным характером этих центральных ядер. Поэтому представляет интерес определение возможности использования таких протеинов масляного тела, как олеозины, для создания средств для выделения рекомбинантных протеинов из

растительных материалов.

Известный уровень техники

Продуцирование чужеродных

(рекомбинантных) пептидов в растениях было исследовано с использованием различных подходов, включая транскрипционные слияния с применением сильного конститутивного растительного промотора (например, из вируса мозаики цветной капусты - Sijmons et al. (1990) *Bio/Technology*, 8: 217-221) и кодирование чужеродного протеина; транскрипционные слияния с органо-специфическими последовательностями (Radke et al. (1988) *Theoret. Appl. Genet.*, 75:685-694); и трансляционные слияния, которые требуют последующего отщепления рекомбинантного протеина (Vander Kerkove et al. (1989) *Bio/Technology*, 7:929-932). Чужеродные протеины, которые были экспрессированы в растительных клетках, включают активные протеины из бактерий (Fraleley et al. (1983) *Proc. Nat'l. Acad. Sci. USA*, 80:4803-4807), животных (Misra and Gedamu (1989) *Theor. Appl. Genet.*, 78:161-168), грибов и других видов растений (Fraleley et al. (1983) *Proc. Nat'l. Acad. Sci. USA*, 80: 4803-4807).

Некоторые протеины, обычно маркеры интеграции, были экспрессированы ткане-специфическим образом, причем некоторые были экспрессированы в семенах (Sen Gupta et al. (1985) *Proc. Nat'l. Acad. Sci. USA*, 82:3320-3324); Radke et al. (1988) *Theor. Appl. Genet.*, 75:685-694). Эти сообщения концентрируются специфически на использовании промоторов протеинов хранения семян как средства обеспечения специфической для семян экспрессии. Используя эти системы, Vanderkerkove et al. (1989) *Bio/Technol.*, 7: 929-932, экспрессировал очень ценный пептид (leu энкефалин) в семенах *Arabidopsis thaliana* and *Brassica napus*. Выход этих пептидов чрезвычайно низок, но он демонстрирует возможность экспрессии пептидного гормона животного в ткани растения. Олеозин кукурузы был экспрессирован в масляных телах семян *Brassica napus* трансформированных геном олеозина кукурузы. Этот ген был экспрессирован под контролем регуляторных элементов из гена *Brassica*, кодирующего напин (napin), основной протеин хранения семян. Как сообщается, временная регуляция и тканеспецифичность экспрессии правильны для промотора/терминатора гена напина. См. Lee et al., *Proc. Nat'l. Acad. Sci. USA*, (1991) 88:6181-6185.

Масляные глобулы, которые образуются в семенах, по-видимому, все одинакового размера, что указывает на их стабилизацию (Huang A.H.C (1985) в *Modern Meths. Plant Analysis* v. 1:145-151 Springer-Verlag, Berlin). При более тщательном исследовании было обнаружено, что они представляют собой не просто масляные глобулы, а скорее масляные тела, окруженные мембраной. Такие масляные тела называли по разному электронные микроскописты: олеосомы, липидные тела и сферосомы (Gurr M.I., (1980) в *The Biochemistry of Plants*, 4:205-248, Acad. Press, Orlando, Fla). Были исследованы масляные тела нескольких видов, и пришли к выводу, что они заключены в необычные "полуфрагментарные" мембраны, которые

представляют собой не классический липидный бислой, а скорее один амфобильный слой с гидрофобными группами с внутренней стороны и гидрофильными группами с внешней (Huang A. H. C. (1985) в *Modern Meths. Plant Analysis*, vol. 1:145-151 Springer-Verlag, Berlin).

Анализ содержания липидных тел продемонстрировал, что помимо триглицеридов и материала мембран они содержат также несколько полипептидов/протеинов, связанных с поверхностью или полостью масляного тела (Bowman-Vance and Huang (1987), *J. Biol. Chem.* 262:11275-11279, Murphy et al. (1989) *Biochem. J.*, 258: 285-293, Taylor et al. (1990) *Planta*, 181:18-26). Протеины масляных тел были идентифицированы для широкого круга таксономически различных видов (Moreau et al. (1980) *Plant Physiol.*, 65: 1176-1180; Qu et al. (1986) *Biochem. J.*, 235:57-65) и было показано, что они расположены только в масляных телах, но не были обнаружены в органеллах растительных тканей. В *Brassica napus* (семена рапса) имеются, по крайней мере, три полипептида, связанные с масляными телами в развивающихся семенах Taylor et al. (1990), *Planta*, 181: 18-26). Количество и размеры протеинов, связанных с масляными телами, могут меняться от вида к виду. Так например, в кукурузе существуют два иммунологически различных класса полипептидов, обнаруженных в масляных телах (Bowman-Vance and Huang (1988) *J. Biol. Chem.* 263:1476-1481).

Было показано, что олеозины содержат участки чередующейся гидрофобности и гидрофильности (Bowman-Vance and Huang (1987) *J. Biol. Chem.*, 262:11275-11279). Аминокислотные последовательные олеозины из кукурузы, семян рапса и моркови были получены. См. Qu and Huang (1990) *J. Biol. Chem.* 265:2238-2243, Hatzopoulos et al (1990) *Plant Cell.*, 2:457-467 соответственно. В таких семенах масличных, как рапс, олеозин может составлять от 8% (Taylor et al. (1990) *Planta*, 181: 19-26) до 20% (Murphy et al. (1989) *Biochem. J.*, 258: 285-293) от полного содержания протеина в семенах. Столь высокое содержание сопоставимо с тем, что было обнаружено для многих протеинов хранения семян.

Сообщалось, что получены гены, кодирующие протеины масляных тел для двух видов: кукуруза (*Zea mays*, Bowman-Vance and Huang (1987) *J. Biol. Chem.* 262: 11275-11279; и Qu and Huang (1990), *J. Biol. Chem.*, 265: 2238-2243) и морковь (Hatzopoulos et al. (1990) *Plant Cell.*, 2:457-467).

Краткое содержание изобретения

Предложены способы и композиции для осуществления этих способов получения пептидов, которые можно легко выделить из протеинов хозяев. Этот способ включает стадии получения химерной конструкции ДНК, которая включает последовательность, кодирующую специфическую последовательность масляного тела, содержащую кодирующую последовательность гена протеина масляного тела, специфического для семян, или последовательность, кодирующую, по крайней

мере, участок гидрофобного ядра протеина масляного тела, и последовательность, кодирующую нужный пептид, из которых можно получить кассету экспрессии, содержащую химерную конструкцию ДНК; трансформация клеток хозяина кассетой экспрессии в условиях геномной интеграции; и выращивания полученного трансгенного растения до получения семян, в которых нужный полипептид экспрессирован как протеин слияния с олеозином.

Представляющий интерес полипептид можно очистить, выделяя масляные тела из клеток семян, разрушая масляные тела, так что при этом высвобождается протеин слияния. Затем протеины масляного тела легко отделяются от других протеинов и полученных из растений материалов за счет разделения фаз. При желании сайт расщепления может быть расположен, по крайней мере, перед N-концом или после C-конца представляющего интерес полипептида, что обеспечит возможность отщеплять полипептид слияния и выделять его за счет фазового разделения на составляющие его пептиды. Таким образом, создана система, обеспечивающая целенаправленный перенос химерного пептида за счет функциональности протеина его масляного тела к масляным телам, что в свою очередь позволяет быстро очистить представляющий интерес полипептид. Эта система получения находит применение при получении многих пептидов, например таких пептидов, которые обладают нужными фармацевтическими, энзиматическими, реологическими и адгезивными характеристиками.

В соответствии с настоящим изобретением предложен полипептид слияния, отличающийся тем, что может быть обеспечена его доставка к масляному телу, содержащий:

а) первый пептид формулы (последовательность 1, см. в конце описания), при условии, что указанный первый пептид отличается от полноразмерного природного 16Kb олеозина из моркови, или 18Kb или 16Kb олеозина из кукурузы, где

aa²⁵ представляет любую аминокислоту;

aa²⁶ представляет нейтральную алифатическую аминокислоту;

aa³¹ представляет нейтральную незамещенную алифатическую аминокислоту, состоящую из 3-6 атомов углерода;

aa³³ представляет нейтральную незамещенную алифатическую аминокислоту, содержащую 3-6 атомов углерода;

aa³⁶ представляет нейтральную алифатическую незамещенную аминокислоту, содержащую 3-5 атомов углерода;

aa³⁷ представляет нейтральную незамещенную аминокислоту;

aa³⁹ представляет нейтральную алифатическую незамещенную аминокислоту;

aa⁴¹ представляют нейтральную алифатическую незамещенную или оксизамещенную аминокислоту;

aa⁴⁴ представляет нейтральную алифатическую незамещенную или оксизамещенную аминокислоту;

aa⁴⁶ представляет нейтральную алифатическую незамещенную аминокислоту или оксизамещенную аминокислоту;

aa⁴⁷ представляет нейтральную алифатическую незамещенную аминокислоту;

aa⁵⁹ представляет нейтральную алифатическую или ароматическую незамещенную аминокислоту;

aa⁷⁶ представляет нейтральную алифатическую незамещенную или тиозамещенную аминокислоту;

aa⁷⁸ представляет нейтральную алифатическую незамещенную аминокислоту;

aa⁸³ представляет нейтральную алифатическую незамещенную или оксизамещенную аминокислоту;

aa⁹² представляет нейтральную алифатическую аминокислоту с оксизамещением;

aa⁹⁶ представляет нейтральную алифатическую тиозамещенную аминокислоту или нейтральную ароматическую гетероциклическую аминокислоту;

aa⁹⁷ представляет нейтральную алифатическую незамещенную или тиозамещенную аминокислоту;

aa⁹⁸ представляет нейтральную алифатическую незамещенную аминокислоту или ароматическую оксизамещенную аминокислоту;

aa⁹⁹ представляет любую аминокислоту;

aa¹⁰⁰ представляет оксизамещенную аминокислоту, либо алифатическую, либо ароматическую;

aa¹⁰¹ представляет нейтральную незамещенную алифатическую или ароматическую аминокислоту;

слитый с

b) вторым пептидом, при условии, что этот второй пептид отличается от части природного олеозинового протеина из *Arabidopsis* или *Brassica*.

В настоящем изобретении предложен также полипептид слияния, характеризуемый как способный осуществлять целенаправленный перенос к масляному телу, содержащий:

а) первый пептид, выбранный из группы, состоящей из:

(1) пептида, содержащего, по крайней мере, восемь последовательных аминокислот, включенных в аминокислотную последовательность 2, см. в конце описания.

(2) пептид, который кодируется ДНК последовательностью, идентифицированной с помощью олигонуклеотидного зонда, сконструированного на основе указанной аминокислотной последовательности (1), или ее фрагмента, при условии, что указанный первый пептид отличается от полноразмерного природного 16Kb олеозина из моркови или 18Kb или 16Kb олеозина из кукурузы, слитый с

b) вторым пептидом, при условии, что указанный второй пептид отличается от части природного олеозинового протеина из *Arabidopsis* или *Brassica*.

Далее в настоящем изобретении предложена химерная ДНК конструкции, содержащая:

а) первую ДНК последовательность, кодирующую олеозин или его часть, достаточную для обеспечения целенаправленного переноса к масляному телу, и

b) вторую ДНК последовательность,

кодирующую пептид, при условии, что указанный пептид отличается от части природного олеозинового протеина из *Arabidopsis* или *Brassica*.

5 Далее в настоящем изобретении предложена кассета экспрессии, содержащая: в качестве компонентов в направлении транскрипции:

- регуляторную ДНК последовательность, содержащую достаточную часть участка 5' до сайта начала трансляции гена, экспрессируемого в семенах, для обеспечения экспрессии ДНК последовательности в семенах;

10 - химерную ДНК последовательность, содержащую

а) первую ДНК последовательность, кодирующую олеозин, или достаточную ее часть для обеспечения целенаправленного переноса к масляному телу, причем указанная ДНК последовательность включает, по крайней мере, один рестрикционный сайт, и

20 b) вторую ДНК последовательность, кодирующую пептид, при условии, что указанный пептид отличается от природного олеозинового протеина из *Arabidopsis* или *Brassica*; и

25 - участок окончания трансляции и транскрипции;

где указанные компоненты операбельно связаны, а экспрессия указанной химерной последовательности ДНК регулируется указанной регуляторной ДНК последовательностью.

30 В настоящем изобретении предложена также кассета экспрессии, содержащая:

- ген протеина масляного тела (oil body protein = ОВР), который включает достаточную часть участка 5' до сайта начала трансляции для обеспечения экспрессии

35 указанного гена в клетки семян и который включает, по крайней мере, один рестрикционный сайт между 5'-концом кодона инициации метионина и 5'-сигналом окончания трансляции указанного ОВР гена, и

40 - ДНК последовательность, встроенную в указанный рестрикционный сайт в рамке считывания с указанным ОВР геном, где указанная ДНК последовательность кодирует пептид, который отличается от части нативного олеозинового протеина из

45 *Arabidopsis* или *Brassica*.

В настоящем изобретении предложена также кассета экспрессии, содержащая: первую ДНК последовательность, кодирующую пептид, при условии, что

50 указанный пептид отличается от части природного олеозинового протеина из *Arabidopsis* или *Brassica*, встроенную в рамке считывания в ген протеина масляного тела, которая включает достаточную часть регуляторного участка 5' до сайта начала трансляции указанного ОВР гена для обеспечения экспрессии указанного гена в семенах, где указанная последовательность

55 встроена по такому сайту указанного гена, чтобы быть экспрессированной под контролем указанного регуляторного участка.

60 В настоящем изобретении предложен способ достижения экспрессии целевого пептида в семенах, причем указанный способ включает трансформацию клеток растения-хозяина экспрессионной кассетой в условиях геномной интеграции, где указанная кассета экспрессии содержит в качестве

компонентов в направлении транскрипции первую ДНК последовательность, содержащую достаточную часть участка 5' до сайта начала трансляции гена, экспрессируемого в семенах, для обеспечения экспрессии ДНК последовательности в семенах, вторую ДНК последовательность, кодирующую олеозин или достаточную его часть, для обеспечения нацеленного переноса к масляному телу, причем указанная вторая ДНК последовательность включает, по крайней мере, один природный или синтетический рестрикционный сайт, в который встроена в рамке считывания третья ДНК последовательность, кодирующая целевой пептид, при условии, что указанный пептид отличается от части природного олеозинового протеина из *Arabidopsis* или *Brassica*; и участок окончания трансляции и транскрипции; где указанные компоненты операбельно связаны, и экспрессия указанной второй ДНК последовательности регулируется указанной первой ДНК последовательностью, для обеспечения экспрессии в семенах.

В настоящем изобретении предложен способ получения очищенного целевого пептида, причем указанный способ включает:

трансформацию клеток растения-хозяина ДНК конструкцией в условиях геномной интеграции, где указанная ДНК конструкция содержит первую ДНК последовательность, кодирующую целевой пептид, при условии, что указанный пептид отличается от части природного олеозинового протеина из *Arabidopsis* или *Brassica*, встроенную в считывающую рамку гена протеина масляного тела, которая содержит достаточную часть регуляторного участка 5' до сайта начала трансляции указанного ОВР гена для обеспечения экспрессии указанного гена в семенах, причем указанная последовательность встроена по такому сайту указанного гена, что экспрессия указанной ДНК последовательности контролируется указанным регуляторным участком, за счет чего указанная ДНК конструкция интегрируется в геном указанных клеток растения;

выращивание указанного растения до получения семян, в которых целевой пептид экспрессирован как протеин слияния с продуктом экспрессии указанного ОВР гена;

выделение масляных тел из клеток указанных семян;

разрушение указанных масляных тел, в результате чего выделяется указанный протеин слияния;

и очистку указанного целевого пептида.

Краткое описание чертежей

Фиг. 1А представляет нуклеотидную последовательность и выведенную аминокислотную последовательность (17 кДа протеин) гена протеина масляного тела (олеозина) из *Arabidopsis thaliana*. Подчеркнуты прямые повторы (R_1 и R_2) инвертный повтор (T), CACA, TATA, TAAT и сигналы полиаденилирования.

Интронная последовательность напечатана мелким шрифтом, а предполагаемый АВА-связывающий сайт указан жирным шрифтом.

На фиг.1В представлено сравнение последовательностей протеина 16Кб масляного тела из моркови и 18Кб и 16Кб протеинов масляного тела из кукурузы и 17Кб

протеина масляного тела из *Arabidopsis thaliana*, где показаны сохраняющиеся (консервативные) и различающиеся участки этих протеинов; аминокислотные последовательности выравнены, чтобы показать консервативность последовательности в центральном участке протеинов.

На фиг.2 представлены конструкции, использованные для слияния генов протеинов масляного тела с генами, кодирующими чужеродные пептиды. IA представляет С-терминальное слияние целевого пептида с ОВР; IB представляет N-терминальное слияние целевого пептида с ОВР; II представляет внутреннее слияние целевого пептида с ОВР; и III представляет междимерное трансляционное слияние целевого пептида, заключенного между двумя практически полными последовательностями протеина масляного тела, обеспечивающими его направленную доставку. В верхней части фиг. (А) представлена ДНК конструкция, использованная для трансляционного слияния целевых пептидов с протеинами масляного тела. В нижней части фиг. (В) представлены конфигурации генных продуктов, показанных на верхнем участке трансляции и доставки к масляным телам.

Ключ к фиг. следующий: заштрихованный снизу слева вверх направо прямоугольник представляет ОВР промотор или другой специфический для семян промотор; заштрихованный снизу справа вверх налево прямоугольник представляет последовательность, кодирующую целевой пептид; незаштрихованный прямоугольник представляет последовательность, кодирующую протеин масляного тела или синтетическую последовательность, обеспечивающую направленную доставку, основанную на ОВР консервативных фрагментах; заштрихованные в клеточку прямоугольники представляют терминатор гена, содержащий сигнал полиаденилирования; заштрихованный кружок представляет фрагмент распознавания протеазы; спиральная линия представляет нативный С- или N-конец ОВР.

Фиг. 3 представляет подробную структуру конструкции С-терминального слияния. Изображена структура последовательности, кодирующей участок распознавания коллагеназы, в качестве линкора при слиянии типичного гена протеина масляного тела и пептида слияния, который нужно связать, используя *NcoI* для клонирования и экспрессии в растениях.

На фиг.4 схематически представлен процесс конструирования векторов пептида слияния, их введение в растения и последующую экстракцию и анализ целевого рекомбинантного пептида.

На фиг. 5 схематически представлено конструирование рCGOPIIT. Прямоугольник с пунктиром представляет олеозинный промотор; прямоугольник, заштрихованный сверху слева-вправо вниз, представляет олеозинную кодирующую последовательность; заштрихованный в клеточку прямоугольник представляет интрон; прямоугольник со штрихами представляет 3'-нетранслируемую последовательность; прямоугольник с редко расположенными полосами сверху слева - вправо вниз,

представляет последовательность интерлейкина - I-β, снабженную последовательностью, кодирующей сайт расщепления протеазы (Фактор Ха или тромбин, расположенный непосредственно в обратном направлении).

На фиг.6 представлена конструкция олигонуклеотида GVR11. На фиг.3A представлена 3' кодирующая последовательность *A.thaliana oleosin*, трансляционно слитая с кодирующей последовательностью фактора Ха/IL-1-β, после которой следует TAA стоп-кодон. Для целей дальнейшего кодирования включают PvuI и Sall рестрикционные сайты. Создание PvuI рестрикционного сайта приводит к дополнительной кодирующей последовательности для аланина (ala). Подчеркнуты последовательности распознавания рестрикционного энзима. Друг над другом расположены последовательности олеозина *A. thaliana* и последовательность распознавания фактора Ха. Действительный сайт расщепления обозначен звездочкой. На фиг. 3B представлена последовательность GVR11 для осуществления слияния с *A. thaliana* олезином, праймер CVR11 должен быть комплементарен последовательности верхней нити.

На фиг. 7 представлена нуклеотидная последовательность OBPILT. Подчеркнута последовательность, кодирующая IL-1- β; последовательность, кодирующая сайт распознавания фактора Ха, указана жирным шрифтом. Терминаторная последовательность нопалинсинтетазы указана мелкими буквами.

Описание предпочтительного варианта

В соответствии с целью изобретения предложены способы и композиции для получения пептидов, которые легко поддаются очистке. Целевой способ включает стадии получения экспрессионной кассеты, содержащей ДНК последовательность, кодирующую достаточную часть специфической последовательности масляного тела, например олеозина, для обеспечения направленного транспорта к масляному телу, и представляющий интерес пептид; трансформации экспрессионной кассетой клеток растения-хозяина; создания трансгенного растения и выращивание его для получения семян, в которых химерный протеин экспрессирован и перемещен к масляным телам. Химерный пептид включает целевой пептид и такой протеин масляного тела, как олеозин. Целевой пептид обычно является чужеродным пептидом, который обычно не экспрессируется в семенах или не находится на масляных телах. Использование протеина масляного тела в качестве носителя или в качестве средства для направленной доставки обеспечивает простой механизм очистки чужеродного протеина. Химерный протеин выделяют из массы клеточных протеинов в одну стадию (например, за счет центрифугирования или флотации); этот протеин также защищен от разложения в процессах экстракции, так как при разделении удаляются неспецифические протеазы из контакта с масляными телами. Ген, кодирующий чужеродный пептид, может быть получен из любого источника, включая растительные, бактериальные, грибковые или животные источники.

Желательно, чтобы химерный пептид содержал последовательности, которые позволяли бы отщеплять целевой пептид от олеозина. Такой метод можно использовать для экспрессии различных пептидов, которые затем легко поддаются очистке.

Обеспечение направленной доставки чужеродного рекомбинантного протеина к масляному телу обеспечивает несколько преимуществ, включая следующие. Протеины можно выделить из клеточной массы после лизиса клеток за счет центрифугирования. Частицы масляных тел будут плавать на поверхности экстракта. Протеин можно необязательно снабдить пептидным линкером, содержащим сайт распознавания протеазы. Это позволяет отделить пептид от масляного тела. Протеин можно в рекомбинантный полипептид таким образом, чтобы он был внутри липофильного консервативного участка. Это приводит к интернализации рекомбинантного пептида в масляное тело, таким образом защищая его от атаки протеазы.

Экспрессионная кассета обычно будет включать в 5'-3' направлении транскрипции транскрипционный и трансляционный регуляторный участок, обеспечивающие возможность экспрессии в развивающихся семенах, типичным примером которых служит промотор, и расположенные в обратном направлении участки, связанные с протеином масляного тела, что обеспечит экспрессию химерного протеина в семенах; ДНК последовательность, кодирующую химерный пептид, содержащий аминокислотную последовательность для обеспечения направленной доставки к масляному телу, и целевой протеин, транскрипционный и трансляционный терминальные участки, функциональные в растениях. Может также присутствовать один или более из интронов.

Специфическая последовательность масляного тела встречает аналогию во фрагментах протеинов масляного тела, в частности олеозинов. Специфическая последовательность масляного тела может быть той же, что и последовательность, получаемая из протеина масляного тела, и они имеют достаточную гомологию, чтобы обеспечить направленную доставку целевого протеина к масляному телу. Под "получаемой" подразумевают аминокислотную последовательность, которая может быть природной, синтетической или комбинированной, достаточно близкой к аминокислотной последовательности протеина нативного масляного тела, для достижения нужного обеспечения направленной доставки. Особый интерес представляют центральный гидрофобный домен протеинов масляного тела, который, по-видимому, в высшей степени консервативен среди различных видов растений, и его фрагментов и гомологичных последовательностей на аминокислотном уровне.

Выведенную последовательность 3 аминокислот для *Arabidopsis thaliana* протеина масляного тела см. в конце описания.

Аминокислоты от около 25 до 101 содержат центральный гидрофобный домен.

Особый интерес представляют в качестве средств, обеспечивающих направленную

доставку для некоторых применений, специфические последовательности 4 масляного тела или их фрагменты последовательности, которые обеспечивают направленную доставку к масляному телу (см. в конце описания).

где:

pp¹ и pp² одинаковы или различны и могут быть такими же или могут отличаться от природного протеина масляного тела и обычно отличаются; они могут представлять водороды, что указывает на терминальную часть указанного полипептида, или они могут быть полипептидами, содержащими вплоть до 1000 аминокислот, обычно вплоть до около 500 аминокислот, и могут содержать лишь одну аминокислоту или могут индивидуально или раздельно быть полипептидами, содержащими от 1 до 100 аминокислот, обычно от около 1 до 75 аминокислот и особенно от около 5 до 50 аминокислот; эти полипептиды могут иметь специфические применения при модификации специфически описанных последовательностей для заранее определенных целей;

aa²⁵ может быть любой аминокислотой, в частности нейтральной алифатической аминокислотой, обычно содержащей от 3 до 6 атомов углерода, более конкретно лейцином или аланином;

aa²⁶ представляет нейтральную алифатическую аминокислоту, особенно аланин или гидроксизамещенную аминокислоту, содержащую от 3 до 4 атомов углерода, в частности треонин, или основную аминокислоту, содержащую от 5 до 6 атомов углерода, в частности лизин;

aa³¹ представляет нейтральную незамещенную алифатическую аминокислоту, содержащую от 3 до 6 атомов углерода, в частности аланин, валин или лейцин, или ароматическую незамещенную аминокислоту, особенно фенилаланин;

aa³³ представляет нейтральную незамещенную алифатическую аминокислоту, содержащую от 3 до 6 атомов углерода, особенно аланин, валин или лейцин, или оксизамещенную алифатическую аминокислоту, особенно треонин;

aa³⁶ представляет нейтральную алифатическую незамещенную аминокислоту, содержащую от 3 до 5 атомов углерода, особенно лейцин, или нейтральную алифатическую оксизамещенную аминокислоту, содержащую от 3 до 4 атомов углерода, особенно треонин или серин;

aa³⁷ представляет нейтральную незамещенную аминокислоту, особенно лейцин или тиозамещенную аминокислоту, особенно метионин;

aa³⁹ представляет нейтральную алифатическую незамещенную аминокислоту, особенно валин, или ароматическую незамещенную аминокислоту, особенно фенилаланин;

aa⁴¹ представляет нейтральную алифатическую незамещенную или оксизамещенную аминокислоту, особенно аланин, лейцин или серин;

aa⁴⁴ представляет нейтральную алифатическую незамещенную или оксизамещенную аминокислоту, в частности аланин, изолейцин или треонин;

aa⁴⁶ представляет нейтральную

алифатическую незамещенную аминокислоту или оксизамещенную аминокислоту, особенно аланин, валин или треонин;

5 aa⁴⁷ представляет нейтральную алифатическую незамещенную аминокислоту, особенно глицин или аланин;

aa⁵⁹ представляет нейтральную алифатическую или ароматическую незамещенную аминокислоту, особенно лейцин или фенилаланин;

10 aa⁷⁶ представляет нейтральную алифатическую незамещенную или тиозамещенную аминокислоту, особенно аланин, лейцин или метионин;

15 aa⁷⁸ представляет нейтральную алифатическую незамещенную аминокислоту, особенно аланин, или нейтральную алифатическую аминокислоту, содержащую тио- или окси-замещения, особенно метионин или треонин;

20 aa⁸³ представляет нейтральную алифатическую незамещенную или оксизамещенную аминокислоту, особенно глицин, серин или треонин;

25 aa⁹² представляет нейтральную алифатическую аминокислоту с окси-замещением, особенно серин или треонин;

aa⁹⁶ представляет нейтральную алифатическую тиозамещенную аминокислоту или нейтральную ароматическую гетероциклическую аминокислоту, особенно триптофан;

30 aa⁹⁷ представляет нейтральную алифатическую незамещенную или тиозамещенную аминокислоту, особенно валин, лейцин, изолейцин или метионин;

35 aa⁹⁸ представляет нейтральную алифатическую незамещенную аминокислоту или ароматическую окси-замещенную аминокислоту, особенно аланин, лейцин или тирозин;

aa⁹⁹ может представлять любую аминокислоту;

40 aa¹⁰⁰ представляет окси-замещенную аминокислоту, либо алифатическую, либо ароматическую, особенно тирозин или треонин;

45 aa¹⁰¹ представляет нейтральную незамещенную алифатическую или ароматическую аминокислоту, особенно аланин, лейцин или фенилаланин.

50 Особый интерес в качестве источника ДНК кодирующей последовательности, способной обеспечить направленную доставку к протеину масляного тела, представляют гены протеинов масляных тел, получаемые из *Arabidopsis* или *Brassica napus*, которые обеспечивают экспрессию целевого протеина в семенах (см. Taylor et al. (1990), *Planta* 181: 18-26). Нужные участки и аминокислотные последовательности, необходимые для обеспечения способности направленной доставки к масляным телам, должны быть, по-видимому, существенно гидрофобным центральным участком протеинов масляного тела.

55 Для идентификации других генов протеинов масляного тела, обладающих нужными характеристиками, где протеин масляного тела уже был выделен, этот протеин можно частично секвенировать, с тем чтобы можно было сконструировать зонд для идентификации мРНК. Такой зонд особенно

ценен, если он сконструирован таким образом, чтобы нацелить его к кодирующему району центрального гидрофобного домена, который в высшей степени консервативен среди различающихся видов растений. Следовательно, ДНК или РНК зонд для этого участка может быть особенно полезен для идентификации кодирующих последовательностей протеинов масляного тела из других видов растений. Для дальнейшей повышения концентрации мРНК можно получить кДНК и эту кДНК вычлест с мРНК или кДНК из клеток, продуцирующих не масляные тела. Оставшуюся кДНК можно затем использовать для зондирования генома для комплементарных

последовательностей, используя соответствующие библиотеки, полученные из растительных клеток. Затем можно выделить последовательности, которые гибридизуются с кДНК в жестких условиях.

В некоторых случаях, как указано ранее, использующих зонд гена протеина масляного тела (консервативный участок), зонд можно использовать непосредственно для скринирования геномной библиотеки кДНК и идентификации последовательностей, которые гибридизуются с зондом. Выделение можно также осуществить стандартным иммунологическим скринированием библиотеки экспрессии кДНК специфической для семян. Антитела можно легко получить к протеинам масляных тел, используя способы очистки и схемы получения антител, описанные Taylor et al. (Planta, (1990), 181, 18-26). Библиотеку экспрессии кДНК скринируют, используя антитела по способу Huynh et al. (1985 в DNA Cloning, Vol.1, a Practical Approach, ed. D.M. Glover, IRL Press, pp. 49-78). Подтверждение структуры последовательности облегчается за счет высокого консерватизма, существующего в центральном гидрофобном участке (фиг.1). ДНК секвенирование по способу Sander et al. (Proc. Natl. Acad. Sci. USA, (1997) 74:5463-5467 или Maxam and Gilbert (1980), Meth. Enzymol., (1980) 65:497-560) можно осуществить на всех предполагаемых клонах и осуществить поиски гомологов. Гомологи последовательностей, кодирующих центральный гидрофобный домен, обычно составляет $\geq 70\%$ как на аминокислотном, так и на нуклеотидном уровнях между различающимися видами. Если доступно антитело, подтверждение идентичности последовательности можно осуществить в экспериментах отбора гибридов и трансляции из препаратов мРНК семян по способу Sambrook et al. (Molecular Cloning, (1990), 2nd Ed., Cold Spring Harbor Press, pp.8-49 до 8-51).

кДНК клоны, полученные из семян, можно скринировать, используя кДНК зонды, полученные из консервативных кодирующих участков любого доступного гена протеина масляного тела (например, Bowman-Vance and Huang (J. Biol. Chem., (1987) 262: 11275-11279). Отбирают клоны, которые обладают степенью гибридизации с ДНК семян, нежеле с кДНК семян. Скринирование повторяют для идентификации конкретной кДНК, связанной с масляными телами развивающихся семян, используя прямое скринирование антителами или отбор гибридов и трансляцию. мРНК,

комплементарные специфическим кДНК, отсутствуют в других тканях, подвергавшихся тестированию. Затем кДНК используют для скринирования геномной библиотеки и отбирают фрагмент, который гибридизуется с нужной кДНК.

Для достижения экспрессии химерного гена в семена участок, регулирующий начало транскрипции, и участок, регулирующий начало трансляции нетранслируемых 5' последовательностей, "рибосомосвязывающие сайты", ответственные за связывание мРНК с рибосомами, и начало трансляции, получаемые из любого гена, предпочтительно экспрессируемого в семенах, могут быть использованы. Примеры таких генов включают гены протеинов хранения семян, например из напина (Josefsson et al., J. Biol. Chem., (1987) 262:12196-12201; Scoield S. R. and Crouch N. L. J. Biol. Chem. (1987) 262: 12203-12208). Предпочтительно, чтобы участок был получен из протеина масляного тела (протеины масляных тел из Arabidopsis, моркови (Hatzopoulos et al., см. ранее) или кукурузы (Huang et al. 1987 и 1990, см. ранее). Такой участок обычно содержит, по крайней мере, 100 bp 5' до начала трансляции последовательности, кодирующей структурный ген, вплоть до 2,5 kb 5' до того же начала трансляции. Предпочтительно, чтобы в се транскрипционные и трансляционные функциональные элементы участка, контролирующего инициацию, были получены или их можно было бы получить из одного и того же гена. Под выражением "можно получить" подразумевают ДНК последовательность, достаточно сходную с нативной последовательностью для того, чтобы обеспечить целевую специфичность транскрипции ДНК последовательности, кодирующей химерный протеин. Она включает природные и синтетические последовательности и может быть комбинацией синтетических и природных последовательностей.

Уровень транскрипции должен быть достаточным для обеспечения такого количества РНК, которое способно обеспечить получение модифицированных семян. Под термином "модифицированные семена" подразумевают семена, обладающие заметной отличающимся фенотипом от фенотипа семян нетрансформированного растения того же вида, например не содержащих рассматриваемой кассеты экспрессии и в его геноме. Представляют интерес различные изменения в фенотипе. Эти изменения включают избыточную экспрессию протеина масляного тела или ОВР-накопление на масляном теле или полученного химерного протеина в цитоплазме.

Представляющим интерес полипептидом может быть любой протеин, и они включают, например, энзим, антикоагулянт, нейропептид, гормон или адгезивный предшественник. Примеры протеинов включают интерлейкин - I- β , антикоагулянт Hirudin, энзим β -глюкуронидазу или одноцепочечное антитело, содержащее трансляционное слияние V_H или V_L цепей иммуноглобулина. ДНК последовательность, кодирующая целевой полипептид, может быть синтетической или природной или их комбинацией. В зависимости от характера

источника ДНК кодирующей целевой полипептид может оказаться желательным синтезировать ДНК последовательность с предпочтительными растительными кодонами. Предпочтительно растительные кодоны модно определить из кодонов, наиболее часто встречающихся в протеинах, экспрессированных в больших количествах в конкретные виды растений, представляющих интерес в качестве растений-хозяев.

Используемый терминаторный участок должен быть преимущественно одним из обычных, так как во многих случаях терминационные участки, по-видимому, относительно взаимозаменяемы.

Терминационный участок может быть нативным с участком инициирования транскрипции, может быть нативным с ДНК последовательностью, кодирующей целевой полипептид или может быть получен из другого источника. Удобные терминационные участки доступны из Ti - плазмид *A. tumefaciens*, такие как октопин-синтазные и нопалин-синтазные терминационные участки.

Легирование ДНК последовательности, кодирующей обеспечивающую направленную доставку последовательность с геном, кодирующим целевой пептид, может осуществляться различными способами, включая терминальное слияние, внутреннее слияние и полимерное слияние. Во всех случаях такие слияния осуществляют таким образом, чтобы не прерывать считывающую рамку протеина масляного тела, и с тем, чтобы избежать каких-либо сигналов прекращения трансляции внутри или вблизи этих соединений. Различные типы терминальных и внутренних слияний представлены на фиг.2 наряду с представлением их конфигураций *in vivo*.

Во всех описанных случаях легирование гена, кодирующего пептид, предпочтительно должно включать линкер, кодирующий протеазный мишеневый фрагмент. Это позволит осуществить выделение пептида после его экстрагирования в виде протеина слияния. Потенциальные сайты расщепления, которые можно использовать, являются фрагментами для тромбина (leu-val-pro-arg-gly) (Fujikawa et al., *Biochemistry* (1972) 11:4892-4899), фактора Ха (phe-glu-gly-arg-aa) (Nagat et al., *Proc. Nat'l Acad. Sci. USA*, (1985) 82:7252-7255) или коллагеназы (pro-leu-gly-pro) (Scholtissek and Grosse *Gene* (1988) 62:55-64).

За счет соответствующих манипуляций, таких как рестрикция, переваривание (chewing back) или заполнение липких концов для создания тупых концов, легирование линкеров или т.п., комплементарные концы фрагментов можно получить соединением и легированием. При проведении различных стадий используют клонирование с тем, чтобы умножить количество ДНК и обеспечить возможность анализа ДНК, для того чтобы убедиться, что операция прошла должным образом. Доступно множество векторов клонирования, когда вектор клонирования включает функциональную в *E. coli* систему репликации и маркер, который обеспечивает селекцию трансформированных клеток. Иллюстративные векторы включают pBR332, pUC серии, M13mp серии, PACYC 184 и т.д. Так, последовательность можно встроить в

вектор по соответствующему рестрикционному сайту (сайтам), полученную плазмиду использовать для трансформации *E. coli* хозяина, вырастить *E. coli* в соответствующей питательной среде, собрать клетки, провести лизис и выделить плазмиду. Анализ может включать анализ последовательности, рестрикционный анализ, электрофорез или т.п. После каждой манипуляции ДНК последовательность, которая должна быть использована в окончательной конструкции, может быть рестриктирована и соединена со следующей последовательностью, где каждая из частичных конструкций может быть клонирована в ту же самую или в отличающуюся плазмиду.

Для введения ДНК в растительные клетки хозяина существует множество возможных методик. Так например, химерные ДНК конструкции можно вводить в клетки хозяев, полученные из двудольных растений, например табака, и таких масличных видов, как *Brassica napus*, используя стандартные *Agrobacterium* векторы, по такой схеме трансформации, как например та, которая описана Moloney et al. *Plant Cell Rep.*, (1989) 8:238-242 или Hinchee et al/ *Bio/Technol.*, (1988) 6:915-922; или другие методики, известные специалистам. Так например, использование Т-ДНК для трансформации растительных клеток интенсивно исследуется и подробно описано в EPA 120516; Hcekema, B "The Binary Plant Vector System Offset-drukkerij Kanters B.V., Albladderdam, 1985, Chapter V., Knauf, et al., *Genetic Analysis of Host Range Expression by Agrobacterium*, B "Molecular Genetics of the Bacteria. *Plant Interaction*, Puhler, A.ed., Springer-Verlag, NY, 1983, p.245, and An et al., *EMBO J.* (1985), 4:277-284.

Удобно эксплантаты культивировать с *A. tumefaciens* или *A. rhizogenes* для того, чтобы обеспечить перенос транскрипционной конструкции в растительные клетки. После трансформации с использованием *Agrobacterium* растительные клетки диспергируют в подходящей селективной среде для отбора, выращивают до образования каллуса, проростки и молодые растения регенерируют из каллуса, выращивая в среде для образования корней. *Agrobacterium* хозяин будет содержать плазмиду, содержащую *Vir* гены, необходимые для переноса Т-ДНК в растительные клетки, и может содержать или не содержать Т-ДНК. Для инъекции и электропорации (см. далее) обнаруженные Т1 - плазмиды (не содержащие опухолевых генов, особенно Т-ДНК участка) можно вводить в растительные клетки.

Использование не-*Agrobacterium* методик позволяет использовать описанные здесь конструкции для осуществления трансформации и экспрессии в широкий круг однодольных и двудольных растений. Эти методики особенно подходят для таких видов, которые не дают результатов в *Agrobacterium* трансформационной системе. Другие методики для переноса генов включают биолизис (Sanford, *Trends in Biotech.* (1988) 6:299-302), электропорацию (Fromm et al. (1985) *Proc. Nat'l. Acad. Sci. USA* 82:5824-5828; Riggs and Batec (1986), *Proc. Nat'l. Acad. Sci. (USA)* 83, 5603-5606 или

захват ДНК с помощью PEG (Potrykus et al. (1985), Mol. Gen. Genet. 199:169-177).

В качестве клеток хозяев можно использовать клетки любых растений, имеющих семена, причем эти клетки получают на таких частях растений, как стебли, листья, корни или семена, или соответствующих виду репродуктивных структур. Клетки могут быть как выделенными клетками, так и частями растений, например листовыми дисками. В специфических применениях, таких как с *Brassica napus*, клетки хозяева обычно получают из семядольных петиолей, как указано Moloney et al. Plant Cell Rep. (1989) 8:238-242). В других примерах с использованием коммерческих маслосодержащих семян используют трансформацию семядолей в соевые эксплантаты (Hihchee et al. Biotechnology, (1988) 6: 915-922) и трансформацию стебля хлопка (Umbeck et al. Biotechnology, (1981) 5:263-266).

После трансформации клетки, например, в виде дисков листьев, выращивают на селективной среде. После того как начинают образовываться ростки, их иссекают и помещают на корнеобразующую среду. После достаточного развития корней растения переносят в почву. Затем предположительно трансформированные растения тестируют на присутствие маркера. Осуществляют Саузерн-блоттинг на геномной ДНК, используя соответствующий зонд, например, *A.thaliana* олеозинный ген для доказательства интеграции целевой последовательности в геном клетки хозяина.

Экспрессионную кассету обычно соединяют с маркером для селекции в растительных клетках. Удобно, чтобы маркер был устойчивым к гербицидам, например, антибиотику, такому как канамицин, G41B, блеомицин, гигромицин, хлорамфеникол и т.п. Конкретным используемым маркером должен быть такой, который обеспечит селекцию трансформированных клеток при сравнении с клетками, которые не содержат ДНК, которую вводили.

Пептид слияния в кассете экспрессии, сконструированной, как указано ранее, экспрессируется, по крайней мере, предпочтительно, в развивающихся семенах. Соответственно, трансформированным растениям, выращенным в соответствии с принятыми способами, дают возможность завязать семена. См. например, McCormick et al. Plant Cell Reporta (1986) 5:81-84. Норзерн-блоттинг можно вести, используя соответствующий генный зонд с РНК, выделенной из ткани, в которой ожидается транскрипция, например в эмбрионе семени. Затем размеры транскриптов можно сравнить с размерами, предсказанными для транскриптов протеина слияния.

Можно вырастить два или более поколений трансгенных растений и либо опылить их пыльцой от того же трансформированного штамма, либо от других штаммов, контролируя в полученном гибриде нужные фенотипические характеристики для подтверждения того, что нужные фенотипические характеристики стабильно сохраняются и наследуются, а затем семена собирают для того, чтобы выделить представляющий интерес пептид или для того, чтобы получить семена,

обладающие новыми фенотипическими характеристиками.

Целевой протеин можно экстрагировать из семян, которые могут быть гомо- или гетерозиготными, для последующих испытаний различными способами, которые включают использование водных, забуференных экстракционных сред и различные способы измельчения (размалывания, расклевывания, пульверизации) для разрушения клеток семян. Затем продукты экстракции можно разделить (например, центрифугированием или за счет седиментации) на три фракции: осадок, или нерастворимый осадок, водный супернатант и плавающую пленку, содержащую липид хранения семян и масляные тела. Эти масляные тела содержат как нативные протеины масляных тел, так и химерные протеины масляных тел, причем последние содержат чужеродные пептиды. Масляные тела выделяют из водорастворимых протеинов и повторно суспендируют в водном буфере.

Если в кассету экспрессии был включен линкер, содержащий фрагмент распознавания протеазы, тогда к повторно суспендированному буферу добавляют протеазу, специфичную для фрагмента распознавания, полученного в результате трансляции линкерной последовательности. Это приводит к выделению целевого пептида в водную фазу. Теперь в результате второй стадии центрифугирования снова получают плавающую пленку обработанных масляных тел с присоединенными к ним протеинами и остается водный раствор целевого пептида. Этот целевой пептид можно осадить, химически модифицировать или лиофилизировать в зависимости от предполагаемого применения.

В некоторых случаях применения нет необходимости отделять химерный протеин от протеинов масляного тела. Такие применения включают случаи, в которых пептид слияния включает энзим, который толерантен к N- или C-терминальному слиянию и сохраняет свою активность; такие энзимы могут быть использованы без дальнейшего расщепления и очистки. Химерный энзим-ОВР будет контактировать с субстратом как протеин слияния. Возможно также при желании провести очистку энзим-ОВР протеина слияния, используя иммуноафинную колонку, содержащую иммобилизованные антитела с высоким титром против ОВР (см. например, Taylor et al. (1990) ранее).

Возможно другое применение предмета изображения. Так как ОВР содержат высокий процент протеинов семян, оказывается возможным снабдить семена для некоторых целей такими ценными свойствами, как высокое содержание лизина метионина и т.д., просто обогащая протеин слияния нужными аминокислотами. Этот способ может найти применение особенно при модификации зерновых, которые прямо или косвенно используются для корма домашних животных, включая скот, птицу, а также для продуктов питания людей. Может оказаться возможным включать как пептид слияния энзим, который облегчит последующую обработку масла или муки в обычных процессах измельчения и экстракции семян масличных растений,

например, за счет включения термостабильного липид-модифицирующего энзима, который останется активным и при повышенных температурах измельчения, которые обычно наблюдаются в этом процессе обработки семян, и тем самым удастся повысить ценность экстрагированных триглицеридов или протеинов.

Другие применения протеинов слияния включают использование их для повышения устойчивости сельскохозяйственных растений и неблагоприятных условиях. Так например, с протеином масляного тела можно соединить инсектицидный протеин или часть иммуноглобулина, специфичного для сельскохозяйственных вредителей, такую как стенки клеток или мембран грибов, тем самым уменьшая опасность нападения на семена конкретных вредителей растений.

Приводимые далее примеры представлены лишь в целях иллюстрации и не являются ограничивающими.

Экспериментальная часть

Пример 1

Экспрессия терминальных слияний чужеродных пептидов с протеинами масляных тел

А. С-терминальные слияния

Геномный клон гена протеина масляного тела, содержащий, по крайней мере 100bp 5' до начала трансляции, клонируют в плазмидном векторе, способном к репликации в подходящем бактериальном хозяине (например, pUC или pBR322 в *E. coli*). Рестрикционный сайт расположен в участке, кодирующем гидрофильную С-терминальную часть гена. В 19 кДа ОВР этот участок простирается обычно от кодонов 125 до конца клона. Идеальный рестрикционный сайт является уникальным, но это не является абсолютно необходимым. Если в этом участке нет удобного рестрикционного сайта, его можно ввести за счет сайт-направленного мутагенеза по Kunkel Proc. Nat'l. Acad. Sci. USA, (1985) 82:488-492. Единственным основным ограничением по встраиванию этого сайта является то, что он должен быть расположен в направлении 5' от трансляционного стоп сигнала ОВР клона.

Получив такой мутантный клон, можно получить синтетический олигонуклеотидный адаптор, который содержит кодирующую последовательность для сайта распознавания протеазы, например, Pro-Leu-Gly-Pro или его мультимер. Это сайт распознавания для протеазы коллагеназы. Адаптор следует синтезировать таким образом, чтобы обеспечить: 4-основной выступ у 5' конца, совместимый с рестрикционным сайтом 3' конца ОВР клона, 4-основной выступ у 3' конца адаптора для облегчения легирования с последовательностью, кодирующей чужеродный пептид, и дополнительными основаниями при необходимости, чтобы обеспечить отсутствие сдвигов рамки в переходе между ОВР кодирующей последовательностью, сайтом распознавания протеазы и последовательностью, кодирующей чужеродный пептид. Типичная конструкция такого слияния представлена на фиг.3. Представленный пример использует существующий XhoI сайт вблизи стоп кодона ОВР моркови (Hatzopoulos et al. Plant Cell, (1990) 2:457-467. Его переваривают, и он может быть легирован с адаптором,

сконструированным из двух описанных олигонуклеотидов. Этот адаптор образует прекрасный NcoI выступ на конце и не будет нарушать трансляционной рамки. Другой конец образует NcoI выступ, который выбирают произвольно (подходят любые отрезки из 6 оснований), но который содержит ATG из целевого чужеродного пептида.

Окончательный продукт легирования содержит самое большее полный ОВР ген, последовательность, кодирующую фрагмент распознавания коллагеназы, и участок, кодирующий целевой пептид, причем все в единой считывающей рамке. Этот состоящий из трех частей фрагмент клонируют в *Agrobacterium* бинарную плазмиду (Bevan Nucl. Acid Res. (1984) 12: 8711-8721), такую, как обычно используют для переноса чужеродной ДНК в растения Fraley et al. Proc. Nat'l. Acad. Sci. USA (1983) 80:4803-4807), и это используют для трансформации семян масличных растений, таких как рапс, используя способ Moloney et al. Plant Cell Rep. , (1989) 8:238-242) или аналогичную процедуру. Трансгенные растения могут быть получены из этого эксперимента по трансформации и выращены до цветения. Затем получают семена растения в результате самоопыления.

Семенам дают взреть (60-80 дней), а затем собирают и измельчают в водном экстракционном буфере (Taylor et al. Planta, (1990) 181:18-26). Полученную суспензию центрифугируют при 5000xg в течение 20 минут, в результате получают на поверхности пленку. Ее снова выделяют и суспендируют при интенсивном встряхивании в коллагеназном аналитическом буфере (Scholtissek and Grosse, Gene (1998) 62:55-64). Добавляют 5 ед. коллагеназы и суспензию инкубируют со встряхиванием в течение 4 часов. После этого суспензию снова центрифугируют при 5000 xg в течение 20 минут. Удаляют поверхностную пленку и определяют содержание протеина в водной фазе с помощью электрофореза на SDS-полиакриламидном геле. Если обнаруживают полосу приблизительно того же размера, что и нужный пептид, протеин можно осадить, используя сульфат аммония, концентрирование с использованием ультрафильтрации, или лиофилизировать.

В. N-терминальное слияние

Гидрофильный N-терминальный конец протеинов масляных тел позволяет осуществить слияние пептидов с N-концом, обеспечивая при этом сохранение чужеродного пептида на внешней поверхности масляного тела. Конфигурации таких слияний представлены на фиг. 2IB.

Такую конфигурацию можно сконструировать из таких же исходных материалов, что были использованы для С-терминальных слияний, но необходима идентификация удобных рестрикционных сайтов вблизи начала трансляции гена протеина масляного тела. Удобный сайт можно создать для многих генов протеинов масляных тел без изменения в кодирующей последовательности за счет введения изменения одного основания сразу в направлении 5' к первому "ATG". В исследованных до настоящего времени протеинах масляных тел второй аминокислотой является аланин, кодон

которого начинается с "G". Контекст последовательности 5 представлен в конце описания.

Изменение одного основания у аденина перед "ATG" дает в обоих случаях... CCATGG..., что представляет собой NcoI сайт. Таким образом, модификация этого основания с помощью сайт-направленного мутагенеза по Kunkel (Proc. Nat'l. Acad. Sci. USA, (1985) 82:488-492) дает возможность подготовить этот клон для использования, предполагая, что в последовательности нет других NcoI сайтов.

Кодирующая последовательность для чужеродного пептида может потребовать изменений, которые обеспечат ее лигирование непосредственно по NcoI сайту. Обычно это может потребовать изменения одного или двух оснований за счет сайт-направленного мутагенеза (Kunkel, 1985, ранее) для создания NcoI сайта около начала трансляции чужеродного пептида. Затем этот пептид иссекают из его вектора клонирования, используя NcoI и второй энзим, который разрезает вблизи остановки трансляции мишени. И снова, используя описанный ранее способ, второй удобный сайт можно ввести за счет сайт-направленного мутагенеза. Было предложено, Qi and Huang (1990, ранее), что N-терминальный метионин может быть удален в процессинге протеина *in vivo* и что аланин, который следует непосредственно в прямом направлении от него, может быть ацилирован. Чтобы учесть эту возможность, может оказаться необходимым сохранить последовательность Met-Ala по N-терминальному концу протеина. Это легко осуществить, используя различные стратегии введения удобного рестрикционного сайта в кодирующую последовательность в Ala кодон или после него. Так например, за счет сайт-направленного мутагенеза последовательность 6 можно модифицировать (см. в конце описания).

Такое изменение одного кодона позволяет ввести SphI сайт в кодирующую последовательность. Второе изменение, которое можно осуществить в том же цикле мутагенеза, превращает два основания в кодоне 6 до получения GGC, GCC, что представляет NarI сайт. Этот полученный в результате мутаций ген можно затем раскрыть SphI и NarI до получения отрезка прямого клонирования, который исключает три кодона. В этот сайт можно ввести адаптор, содержащий 3' выступ с последовательностью CATG... (совместимую с SphI) GC 5' выступ с противоположного конца.

Точная последовательность 7 этого адаптора представлена в конце описания.

Этот адаптор будет воссоздавать как SphI так и NarI рестрикционные сайты, которые можно использовать для диагностических целей. SphI сайт теперь можно использовать для раскрытия плазмиды и клона в рамке ДНК фрагмента, содержащего последовательность для нужного пептида. Затем следует проанализировать ориентацию клонирования за счет нарезания по любым несимметрично расположенным сайтам и NarI плазмиды.

Полученные из этих N-терминальных слияний конструкции могут быть типичными примерами IB фиг. 2. Они должны содержать ОВР промоторную последовательность в

первых нескольких кодонах ОВР гена, представляющего высокую ценность пептида с его собственной АТГ в качестве стартового сигнала при необходимости и остальную часть ОВР гена и терминатор.

Этот модифицированный ген вводят в бинарную *Agrobacterium* плазмиду (Bevan, (1984, ранее) и мобилизуют в *Agrobacterium*. Трансформации осуществляют в соответствии с вышеприведенным описанием. Выделение ценного пептида из семян осуществляют по способу, описанному в разделе "С-терминальные слияния".

С. Внутренние трансляционные слияния

Третий тип слияний включает помещение кодирующей ценный пептид последовательности внутрь кодирующей последовательности ОВР. Этот тип слияния требует той же стратегии, что и N-терминальное слияние, но может быть функциональным лишь при модификациях в участках с низкой консервативностью, так как считают, что участки с высокой степенью консервативности в этих ОВР существенны для обеспечения направленной доставки зрелого протеина.

Ключевое отличие в этом типе слияния состоит в необходимости фланкирующих сайтов распознавания коллагеназы для высвобождения протеина. Это означает, что вместо стандартной коллагеназной системы линкер/адаптор, описанной ранее, необходимо иметь линкер вида, представленного в последовательности 8 (см. в конце описания).

Липкие концы 1 и 2 будут использованы для клонирования адаптора в ОВР клон непосредственно. Затем встроенный рестрикционный сайт используют для введения последовательности, кодирующей ценный пептид, фланкированный рестрикционными сайтами или линкерами. Ориентацию контролируют, используя асимметрично расположенный рестрикционный сайт в последовательности, кодирующей ценный пептид, и один из двух рестрикционных сайтов, фланкирующих последовательность, кодирующую фрагмент распознавания коллагеназы.

Мобилизация этих конструкций в *Agrobacterium* плазмиды, а затем в растения идентична ранее описанной процедуре. Выделение ценного протеина из семян трансгенных растений несколько отличается тем, что после того как масляные тела выделяют и промывают, может понадобиться удалить липиды с масляных тел для того, чтобы сделать доступными сайты распознавания коллагеназы, которые могут быть спрятаны внутри масляных тел в липидной фазе. Эта стадия может уменьшить некоторые преимущества использования протеинов масляных тел в качестве носителей, но может, с другой стороны, оказаться очень удобной для последовательностей протеинов, которые лабильны в водной среде или в растительных цитоплазмах.

Д. Интердимерные трансляционные слияния

Возможно создать конструкцию, в которой целиком повторяется кодирующая последовательность ОВР. Димерный протеин, получаемый из такой конструкции, может все еще содержать все необходимые факторы

для обеспечения направленной доставки ОВР к масляным телам. Такая конструкция должна содержать промоторный участок, полную или почти полную открытую рамку считывания для ОВР, но без остановки трансляции, и затем полную открытую считывающую рамку второго ОВР, на этот раз укомплектованного трансляционным "стоп" и терминаторным участком.

При конструировании этого химерного гена пару различающихся рестрикционных сайтов либо находят, либо создают в участке соединения этих двух копий. Эти сайты используют для обеспечения возможности введения такого линкера, как описанный ранее для внутренних трансляционных слияний. Линкеры содержат не только наборы фрагментов распознавания коллагеназы, но также внутренний рестрикционный сайт, в который вставляют последовательность, кодирующую представляющий высокую ценность протеин. Форма такой конструкции представлена на фиг. 2 III. Введение этой конструкции в *Agrobacterium*, а затем в растении осуществляется описанным ранее способом. Выделение представляющего высокую ценность протеина из семян трансформированных растений осуществляют по способу, описанному ранее для процедуры С-терминального слияния.

Пример 2.

Стратегия клонирования и экспрессии Интерлейкина-1-β (IL-1-β) в результате слияния с олеозинами в растениях

А. Клонирование и секвенирование гена олеозина

Arabidopsis thaliana.

Ген олеозина *Brassica napus* (Murphy et al., (1991) *Biochim Biophys Acta* 1086: 86-94) используют для скринирования геномной библиотеки *A.thaliana* (cv. Columbia) в EMBL3A (Stratagene). В результате скринирования получают выделенный KMBL3A клон (λ 2,1), содержащий 15 kb геномный фрагмент из *A. thaliana*. Олеозин картируют внутри 6,6 kb *Kpn1* вставки, внутри этого 15 kb фрагмента (фиг.5). 1,8 kb *NcoI/Kpn1* фрагмент, содержащий ген олеозина, имеет тупой конец, и его субклонировать в *SmaI* сайт RFM13mp19. 1,8 kb вставку переваривают подходящими рестрикционными ферментами и субклонировать в M13mp19 для секвенирования. Последовательность 1800 kb гена олеозина *A.thaliana* представлен на фиг.1а. Все процедуры клонирования осуществляют по Sambrook et al., (1989) (*Molecular Cloning: A laboratory manual 2nd ed.* Cold Spring Harbor Laboratory Press.).

В. Конструирование олигонуклеотида, кодирующего IL-1-β

IL-1-β состоит из 9 аминокислот (aa); val-gln-gly-glu-glu-ser-asn-asp-lys (Antoni et al., (1986) *J. Immunol.* 137:3201-3204). Протеазный фактор Ха может расщепить последовательность протеина, которая содержит последовательность аминокислот ile-glu-gly-arg. Расщепление происходит после аминокислоты arg. На основании этих последовательностей конструируют олигонуклеотид (GVR11, фиг.5), который содержит дополнительно к IL-1-β кодирующей последовательности кодирующую последовательность для сайта расщепления фактора Ха и 18 нуклеотидов 3' кодирующего

участка *A. thaliana* олеозина (положения оснований 742-759). Сконструирована IL-1-β кодирующая последовательность, используя оптимальное использование кодона для *B.napus* и *A. thaliana* олеозина (таблица 1)

С. Создание *A. thaliana* олеозин IL-1-β слияния. На основании

последовательности 9 была проведена следующая операция (см. в конце описания). GVR10 содержит *Pst1* рестрикционный сайт (подчеркнут) для облегчения клонирования. Полимеразная цепная реакция (PCR) была использована для амплификации участка между CVR10 и CVR11. Реакционная смесь содержит: 16 мкл dNTR_s (1,25 mM), 10 мкл 10X PCR буфер (100 mM Tris-HCl pH 8,3, 500 mM, KC1, 15 mM MgCl₂, 0,1 % (вес/объем) желатин, 5 мкл H₂O. Реакцию ведут в течение 30 циклов. Каждый цикл включает 1 минуту денатурирования при 92°C, 1 минуту отжига при 45°C и 3 минуты удлинения при 72°C. В результате реакции PCR получают один фрагмент длиной 1652 нуклеотида.

Д. Клонирование *A. thaliana* олеозин - IL-1-β (OBPIL) слияния

5' *Sall*-нопалин-синтеазную (nos)

терминаторную *EcoRI* 3' последовательность выделяют из pBI121 (Clontech laboratories) и клонируют в *Sal/EcoRI* сайты pUC19. Плазмиду называют pTerm. Фрагмент 1652 bp (описан в разделе С) выделяют и переваривают рестрикционными ферментами *PstI* и *Sall*. Этот фрагмент клонируют в pTerm. Полученную плазмиду называют pUCOBPILT (фиг.5). Эту плазмиду переваривают *EcoRI* и *PstI* и получают в переваре pUC19 вектор и *EcoRI*- *A. thaliana* олеозин-IL-1-β - nos - *PstI* слитый *PstI*

(OBPILT). Полная последовательность OBPILT представлена на фиг.7. OBPILT субклонировать в *EcoRI/PstI* сайты pBluescript+. Эту плазмиду (pBIOBPILT) переваривают *PstI* или *HindIII* и фрагмент *PstI*-OBPILT-*HindIII* субклонировать в бинарную *Agrobacterium* плазмиду (Bin 19) (Bevan. M., (1984) *Nucl.Acid.Res.* 12:8711-8721),

содержащую селекционный маркер (неомицинфосфотрансферазу) и *PstI*-*HindIII* уникальные сайты. Полученную плазмиду называют pCGOBPILT. Схема процедуры клонирования представлена на фиг. 5. Для описания различных бинарных плазмид смотри pGA642 или 645; An et al. (1985) *EMBO. J.* 4 277-288 или pCGN1558 или 1559; MacBride and Summerfeldt (1990) *Plant. Molec. Biol.* 14.269-276.

Ф. Трансформация pCGOBPILT в *Agrobacterium* штамм EHA101.

Отдельную EHA101 колонию (Hood et al., (1986) *J. Bact.* 168:1291-1301) используют для того, чтобы инокулировать 5 мл LB+100 мкг/мл канамицина. Эту культуру выращивают в течение 48 часов при 28°C. 5 мл этой культуры используют для инокулирования 500 мл LB+100 мкг/мл канамицина. Эту культуру выращивают при 28°C до тех пор, пока плотность культуры не достигает OD₆₀₀ = 0,5 (приблизительно 4 часа). В результате центрифугирования клетки осаждают (10 мин, 5000xg) и снова суспендируют в 500 мл стерильной H₂O (повторяют 2 раза). Клетки снова центрифугируют и снова суспендируют в 3 мл стерильной H₂O, содержащей 10%

глицерина. Аликвоты по 40 мкл помещают в ампулы Эппендорфа и либо непосредственно используют для электропорации, либо хранят при -80°C до дальнейшего использования. Электропорацию осуществляют по способу Bower et al., Nucl. Acid. Res. (1888) 16 6127-6145. Мощность импульсного генератора устанавливают 25 мкФ, 2,5 кВ и 200 Ом параллельно с камерой образца.

G. Трансформация *Nicotina tabacum* (табак) с помощью pCGOVPILT

Для трансформации дисков листьев табака используют EHA101, содержащий pCGOVPILT. Листья табака длиной 8-10 см берут с растений, растущих в оранжерее, стерилизуют в 70% этаноле в течение 20 секунд, а затем в 10% хлорной извести (например, Javex) в течение 8 минут. Затем листья 6 раз промывают стерильной водой. Края листьев, а также среднюю жилку иссекают, а остальную часть листовой пластины нарезают на квадраты 5x7 мм или диски диаметром 5 мм. Собирают около 30 листовых дисков и помещают в небольшую чашку Петри. Затем на диски выливают раствор *Agrobacterium* и инкубируют в течение 9 минут. Затем осуществляют блоттинг на стерильной Ватманской фильтровальной бумаге и помещают плоской стороной вниз на среду I (MS, 3% сахарозы и 2 мг/л 2,4-Д). Такое совместное культивирование ведут в течение последующих 48 часов. В этот момент диски листьев переносят на селекционную среду (MD, 3% сахарозы, 2,5 мг/л Ва, 0,1 мг/л NAA, 500 м г/л карбенициллина и 100 мг/л канамицина), где и оставляют затем на 3-4 недели. После появления проростков их иссекают и помещают на корнеобразующую среду (MS, 3% сахарозы, 0,1 мг/л NAA, 500 мг/л карбенициллина и 50 мг/л канамицина). После образования достаточной корневой системы растения табака переносят в почву.

H. Трансформация *B. parus* с помощью pCGOVPILT

Трансформация на *B. parus* ведут по способу Noloney et al., (1989) Plant Cell Rep., 8: 238-242 (включено сюда по ссылке).

Процедура трансформации

Отдельные колонии *Agrobacterium tumefaciens* штамма EHA 101, содержащие бинарную плазмиду, выращивают в течение ночи при 28°C на АВ среде. Образец 50 мкл надосадочной жидкости этой среды выращивают в течение ночи при 28°C в 5 мл бульона MG/L, дополненного соответствующими антибиотиками. Эту бактериальную суспензию осаждают центрифугированием в течение 15 минут при 10000xg, затем повторно суспендируют в 10 мл MS среды, содержащей 3% сахарозы при pH 5,8. Тонкую пленку этой суспензии используют для выстилания дна чашки Петри диаметром 5 см. Отдельные вырезанные семядоли берут с описанных ранее пластин и срезанную поверхность их петиолей погружают в эту бактериальную суспензию на несколько секунд. Их немедленно возвращают на те же самые MS пластины, с которых их взяли. Семядоли культивируют совместно с *Agrobacterium* в течение 72 часов; при этом не используют питательных слоев.

После совместного культивирования семядоли переносят на регенерационную среду, содержащую MS среду, дополненную

20 мкМ бензиладенина, 3% сахарозы, 0,7% фитагара, pH 5,8 и 500 мг/л карбенициллина (Puopen, Ayerat) и 15 мг/л канамицин-сульфата (Boehringer-Mannheim). Снова петиоли осторожно вставляют в агар на глубину 2 мм. Плотность растений поддерживают около 10 эксплантатов на пластину. Более высокая плотность снижает регенерируемость.

Отбор и регенерация растений

Эксплантаты выдерживают на регенерационной среде в специфических условиях освещенности и температуры в течение 2-3 недель. За это время более чем на половине эксплантатов появляются ростки при относительно небольшом образовании каллуса. Некоторые из этих ростков обесцвечиваются к четвертой неделе культивирования. Оставшиеся зеленые ростки субкультивируют на среде для удаления корней, которая представляет собой регенерационную среду, но без бензиладенина. Одна-две недели на этой среде обеспечивают установление апикального доминирования из образовавшихся ростковых кластеров. Полученные таким образом ростки переносят на "корнеобразующую" среду, содержащую MS среду, 3% сахарозы, 2 мг/л индолмасляной кислоты, 0,7% фитагара и 500 мг/л карбенициллина. На этой стадии не используют канамицин, так как было обнаружено, что без селективирующего агента происходит более быстрое корнеобразование, тогда как очень мало "промахов" реально наблюдается на корнеобразовании после двух раундов отбора на регенерационной и удлиняющей корни среде.

I. Стабильная интеграция OVPILT в геномы табака и *B. parus*

Предположительно трансформированные растения тестируют на активность неомицин-фосфотрансферазы. Выделяют геномную ДНК из растений, демонстрирующих такую активность. Саузерн-блоттинг осуществляют для демонстрации того, что последовательности между T-ДНК границами (OVPILT и геном неомицин-фосфотрансферазы) стабильно интегрированы в геномы *B. parus* и табака. Саузерн табака зондируют *A. thaliana* олеозиновым геном, а геном неомицин-фосфотрансферазы. *B. parus* Саузерн зондируют геном неомицин-фосфотрансферазы.

J. Экспрессия олеозин-IL-1-β слияния в растениях табака РНК выделяют из развивающихся эмбрионов, полученных из трансформированных и нетрансформированных растений. Нозерн-блоттинг осуществляют, используя *A. thaliana* олеозин в качестве генного зонда. Во всех тестируемых трансформированных растениях можно обнаружить 850-нуклеотидный транскрипт. Размеры этих транскриптов соответствуют ожидаемым размерам мРНК олеозин-IL-1-β. Эти транскрипты обнаруживаются в нетрансформированных растениях.

K. Накопление олеозин -IL-1-β протеина

Протеины масляных тел выделяют из трансформированных семян табака (Holbrook et al., (1991) Plant Physiol 97:1051-1058). Осуществляют PAGE и протеин переносят из геля на PVDF мембраны. Для определения

слияния олеозин-IL-1-β в семенах табака используют антитела, которые вырабатываются против 22 кДа олеозина *B. parvus*. Эти антитела распознают все основные олеозины в *B. parvus* и *A. thaliana*. Кроме того, эти антитела распознают олеозины табака. Олеозины табака имеют размеры, отличающиеся от олеозинов *A. thaliana* и *B. parvus*. В трансформированных семенах табака анти-22 кДа антитела распознают 20 кДа - протеин, который отсутствует в семенах нетрансформированных растений.

Предсказанный размер олеозин-IL-1-β слияния составляет 20,1 кДа. Полученные результаты представлены в таблице 2.

За счет экспрессии представляющего интерес пептида, конъюгированного с протеином масляного тела или его частью, достаточной для обеспечения его направленного транспорта к масляным телам, представляющий интерес пептид можно легко очистить, то есть практически отделить его от других клеточных компонентов. Протеин слияния можно расщепить после очистки или его можно использовать далее без расщепления. Рассматриваемые способы и композиции обеспечивают быстрый простой способ очистки представляющего интерес полипептида.

Все публикации и патентные заявки, указанные в этом описании, включены сюда по ссылке.

Далее настоящее изобретение раскрыто полностью, но специалистам должно быть ясно, что не выходя за рамки объема и сути прилагаемой формулы изобретения, можно осуществить множество изменений и модификаций.

Формула изобретения:

1. Слитый полипептид, способный к целенаправленному переносу к масляному телу, включающий а) первый пептид, имеющий формулу

M-M-G-R-D-R-D-Q-Y-Q-M-S-G-R-G-S-D-Y-S-K-S-R-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-T-V-A-T-P-L-L-V-I-F-S-P-I-L-V-P-A-L-I-T-A-V-L-L-I-T-G-F-L-S-S-G-G-F-G-I-A-A-I-T-V-F-S-W-I-Y-K-Y-A-T-G-E-H-P-Q-G-S-D-K-L-D-S-A-R-M-K-L-G-S-K-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-T-T,

и б) второй пептид, выбираемый из группы, включающей цитокин, фермент, антикоагулянт, нейропептид, гормон, антитело или предшественник молекулы адгезии.

2. Слитый полипептид по п. 1, отличающийся тем, что второй пептид (б), являющийся цитокином, представляет собой интерлейкин-1β.

3. Слитый полипептид по п. 1, отличающийся тем, что второй пептид (б), являющийся антикоагулянтом, представляет собой гирудин.

4. Слитый полипептид по п. 1, отличающийся тем, что второй пептид (б), являющийся ферментом, представляет собой β-глюкуронидазу.

5. Слитый полипептид по п. 1, отличающийся тем, что второй пептид (б), являющийся антителом, представляет собой одноцепочечное антитело, включающее

трансляционный гибрид доменов V_H или V_L иммуноглобулина.

6. Слитый полипептид по п. 2, отличающийся тем, что интерлейкин-1β характеризуется последовательностью V-Q-G-E-E-S-N-D-K.

7. Химерная ДНК-конструкция, включающая а) первую последовательность ДНК, кодирующую часть олеозина, достаточную для обеспечения переноса к масляному телу, причем данная часть имеет следующую последовательность:

M-M-G-R-D-R-D-Q-Y-Q-M-S-G-R-S-G-D-Y-S-K-S-R-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-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-S-G-G-F-G-I-A-A-I-T-V-F-S-W-I-Y-K-Y-A-T-G-E-H-P-Q-G-S-D-K-L-D-S-A-R-M-K-L-G-S-K-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-T-T,

и б) вторую последовательность ДНК, кодирующую пептид, выбираемый из группы, включающей цитокин, фермент, антикоагулянт, нейропептид, гормон, антитело или предшественник молекулы адгезии.

8. Химерная ДНК-конструкция по п. 7, отличающаяся тем, что второй пептид (б), являющийся цитокином, представляет собой интерлейкин-1β.

9. Химерная ДНК-конструкция по п. 7, отличающаяся тем, что второй пептид (б), являющийся антикоагулянтом, представляет собой гирудин.

10. Химерная ДНК-конструкция по п. 7, отличающаяся тем, что второй пептид (б), являющийся ферментом, представляет собой β-глюкуронидазу.

11. Химерная ДНК-конструкция по п. 7, отличающаяся тем, что второй пептид (б), являющийся антителом, представляет собой одноцепочечное антитело, включающее трансляционный гибрид доменов V_H или V_L иммуноглобулина.

12. Химерная ДНК-конструкция по п. 7, отличающаяся тем, что интерлейкин-1β характеризуется последовательностью V-Q-G-E-E-S-N-D-K.

13. Экспрессирующая кассета, включающая в качестве компонентов, расположенных по направлению транскрипции, 1) регуляторную область с последовательностью промотора, способного обеспечить экспрессию в семени растения; 2) химерную конструкцию ДНК по п. 7 и 3) сегмент терминации транскрипции и транскрипции, причем указанные компоненты оперативно соединены друг с другом и экспрессия указанной химерной последовательности ДНК осуществляется под контролем указанной регуляторной последовательности ДНК.

14. Экспрессирующая кассета по п. 13, отличающаяся тем, что указанная регуляторная последовательность ДНК происходит из состава гена, экспрессируемого в клетке семени зернового растения.

15. Экспрессирующая кассета по п. 13, отличающаяся тем, что указанная регуляторная последовательность ДНК происходит из генома резуховидки *Arabidopsis thaliana*.

Последовательность 1

aa²⁵ - aa²⁶ - V - V - T - L - aa³¹ - ρ -
 A A A T
 aa³⁴ - G - G - aa³⁶ - L - L - aa³⁹ - L - aa⁴¹
 M
 G - I - aa⁴⁴ - L - aa⁴⁶ - aa⁴⁷ - T - L - I -
 S L S V V
 aa⁵¹ - L - aa⁵³ - V - A - T - P - L - aa⁵⁹ - L -
 V V
 L - F - S - P - V - L - V - P - A - A - L - aa⁷³ -
 I I L I
 aa⁷⁴ - aa⁷⁵ - aa⁷⁶ - aa⁷⁷ - aa⁷⁸ - G - F - L -
 G L
 S - S - aa⁸⁷ - G - V - aa⁸⁹ - aa⁹⁰ - L - S -
 T I I T
 aa⁹³ aa⁹⁴ - S - aa⁹⁶ - aa⁹⁷ - aa⁹⁸ - aa⁹⁹ - aa¹⁰⁰ - aa¹⁰¹
 T

Последовательность 2

M-M-G-R-D-R-D-Q-Y-Q-M-S-G-R-G-S-D-Y-S-K-
 S-R-Q-I-A-K-A-A-T-A-V-T-M-G-G-S-L-L-V-L-
 S-S-L-T-L-V-G-T-V-I-A-L-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-S-G-G-F-G-I-A-A-I-T-V-F-S-W-I-Y-K*Y-
 A-T-G-E-H-P-Q-G-S-D-K-L-D-S-A-R-M-K-L-G-
 S-K-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-T-T;

RU 2201449 C2

RU ?201449 C2

Последовательность 3

10
M-M-G-R-D-R-D-Q-Y-Q-M-S-G-R-G-S-D-Y-S-K-
30 40
S-R-Q-I-A-K-A-A-T-A-V-T-A-G-G-S-L-L-V-L-
50 60
S-S-L-T-L-V-G-T-V-I-A-L-T-V-A-T-P-L-L-V-
70 80
I-F-S-P-I-L-V-P-A-L-I-T-V-A-L-L-I-T-G-F-
90 100
L-S-S-G-G-F-G-I-A-A-I-T-V-F-S-W-I-Y-K*Y-
110 120
A-T-G-E-H-P-Q-G-S-D-K-L-D-S-A-R-M-K-L-G-
130 140
S-K-A-Q-D-L-K-D-R-A-Q-Y-Y-G-Q-Q-H-T-G-G-
150
E-H-D-R-D-R-T-R-G-G-Q-H-T-T

Последовательность 4

pp¹ - aa²⁵ - aa²⁶ - V - V- T - L- aa³¹ - p-
A A A T
aa³⁴ - G- G- aa³⁶ - L- L- aa³⁹ - L- aa⁴¹ -
M
G - I - aa⁴⁴ - L- aa⁴⁶ - aa⁴⁷ - T - L - I -
S L S V V
aa⁵¹ - L- aa⁵³ - V- A- T- p- L - aa⁵⁹ - L -
V V
L - F- S- P- V - L- V- P- A- A - L - aa⁷³ -
I I L I
aa⁷⁴ - aa⁷⁵ - aa⁷⁶ - aa⁷⁷ - aa⁷⁸ - G- F- L-
G
S - S- aa⁸⁷ - G- V - aa⁸⁹ - aa⁹⁰ - L - S -
T I I T
aa⁹³ aa⁹⁴ - S - aa⁹⁶ - aa⁹⁷ - aa⁹⁸ - aa⁹⁹ -
T
aa¹⁰⁰ - aa¹⁰¹ - pp²

RU 2201449 C2

RU 2201449 C2

A-C переход здесь дает NcoI сайт

3' ... TC TCA ACA ATG CCA...ОВР моркови
3' ... CC GCA GCA ATC GCG...18 кПа ОВР кукурузы

3' ... TC TCA ACA ATG GCA GAA CGA GGC ACT TAT ...
изменить мутацией на

3' ... TC TCA ACA ATG GCA TGC CGA GGC GCC TAT ...
Sph1

Sph1 n раз Nar1
CTACG CCG CTC GGT CCG CG
GGC GAG GCA GGC CCGC

Липкий конец 1 n раз Рестрикционный сайт n раз Липкий конец 2
CCG CTC GGT CCG CCG CTC GGT
GGC GAG CCA GGC GGC GAG CCA GGC

5' CACACCAGCAACTCTCTGGTAAGC 3'
(положение оснований - 838 до - 814), олигонуклеотид G-VR10
5' CACTGCAGGAACCTCTCTGGTAAGC 3'

Таблица 1

Использование кодонов *A. thaliana*¹ и *B. napus oleosin*²
Codon usage of *A. thaliana*¹ and *B. napus oleosin*²

TTT phe F 5	TCT ser S 8	TAT tyr Y 6	TCT cys C -
TTC phe F 3	TCC ser S 10	TAC tyr Y 11	TGC cys C -
TTA leu L -	TCC ser S 2	TAA och z -	TGA opa z -
TTG leu L 4	TGA ser S -	TAG amb z -	TGG trp w 2
CTT leu L 10	CCT pro p 3	CAT his H 4	CGT arg R 6
CTC leu L 11	CCC pro p 1	CAC his H 6	CGC arg R -
CTA leu L -	CCA pro p 4	CAA gln Q 4	CGA arg R 1
CTG leu L 6	CCG pro p 2	CAG gln Q 17	CGG arg R -
ATT ile I 7	ACT thr T 11	AAT asn N -	AGT ser S 5
ATC ile I 13	ACC thr T 14	AAC asn N 1	AGC ser S 3
ATA ile I 3	ACA thr Y 7	AAA lys K 5	AGA arg R 5
ATG met M 11	ACG thr T 5	AAG lys K 10	AGG arg R 3
GTT val V 11	GCT ala A 17	GAT asp D 8	GGT gly G 0
GTC val V 9	GCC ala A 2	GAC asp D 19	GGC gly G 9
GTA val V -	GCA ala A 10	GAA glu E 2	GGA gly G 14
GTG val V -	GCG ala A 2	GAG gly E 2	GGG gly G 5

1 See Figure 1A. См. фиг. 1A

2 Lee and Huang (1991) Plant Physical 96: 1395-1397.

Результаты экспрессии *)

	Стабильная интеграция Саузерн-блоттинг	Экспрессия олеозин-IL-1-β Нозерн-блоттинг	Определение 20 кДа протеина Бестерн-блоттинг
1	2	3	4
<i>B. parus</i> :			
не трансформировано	-	NT	NT
трансформант № 12	+	NT	NT
Табак			
не трансформировано	-	+	+
трансформант А	+	+	+
трансформант В	+	+	+
трансформант С	+	+	+
трансформант L	+	+	+

*) Стабильная интеграция, определенная, как описано в 1.
 Экспрессию олеозин-IL-1-β слияния определяют по способу J.
 Детектирование предполагаемого протеина слияния олеозин-IL-1-β ведут по способу К.
 NT - не тестировали.

-867 NcoI
 CCATGGCTATACCCAACCTCGGTCTTGGTCACACCAGGAACCTCTCGGTAAGCTAGCTCCACTCCCCAGAAACAACCGGCGCCAAATTGG
 -777 CGGAATTGCTGACCTGAAGACGGAACATCATCGTCGGGTCTTGGGGGATTCGGGGGGAAGATGGGTCAGCTTGGGCTTGAGGACGAGAC
 -687 CGGAATCGAGTCTGTTGAAAGGTTGTTTCATTGGGATTTGTATACGGAGATTGGTCGTCGAGAGGTTTGAGGGAAGGACAAATGGGTTTG
 R1
 -597 GCTCTGGAGAAAGAGAGTGGCGCTTTAGAGAGAGAATTGAGAGGTTTAGAGAGAGATGCGGGCGCGATGACGGGAGGAGAGACGACGAGG
 R2 R1 R2
 -507 ACCTGCATTATCAAAGCAGTGACGTGGTGAATTTGAACTTTTAAGAGGCAGATAGATTTATTATTTGTATCCATTTTCTTCATTGTTC
 -417 TAGAATGTCGGCGGAACAAATTTAAACTAAATCCTAAATTTTCTAATTTTGTTCGCAATAGTGGATATGTTGGCCGATAGAAGGAAT
 -327 CTATTGAAGGGCCAAACCCATACTGACGAGCCCAAAGGTTGTTTTGCGTTTTATGTTTTCGGTTTCGATGCCAACGCCACATTCTGAGCTA
 I
 -237 GGCAAAAAACAACCGTGTCTTTGAATAGACTCCTCTCGTTAACACATGCAGCGGCTGCATGGTGACGCCATTAACACGTTGGCCTACAAAT
 -147 GCATGATGCTCCATTGACACGCTGACTTCTCGTCTCCTTTCTTAATATATCTAACAAACACTCCTACCTCTTCCAAAATATATACACATC
 M A D T A R G T H H D
 -57 TTTTGTATCAATCTCTCATTCAAATCTCATTCTCTCTAGTAAACAAGAACAATAAATGGCGGATACAGCTAGAGGAACCCATCAGGAT
 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
 34 ATCATCGGCAGAGACCAGTACCCGATGATGGGCGGAGACCAGACCCAGTACCAGATGTCGGGACGAGGATCTGACTACTCCAAGTCTAGG
 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 CAGATTGCTAAAGCTGCAACTGCTGTACAGTGGTGGTTCCTCTCTGTTCTCTCCAGCCTTACCCTTGTGGAAGTGTATAGCTTTG
 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 ACTGTTGCAACACCTCTGCTCGTTATCTTCAGCCCAATCCTTGTCCCGGCTCTCATCACAGTTGCACTCCTCATCACCAGTTTCTCTCC
 S G G F G I A A I T V F S W I Y K
 304 TCTGGAGGTTTGGCATTGCCGCTATAACCGTTTTCTCTGGATTTACAAGtaagcacacatttatcatcttacttcataattttgtgca
 394 atatgtgcatgcatggttgagccagtagctttggatcaatTTTTTGGTcgaataacaaatgtaacaataagaatgcaattctag
 484 gaacatttgggtaactaaatcgaatattgacctagcttagcttgaatgtgtctgtgtatcatcttatataggtaaaatgcttggtatga
 Y A T G E H P Q G S D K L D S A R M K L G S K
 574 tacctattgattgtgaatagGTACGCAACGGGAGAGCACCCACAGGGATCAGACAAAGTTGGACAGTGAAGTTGGGAAGCAAA
 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 GCTCAGGATCTGAAAGACAGAGCTCAGTACTACGGACAGCAACATACTGGTGGGAACATGACCGTGACCGTACTCGTGGTGGCCAGCAC
 T T *
 754 ACTACTTAAGTTACCCCACTGATGTCATCGTCAATAGTCCAAATAACTCCAATGTCCGGGAGTTAGTTTATGAGGAAATAAAGTGTITAGAAAT
KpnI
 844 TTGATCAGGGGGAGATAATAAAGCCGAGTTTGAATCTTTTGTATAAGTAATGTTTATGTGTGTTTCTATATGTTGTCAAATGGTACC

RU 2201449 C2

RU 2201449 C2

Фиг. 1А

carrot 19KDa
 МОРКОВЬ
 Maize 18KDa
 кукуруза
 Maize 16KDa
 кукуруза
 Arab. 17KDa

MAERGTYAHQVQVHPQQTANQPGGVKS
 LIPKNSPSTSQV
 MADRRDRSGIYGGAHATYGGQQQQGGGGRPMGEQVKKG
 MLHDKGPTASQA
 RGGGGYGD LORGGGMHGEAA
 QQQKQCAMMTA
 MMGRDRDQYQMS
 GRGSDYSKSRQI

LAVVTLVPVGGTLLFLAGITLVGTIIGLAVATPLFLL
 FSPVLVPAALTIGLAVTGLGSGAFGLTGLSSLSWVLS-YF
 LTVATLFPPLGGLLLVSGIALTASVVGLAVATPVFLIF
 SPVLVPAALLIGTAVMGFLTSGALGLGGLSSLTCLAN-TA
 LKAATAATTGGSMVLVSLILAGTVIALIVATPVLVI
 FSPVLVPAALALALMAAGFVTSGGLGVAALSVFSPWYK-YL
 AKAATAVTAGGSLVLLSLTLVGTVIALTVATPLLVI
 FSPILVPALITVALLITGFLSSGGFGIAAITVFSWIYK*YL

N-terminal hydrophilic
 domain
 - терминальный гидрофильный
 домен

Central hydrophobic
 domain
 Центральный гидрофобный
 домен

*=Intron splice site
 сайт интронного сплайсинга

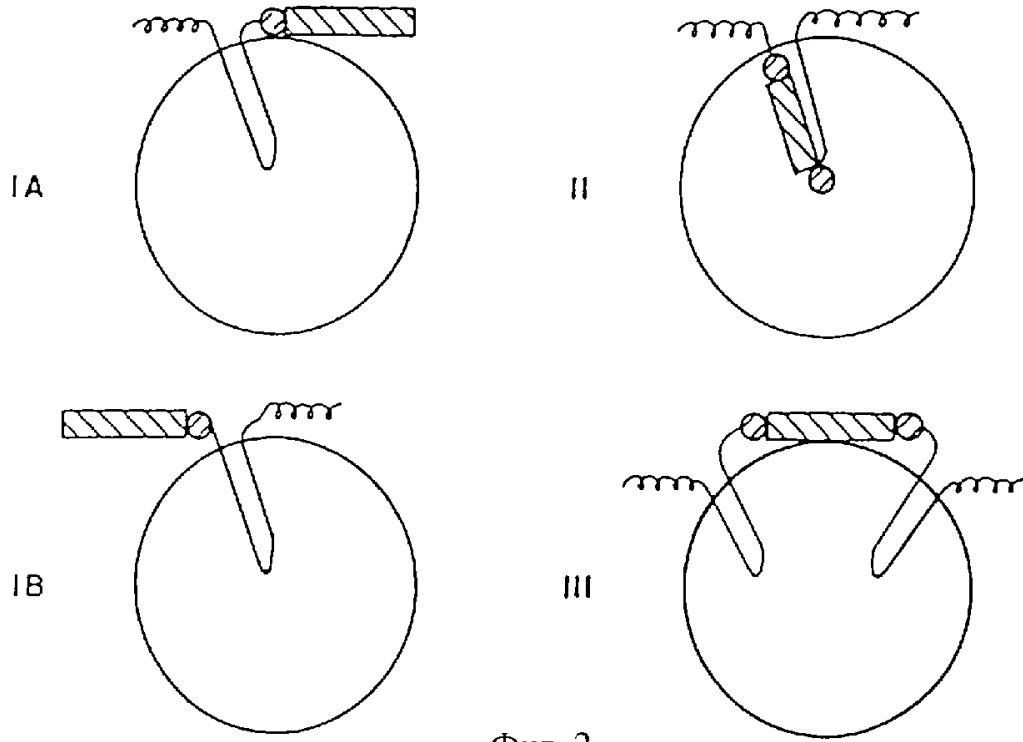
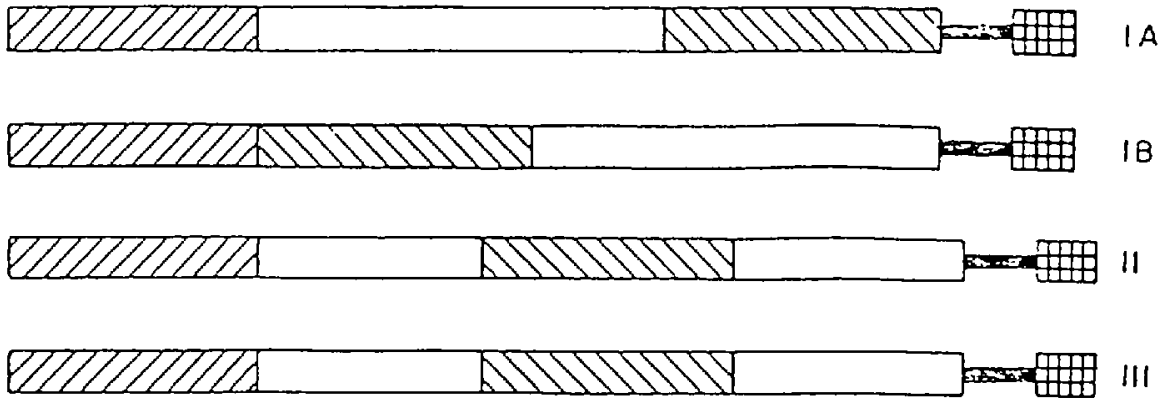
RQASQRVPDQIELAKKRAQAMAAYAGQKTKEVGDTIQSKAAQAQDITATTGRDTRSTARDTSRT
 RQAFQRTPDYVEEARRRMAEAAAQAGHKTAQAGQAIQGRAQEAGTGGGAGRASS
 TGKHPGADQLDHAARLASKARDIKDAAQHRIDQAQGS
 LIENPQGSCLKDSARMKLGSKAQDLKDRAYYQQHTGWENHRRDRTRGGQHTT

C-terminal amphipathic
 domain
 С-терминальный амфипатический
 домен

Фиг. 1В

RU 2201449 C2

RU 2201449 C2



Фиг. 2

Xho I

....AGC ACT GCT CGA GAC ACT TCA AGG ACT....Carrot OBP
(see Hatzopoulos et Ser Thr Ala Arg Asp Thr Ser Arg Thr.... al. 1990)

3" ..AGC ACT GC Xho I TCGA CCG CTC GGT CCG GC Nco I
TCG TGA CGAGCT GGC GAG CCA GGC CGGTAC

Pro Leu Gly Pro

Collagenase recognition motif
Фрагмент распознавания коллагеназы
Фиг. 3

Конструирование гена слияния с использованием протеазных/ например, коллагеназных/ , кодирующих мишеневый фрагмент линкеров, клонируемых в E. coli совместимые плазмиды

↓
Встраивание конструкции в репликон хозяина, содержащий T-ДНК границы / то есть Agrobacterium бинарный вектор/

↓
Трансформация клеток растения с использованием листовых, стеблевых, семя-дольных или петиольных эксплататов

↓
Регенерация трансгенных растений

↓
Получение семян

↓
Измельчение семян в водном экстракционном буфере /Taylor et al., 1990/

↓
Центрифугирование или другой способ выделения масляных тел

↓
Суспендирование промытых масляных тел в коллагеназном аналитическом буфере и добавление 5 единиц очищенной коллагеназы

↓
Центрифугирование для выделения остатка масляных тел

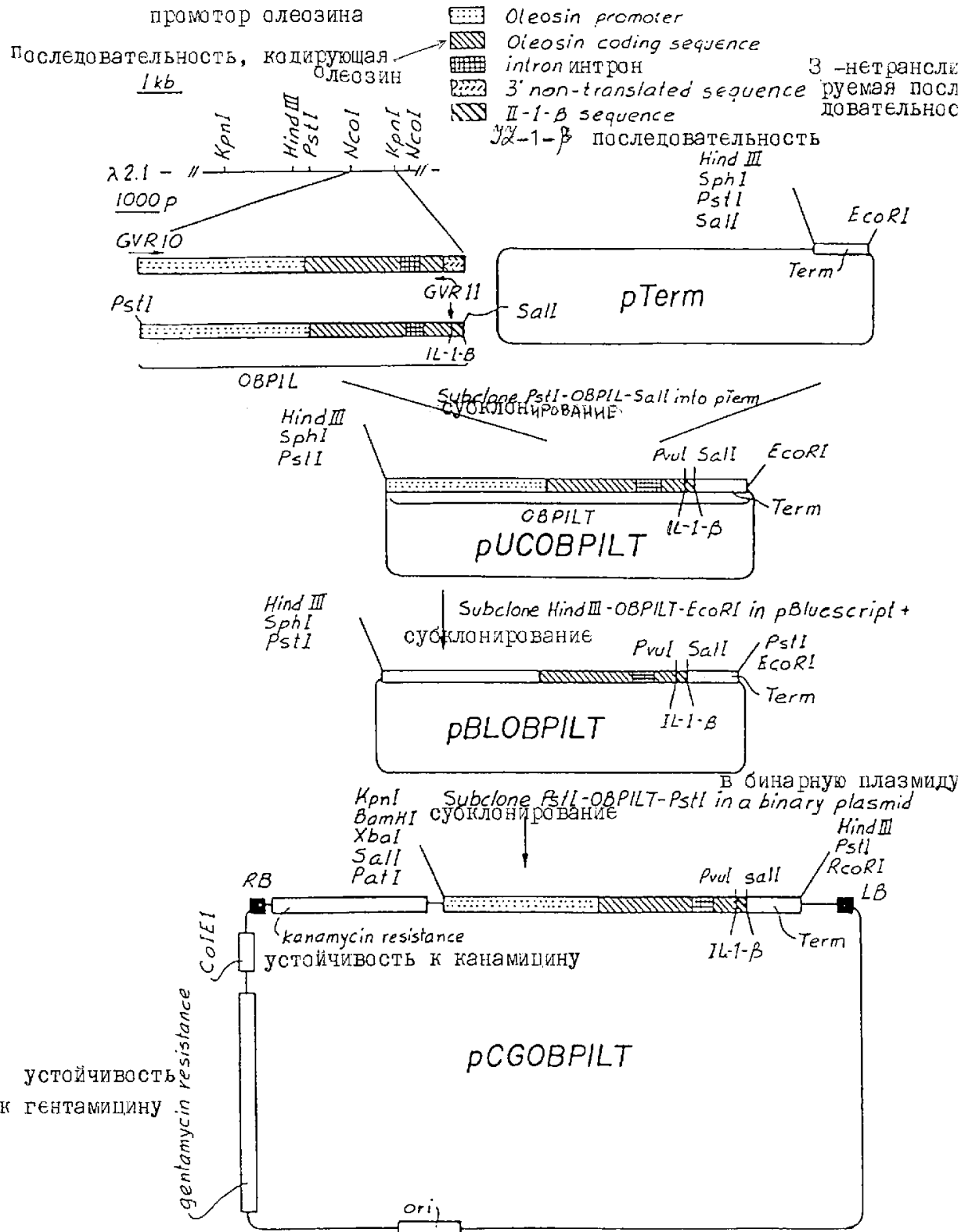
↓
Выделение высвободившихся пептидов после обработки протеазой за счет осаждения сульфатом аммония, с помощью хроматографической колонки и т.д.

↓
Проведение биологического анализа выделенного рекомбинатного пептида

Фиг. 4

RU 2201449 C2

RU 2201449 C2



Фиг. 5

RU 2201449 C2

RU 2201449 C2

INTERLEUKIN 1- β gvr11 (71-mer)
 ИНТЕРЛЕЙКИН

a) A. thaliana oleosin Factor Xa Фактоор Ха
 ОЛЕОЗИН
 gly gly gln his thr thr ala ile glu gly arg val gln
 gly glu glu ser asn asp lys OCH val asp
 GGT GGC CAG CAC ACT ACT GCT ATC GAA GGG AGA GTT CAG
 PvuI
 GGA GAA GAA TCT AAC GAC AAG TAA GTC GAC GG
 Sall

b) 3'CCA CCG GTC GTG TGA TGA CGC TAG CTT CCC TCT CAA GTC
 [.....
 CCT CTT CTT AGA TTG CTG TTC ATT CAG CTG CC 5'
]

Фиг. 6

RU 2201449 C2

RU 2201449 C2

PstI

CACTGCAGGAACTCTCTGGTAAGCTAGCTCCACTCCCAGAAACAACCGGCGCCAAATTGCC
GGAATTGCT
GACCTGAAGACGGAACATCATCGTCGGGTCCTTGGGCGATTGCGGGCGGAAGATGGGTCAGCT
TGGGCTTGAG
GACGAGACCCGAATCGAGTCTGTTGAAAGGTTGTTCAATTGGGATTGTATACGGAGATTGGTC
GTCGAGAGG
TTTGAGGGAAAGGACAAATGGGTTTGGCTCTGGAGAAAGAGAGTGCGGCTTTAGAGAGAGAA
TTGAGAGGTT
TAGAGAGAGATGCGGGCGCGATGACGGGAGGAGAGACGACGAGGACCTGCATTATCAAAGCA
GTGACGTGGT
GAAATTTGGAACTTTTAAGAGGCAGATAGATTTATTATTTGTATCCATTTTCTTCATTGTTT
TAGAATGTCTG
CGGAACAAATTTTAAACTAAATCCTAAATTTTTCTAATTTTGTGGCCAATAGTGGATATGT
GGGCCGTATA
GAAGGAATCTATTGAAGGCCCAAACCCATACTGACGAGCCCAAAGGTTTCGTTTTGCGTTTTA
TGTTTCGGTT
CGATGCCAACGCCACATTCTGAGCTAGGCCAAAAACAACGTGTCTTTGAATAGACTCCTCT
CGTTAACACA
TGCAGCGGCTGCATGGTGACGCCATTAACACGTGGCCTACAATTGCATGATGTCTCCATTGA
CACGTGACTT
CTCGTCTCCTTTCTTAATATATCTAACAAACACTCCTACCTCTTCCAAAATATATACACATC
TTTTTGATCA
ATCTCTCATTCAAAATCTCATTCTCTCTAGTAAACAAGAACAATAAATGGCGGATACAGCT
AGAGGAACCC
ATCACGATATCATCGGCAGAGACCAGTACCCGATGATGGGCCGAGACCGAGACCAGTACCAG
ATGTCCGGAC
GAGGATCTGACTACTCCAAGTCTAGGCAGATTGCTAAAGCTGCAACTGCTGTACAGCTGGT
GGTCCCTCC
TTGTTCTCTCCAGCCTTACCCTTGTTGGAAGTGTACATAGCTTTGACTGTTGCAACACCTCTG
CTCGTTATCT
TCAGCCCAATCCTTGTCGCCGGCTCTCATCACAGTTGCACTCCTCATCACCGGTTTTCTTTCC
TCTGGAGGGT
TTGGCATTGCCGCTATAACCGTTTTCTCTTGGATTTACAAGTAAGCACACATTTATCATCTT
ACTTCATAAT
TTTGTGCAATATGTGCATGCATGTGTTGAGCCAGTAGCTTTGGATCAATTTTTTTGGTCGAA
TAACAAATGT
AACAAATAAGAAATTGCAAATCTAGGGAACATTTGGTTAACTAAATACGAAATTTGACCTAG
CTAGCTTGAA
TGTGTCTGTGTATATCATCTATATAGGTAAAATGCTTGGTATGATACCTATTGATTGTGAAT
AGGTACGCAA
CGGGAGAGCACCCACAGGGATCAGACAAGTTGGACAGTGCAAGGATGAAGTTGGGAAGCAAA
GCTCAGGATC
TGAAAGACAGAGCTCAGTACTACGGACAGCAACATACTGGTTGGGAACATGACCGTGACCGT
ACTCGTGGTG
GCCAGCACACTACTGCGATCGAAGGGAGAGTTCAGGGAGAAGAATCTAACGACAAGTAAGTC
GACTCTAG
ACGGATCTCCCgategttcaaacatttggcaataaagtttcttaagattgaatcctggtgccc
ggtcttgcg
tgattatcatataatttctggttgaattacgttaagcatgtaataattaacatgtaatgcatg
acgttattta
tgagatgggtttttatgattagagtcccgcaattatacatttaatacgcgatagaaaacaaa
atatagecgcg
caactaggataaattatecgcgcgcggtgtcatctatgttactagatcGGAATTC

EcoRI

Фиг. 7

RU 2201449 C2

RU 2201449 C2