



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21)(22) Заявка: 2013150784, 13.04.2012

(24) Дата начала отсчета срока действия патента:
13.04.2012Дата регистрации:
05.10.2017

Приоритет(ы):

(30) Конвенционный приоритет:
15.04.2011 US 61/475,921

(43) Дата публикации заявки: 20.05.2015 Бюл. № 14

(45) Опубликовано: 05.10.2017 Бюл. № 28

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 15.11.2013(86) Заявка РСТ:
US 2012/033458 (13.04.2012)(87) Публикация заявки РСТ:
WO 2012/142371 (18.10.2012)Адрес для переписки:
129090, Москва, ул. Б. Спасская, 25, строение 3,
ООО "Юридическая фирма Городисский и
Партнеры"

(72) Автор(ы):

ЛАРРИНУА Игнасио Марио (US),
МЕРЛО Дональд Дж. (US),
РЕДДИ Авуту С. (US),
ТХИРУМАЛАЙСВАМИСЕКХАР Арвинд
Кумар (US),
ВУСЛИ Аарон Тодд (US)

(73) Патентообладатель(и):

ДАУ АГРОСАЙЕНСИЗ ЭлЭлСи (US)

(56) Список документов, цитированных в отчете
о поиске: WO 1998056921 A1 от 17.12.1998.
EP 1203077 B1 от 07.01.2009. US 20100223694
A1 от 02.09.2010. US 0006939711 B2 от
06.09.2005. RU 2006137474 A от 27.04.2008.

(54) СИНТЕТИЧЕСКИЕ ГЕНЫ

(57) Реферат:

Изобретение относится к области биохимии, в частности к синтетической ДНК для экспрессии белковых токсинов Cgu. Также раскрыты ДНК-конструкция для экспрессии белковых токсинов Cgu и трансгенное растение, имеющее устойчивость к насекомым-вредителям,

чувствительным к белковым токсинам Cgu. Раскрыт способ борьбы с насекомыми-вредителями зерна или семян, чувствительными к белковым токсинам Cgu. Изобретение позволяет бороться с насекомыми-вредителями. 4 н.п. ф-лы, 29 табл., 17 пр.



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.
C12N 15/82 (2006.01)
C12N 5/04 (2006.01)
C12N 5/10 (2006.01)

(12) **ABSTRACT OF INVENTION**

(21)(22) Application: **2013150784**, 13.04.2012

(24) Effective date for property rights:
13.04.2012

Registration date:
05.10.2017

Priority:

(30) Convention priority:
15.04.2011 US 61/475,921

(43) Application published: **20.05.2015** Bull. № 14

(45) Date of publication: **05.10.2017** Bull. № 28

(85) Commencement of national phase: **15.11.2013**

(86) PCT application:
US 2012/033458 (13.04.2012)

(87) PCT publication:
WO 2012/142371 (18.10.2012)

Mail address:
**129090, Moskva, ul. B. Spasskaya, 25, stroenie 3,
OOO "Yuridicheskaya firma Gorodisskij i Partnery"**

(72) Inventor(s):

**LARRINUA Ignacio Mario (US),
MERLO Donald Dzh. (US),
REDDI Avutu S. (US),
TKHIRUMALAISVAMISEKKHAR Arvind
Kumar (US),
VUSLI Aaron Todd (US)**

(73) Proprietor(s):

DAU AGROSAJENSIZ EIEISi (US)

(54) **SYNTHETIC GENES**

(57) Abstract:

FIELD: biotechnology.

SUBSTANCE: invention relates to biochemistry, in particular to a synthetic DNA for expression of Cry protein toxins. A DNA construct for Cry protein toxins expression and a transgenic plant having resistance to insect pests susceptible to Cry protein toxins are

disclosed as well. A method to control insect pests of grain or seeds sensitive to Cry protein toxins is disclosed.

EFFECT: invention allows to deal with insect pests.
4 cl, 29 tbl, 17 ex

C 2
1 7 5 2 2 6 2 6 3 2 5 7 1
R U

R U
2 6 3 2 5 7 1
C 2

УРОВЕНЬ ТЕХНИКИ

Было обнаружено, что для достижения необходимых уровней экспрессии гетерологичных белков в трансгенных растениях целесообразно изменять нативную кодирующую последовательность ДНК, иногда называемую последовательностью дикого типа или исходной последовательностью, различными способами, например, таким образом, чтобы использование кодонов точнее соответствовало использованию кодонов для вида растения-хозяина, и/или чтобы содержание G+C в кодирующей последовательности точнее соответствовало уровню G+C, обычно имеющемуся в кодирующих последовательностях для вида растения-хозяина, и/или чтобы определенные последовательности, которые дестабилизируют мРНК, были удалены. Например, экспрессия в растениях кристаллических белковых токсинов *Bacillus thuringiensis* (B.t.) против насекомых было улучшено с применением одного или более указанных способов. См., например, патент США No. 5380301, патент США No. 5625136, патент США No. 6218188, патент США No. 6340593, патент США No. 6673990, патент США No. 7741118. Вырожденность кодонов позволяет создавать синтетические последовательности ДНК, которые кодируют представляющий интерес белок с использованием кодонов, которые отличаются от используемых в исходной кодирующей последовательности ДНК.

Что касается удаления последовательностей, которые могут дестабилизировать мРНК, в патенте США No. 7741118 изложен список из 16 последовательностей сигналов полиаденилирования (столбец 15, таблица II) и требования по снижению количества таких последовательностей в синтетических кодирующих последовательностях, которые предназначены для экспрессии в растениях. Последовательности сигналов полиаденилирования, приведенные в таблице II US 7741118, приведены ниже в таблице 1:

25

Последовательности сигналов полиаденилирования, приведенные в таблице II US 7741118							
1	AATAAA	6	ATACTA	11	ATACAT	16	CATAAA
2	AATAAT	7	ATAAAA	12	AAAATA		
3	AACCAA	8	ATGAAA	13	ATTAAA		
4	ATATAA	9	AAGCAT	14	AATTAA		
5	AATCAA	10	ATTAAT	15	AATACA		

30

В US 7741118 также, предпочтительно, требуется удаление последовательности АТТТА (известной как последовательность Шоу-Камена), поскольку она была идентифицирована как потенциально дестабилизирующая мРНК.

35

В отличие от изложенного в US 7741118, мы обнаружили, что снижение количества последовательностей сигналов полиаденилирования, идентифицированных выше в таблице 1, не является ни необходимым, ни достаточным для обеспечения повышенной экспрессии синтетических генов в растениях.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

40

Ниже в таблице 2 идентифицировано 20 потенциальных последовательностей сигналов полиаденилирования, которые часто появляются в генах кукурузы.

45

Потенциальные последовательности сигналов полиаденилирования, найденные в генах кукурузы							
1	АТАГАТ	6	ТАТТТТ	11	ТААТАА	16	ТАГТАТ
2	ТТГТТТ	7	ТТТТТТ	12	АТТТАТ	17	ТГТТТГ
3	ТТТТГТ	8	АТТТТТ	13	ТАГАТТ	18	ТТАГАТ
4	ТГТТТТ	9	ТТАТТТ	14	ТТТТАТ	19	ТГТААТ
5	ТАГАТА	10	ТТТАТТ	15	АТАТТТ	20	АААТАА

Ниже в таблице 3 идентифицировано 20 потенциальных последовательностей сигналов полиаденилирования, которые часто появляются в генах сои.

Потенциальные последовательности сигналов полиаденилирования, найденные в генах сои							
1	АТТТТТ	6	ТТТГАТ	11	АААТТТ	16	АТАТАТ
2	ТАТТТТ	7	ААТТТТ	12	АААТАА	17	АТТАТТ
3	ТТАТТТ	8	ТТТТТА	13	АТАТТТ	18	АТТТТА
4	ТТТАТТ	9	ТААТТТ	14	ТТТГТТ	19	ТТТААТ
5	ТТТТТТ	10	ТГААТТ	15	ТТГТТТ	20	ТТТТАА

Настоящее изобретение предоставляет синтетическую последовательность ДНК для экспрессии представляющего интерес белка в клетках кукурузы, которая содержит:

а) оптимизированную по кодомам последовательность ДНК, кодирующую представляющий интерес белок,

б) по меньшей мере, одну последовательность сигнала полиаденилирования, выбранную из группы, состоящей из класса I и класса II, где

класс I выбирают из группы, состоящей из ААТААА, ААТААТ, ААССАА, АТАТАА, ААТСАА, АТАСТА, АТАААА, АТГААА, ААГСАТ, АТТААТ, АТАСАТ, ААААТА, АТТААА, ААТТАА, ААТАСА и САТААА; и

класс II выбирают из группы, состоящей из АТАТАТ, ТТГТТТ, ТТТТГТ, ТГТТТТ, ТАТАТА, ТАТТТТ, ТТТТТТ, АТТТТТ, ТТАТТТ, ТТТАТТ, ТААТАА, АТТТАТ, ТАТАТТ, ТТТТАТ, АТАТТТ, ТАГТАТ, ТГТТТГ, ТТАТАТ, ТГТААТ и АААТАА; и

где указанная оптимизированная по кодомам последовательность ДНК содержит, по меньшей мере, одну последовательность сигнала полиаденилирования из класса II, и где указанная синтетическая последовательность ДНК содержит меньше последовательностей сигналов полиаденилирования класса II, чем нативная последовательность ДНК белка, и содержит такое же количество последовательностей сигналов полиаденилирования класса I по сравнению с нативной последовательностью ДНК.

Настоящее изобретение также предоставляет синтетическую последовательность ДНК для экспрессии представляющего интерес белка в клетках сои, которая содержит:

а) оптимизированную по кодомам последовательность ДНК, кодирующую представляющий интерес белок,

б) по меньшей мере, одну последовательность сигнала полиаденилирования, выбранную из группы, состоящей из класса I и класса III, где

класс I выбирают из группы, состоящей из ААТААА, ААТААТ, ААССАА, АТАТАА, ААТСАА, АТАСТА, АТАААА, АТГААА, ААГСАТ, АТТААТ, АТАСАТ, ААААТА, АТТААА, ААТТАА, ААТАСА и САТААА; и

класс III выбирают из группы, состоящей из АТТТТТ, ТАТТТТ, ТТАТТТ, ТТТАТТ, ТТТТТТ, ТТТТАТ, ААТТТТ, ТТТТТА, АТАТАТ, ТААТТТ, ТТААТТ, АААТТТ, АААТАА, АТАТТТ, ТТТГТТ ТТГТТТ, АТТАТТ, АТТТТА, ТТТААТ и ТТТТАА, и

где указанная оптимизированная по кодомам последовательность ДНК содержит, по меньшей мере, одну последовательность сигнала полиаденилирования из класса III, и где указанная синтетическая последовательность ДНК содержит меньше последовательностей сигналов полиаденилирования из класса III, чем нативная последовательность ДНК белка, и содержит такое же количество последовательностей сигналов полиаденилирования класса I по сравнению с нативной последовательностью ДНК.

Изобретение также предоставляет способ создания синтетической последовательности

ДНК, которая кодирует представляющий интерес белок, который включает (а) сначала, использование аминокислотной последовательности представляющего интерес белка, полученной из встречающегося(ихся) в природе полипептида(ов), кодируемого(ых) нативной(ыми) последовательностью(ями), которые содержат, по меньшей мере, одну последовательность сигнала полиаденилирования, приведенную в таблице 2, и (б) создание синтетической последовательности ДНК, которая кодирует указанную аминокислотную последовательность и содержит меньше последовательностей сигналов полиаденилирования, приведенных в таблице 2, по сравнению с соответствующей кодирующей последовательностью нативной(ых) последовательности(ей), и содержит такое же количество последовательностей сигналов полиаденилирования, приведенных в таблице 1.

В другом варианте осуществления изобретение предоставляет способ создания синтетической последовательности ДНК, которая кодирует представляющий интерес белок, который включает (а) сначала, использование аминокислотной последовательности представляющего интерес белка, полученной из встречающегося(ихся) в природе полипептида(ов), кодируемого(ых) нативной(ыми) последовательностью(ями), которые содержат, по меньшей мере, одну последовательность сигнала полиаденилирования, приведенную в таблице 3, и (б) создание синтетической последовательности ДНК, которая кодирует указанную аминокислотную последовательность и содержит меньше последовательностей сигналов полиаденилирования, приведенных в таблице 3, по сравнению с соответствующей кодирующей последовательностью нативной(ых) последовательности(ей), и содержит такое же количество последовательностей сигналов полиаденилирования, приведенных в таблице.

В некоторых вариантах осуществления синтетические последовательности ДНК, предоставляемые в изобретении, не содержат последовательностей сигналов полиаденилирования, приведенных в таблице 2 и/или таблице 3, или количество последовательностей сигналов полиаденилирования, идентифицированных в таблице 2 и/или таблице 3, снижено насколько это возможно при условии поддержания того же количества последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, и поддержания последовательностей из таблицы 1 в их исходных положениях в последовательности.

В некоторых вариантах осуществления синтетические последовательности ДНК, предоставляемые изобретением, кодируют инсектицидный белок, необязательно, полученный из *Bacillus thuringiensis*, а также последовательности ДНК, полезные для устойчивости к гербицидам, недостатку воды и/или устойчивости к тепловому стрессу, модификации масел для здорового питания и для маркерных генов и селективируемых маркерных генов трансформации.

Синтетические последовательности ДНК по изобретению могут быть применены в ДНК-конструкции для экспрессии представляющего интерес белка, при этом конструкция содержит 5'-нетранслируемую последовательность, синтетическую последовательность ДНК по изобретению и 3'-нетранслируемую область, и указанная 5'-нетранслируемая последовательность содержит промотор, функциональный в растениях, и указанная 3'-нетранслируемая последовательность содержит сигнал терминации транскрипции и полиаденилирования.

Изобретение также предоставляет трансгенное растение, содержащее синтетические последовательности ДНК по изобретению.

Также представлен способ борьбы с насекомыми-вредителями в растении, который

включает экспрессию синтетической последовательности ДНК по изобретению в растении, где синтетическая последовательность ДНК кодирует токсины против насекомых, например, белок Cry Bacillus thuringiensis.

5 Также представлен способ для устойчивости к гербицидам у растения, который включает экспрессию синтетической последовательности ДНК по изобретению в растении, где синтетическая последовательность ДНК кодирует известный фермент устойчивости к гербицидам, например, такой фермент, как арилоксиалканоат диоксигеназа (AAD1), см. WO/2005/107437, или фосфинотрицин ацетилтрансфераза, или 5-энолпирувилшикимат-3-фосфат синтаза.

10 Также представлен способ для изменения масляных профилей в растении, который включает экспрессию одной или более синтетических последовательностей ДНК по изобретению в растении, где синтетическая последовательность ДНК кодирует один или более известных ферментов для изменения масляных профилей в растениях, например, десатуразу жирных кислот.

15 Также представлен способ для устойчивости к стрессу в растении, который включает экспрессию синтетической последовательности ДНК по изобретению, где синтетическая последовательность ДНК кодирует известные гены устойчивости к стрессу недостатка воды и/или тепловому стрессу, например, для ассоциированного со стрессом белка (SAP1); патентная публикация США NO:2010/0275327, и для белков 1-Cys пероксиредоксина (Per1) (Mowla, et al, 2002, Planta 215:716-726).

20 Также представлен способ добавления генов-репортеров к растению, который включает экспрессию синтетической последовательности ДНК по изобретению в растении, где синтетическая последовательность ДНК кодирует известный маркерный белок трансформации, являющийся функциональным в растениях, например, зеленый флуоресцентный белок (GFP) или фермент бета-глюкуронидазу.

25 Также представлен способ защиты от вредителей для зерен или семян, который включает получение указанных зерен или семян у растений, содержащих синтетический ген по изобретению, который экспрессирует токсин против насекомых, и способ защиты от вредителей для муки крупного и мелкого помола, который включает получение 30 указанной муки крупного и мелкого помола из зерна, содержащего синтетический ген по изобретению, который экспрессирует токсин против насекомых.

35 Также представлена композиция, полученная из трансгенных растений, содержащая синтетическую ДНК по изобретению, где указанная композиция представляет собой товарный продукт, выбранный из группы, состоящей из муки крупного помола, муки мелкого помола, белкового концентрата и масла.

В некоторых случаях количество сигналов полиаденилирования, приведенных в таблице 1, может поддерживаться в синтетических последовательностях ДНК по изобретению путем удаления экземпляров ААТААА и их замены на другие последовательности сигналов полиаденилирования, приведенные в таблице 1. Пример 40 указанного показан в примере 1, SEQ ID NO:5.

ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

SEQ ID NO:1 представляет собой нативную последовательность ДНК, кодирующую ядерный токсин Cry1Fa Bacillus thuringiensis.

45 SEQ ID NO:2 представляет собой последовательность ядерного токсина Cry1Fa Bacillus thuringiensis.

SEQ ID NO:3 представляет собой синтетическую последовательность ДНК, кодирующую ядерный токсин Cry1Fa Bacillus thuringiensis с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности

из таблицы 1.

SEQ ID NO:4 представляет собой последовательность ядерного токсина Cry1Fa *Bacillus thuringiensis*.

5 SEQ ID NO:5 представляет собой синтетическую последовательность ДНК по изобретению, кодирующую ядерный токсин Cry1Fa *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой удалены последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1.

10 SEQ ID NO:6 представляет собой последовательность ядерного токсина Cry1Fa *Bacillus thuringiensis*.

SEQ ID NO:7 представляет собой нативную последовательность ДНК, кодирующую токсин Cry34Ab1 *Bacillus thuringiensis*.

SEQ ID NO:8 представляет собой последовательность токсина Cry34Ab1 *Bacillus thuringiensis*.

15 SEQ ID NO:9 представляет собой синтетическую последовательность ДНК, кодирующую токсин Cry34Ab1 *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1.

20 SEQ ID NO:10 представляет собой последовательность токсина Cry34Ab1 *Bacillus thuringiensis*.

SEQ ID NO:11 представляет собой синтетическую последовательность ДНК по изобретению, кодирующую токсин Cry34Ab1 *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой удалены последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1.

SEQ ID NO:12 представляет собой последовательность токсина Cry34Ab1 *Bacillus thuringiensis*.

SEQ ID NO:13 представляет собой нативную последовательность ДНК, кодирующую токсин Cry35Ab1 *Bacillus thuringiensis*.

30 SEQ ID NO:14 представляет собой последовательность токсина Cry35Ab1 *Bacillus thuringiensis*.

SEQ ID NO:15 представляет собой синтетическую последовательность ДНК, кодирующую токсин Cry35Ab1 *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1.

SEQ ID NO:16 представляет собой последовательность токсина Cry35Ab1 *Bacillus thuringiensis*.

40 SEQ ID NO:17 представляет собой синтетическую последовательность ДНК по изобретению, кодирующую токсин Cry35Ab1 *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой удалены последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1.

SEQ ID NO:18 представляет собой последовательность токсина Cry35Ab1 *Bacillus thuringiensis*.

SEQ ID NO:19 представляет собой нативную последовательность ДНК, кодирующую ядерный токсин Cry1Ab1 *Bacillus thuringiensis*.

SEQ ID NO:20 представляет собой последовательность ядерного токсина Cry1Ab1 *Bacillus thuringiensis*.

SEQ ID NO:21 представляет собой синтетическую последовательность ДНК, кодирующую ядерный токсин Cry1Ab1 *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1.

5 SEQ ID NO:22 представляет собой последовательность ядерного токсина Cry1Ab1 *Bacillus thuringiensis*.

SEQ ID NO:23 представляет собой синтетическую последовательность ДНК по изобретению, кодирующую ядерный токсин Cry1Ab1 *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой удалены
10 последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1.

SEQ ID NO:24 представляет собой последовательность ядерного токсина Cry1Ab1 *Bacillus thuringiensis*.

15 SEQ ID NO:25 представляет собой нативную последовательность ДНК, кодирующую ядерный токсин Cry1Ca *Bacillus thuringiensis*.

SEQ ID NO:26 кодирует последовательность ядерного токсина Cry1Ca *Bacillus thuringiensis*.

SEQ ID NO:27 представляет собой синтетическую последовательность ДНК, кодирующую ядерный токсин Cry1Ca *Bacillus thuringiensis* с использованием кодонов,
20 оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1.

SEQ ID NO:28 кодирует последовательность ядерного токсина Cry1Ca *Bacillus thuringiensis*.

25 SEQ ID NO:29 представляет собой синтетическую последовательность ДНК по изобретению, кодирующую ядерный токсин Cry1Ca *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой удалены последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1.

30 SEQ ID NO:30 кодирует последовательность ядерного токсина Cry1Ca *Bacillus thuringiensis*.

SEQ ID NO:31 представляет собой нативную последовательность ДНК, кодирующую токсин Cry6Aa *Bacillus thuringiensis*.

SEQ ID NO:32 представляет собой последовательность токсина Cry6Aa *Bacillus thuringiensis*.

35 SEQ ID NO:33 представляет собой синтетическую последовательность ДНК, кодирующую токсин Cry6Aa *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1.

40 SEQ ID NO:34 представляет собой последовательность токсина Cry6Aa *Bacillus thuringiensis*.

SEQ ID NO:35 представляет собой синтетическую последовательность ДНК по изобретению, кодирующую токсин Cry6Aa *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой удалены последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы
45 1.

SEQ ID NO:36 представляет собой последовательность токсина Cry6Aa *Bacillus thuringiensis*.

SEQ ID NO:37 представляет собой нативную последовательность ДНК, кодирующую

белок AAD1 *Sphingobium herbicidovorans*.

SEQ ID NO:38 представляет собой последовательность белка AAD1 *Sphingobium herbicidovorans*.

5 SEQ ID NO:39 представляет собой синтетическую последовательность ДНК, кодирующую белок AAD1 *Sphingobium herbicidovorans* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1 и таблицы 2.

SEQ ID NO:40 представляет собой последовательность белка AAD1 *Sphingobium herbicidovorans*.

10 SEQ ID NO:41 представляет собой синтетическую последовательность ДНК по изобретению, кодирующую белок AAD1 *Sphingobium herbicidovorans* с использованием кодонов, оптимизированных для кукурузы, и в которой удалены последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1.

15 SEQ ID NO:42 представляет собой последовательность белка AAD1 *Sphingobium herbicidovorans*.

SEQ ID NO:43 представляет собой нативную последовательность ДНК, кодирующую белок дельта-9-десатуразы жирных кислот *Aspergillus nidulans*.

20 SEQ ID NO:44 представляет собой последовательность белка дельта-9-десатуразы жирных кислот *Aspergillus nidulans*.

SEQ ID NO:45 представляет собой синтетическую последовательность ДНК, кодирующую белок дельта-9-десатуразы жирных кислот *Aspergillus nidulans* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1 и таблицы 2.

25 SEQ ID NO:46 представляет собой последовательность белка дельта-9-десатуразы жирных кислот *Aspergillus nidulans*.

30 SEQ ID NO:47 представляет собой синтетическую последовательность ДНК по изобретению, кодирующую белок дельта-9-десатуразы жирных кислот *Aspergillus nidulans* с использованием кодонов, оптимизированных для кукурузы, и в которой удалены последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1.

SEQ ID NO:48 представляет собой белок дельта-9-десатуразы жирных кислот *Aspergillus nidulans*.

35 SEQ ID NO:49 представляет собой нативную последовательность ДНК, кодирующую белок SAP1 *Xerophyta viscosa*.

SEQ ID NO:50 представляет собой последовательность белка SAP1 *Xerophyta viscosa*.

40 SEQ ID NO:51 представляет собой синтетическую последовательность ДНК, кодирующую белок SAP1 *Xerophyta viscosa* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1 и таблицы 2.

SEQ ID NO:52 представляет собой последовательность белка SAP1 *Xerophyta viscosa*.

45 SEQ ID NO:53 представляет собой синтетическую последовательность ДНК по изобретению, кодирующую белок SAP1 *Xerophyta viscosa* с использованием кодонов, оптимизированных для кукурузы, и в которой удалены последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1

SEQ ID NO:54 представляет собой последовательность белка SAP1 *Xerophyta viscosa*.

SEQ ID NO:55 представляет собой нативную последовательность ДНК, кодирующую

белок GFP1 *Aequorea victoria*.

SEQ ID NO:56 представляет собой последовательность белка GFP1 *Aequorea victoria*.

SEQ ID NO:57 представляет собой синтетическую последовательность ДНК, кодирующую белок GFP1 *Aequorea victoria* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1 и таблицы 2.

SEQ ID NO:58 представляет собой последовательность белка GFP1 *Aequorea victoria*.

SEQ ID NO:59 представляет собой синтетическую последовательность ДНК по изобретению, кодирующую белок GFP1 *Aequorea victoria* с использованием кодонов, оптимизированных для кукурузы, и в которой удалены последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1.

SEQ ID NO:60 представляет собой последовательность белка GFP1 *Aequorea victoria*.

SEQ ID NO:61 представляет собой нативную последовательность ДНК, кодирующую белок дельта-9-десатуразы жирных кислот *Leptosphaeria nodorum*.

SEQ ID NO:62 представляет собой последовательность белка дельта-9-десатуразы жирных кислот *Leptosphaeria nodorum*.

SEQ ID NO:63 представляет собой синтетическую последовательность ДНК, кодирующую белок дельта-9-десатуразы жирных кислот *Leptosphaeria nodorum* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1 и таблицы 2.

SEQ ID NO:64 представляет собой последовательность белка дельта-9-десатуразы жирных кислот *Leptosphaeria nodorum*.

SEQ ID NO:65 представляет собой синтетическую последовательность ДНК по изобретению, кодирующую белок дельта-9-десатуразы жирных кислот *Leptosphaeria nodorum* с использованием кодонов, оптимизированных для кукурузы, и в которой удалены последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1

SEQ ID NO:66 представляет собой последовательность белка дельта-9-десатуразы жирных кислот *Leptosphaeria nodorum*.

SEQ ID NO:67 представляет собой нативную последовательность ДНК, кодирующую белок PER1 *Xerophyta viscosa*.

SEQ ID NO:68 представляет собой последовательность белка PER1 *Xerophyta viscosa*.

SEQ ID NO:69 представляет собой синтетическую последовательность ДНК, кодирующую белок PER1 *Xerophyta viscosa* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1 и таблицы 2.

SEQ ID NO:70 представляет собой последовательность белка PER1 *Xerophyta viscosa*.

SEQ ID NO:71 представляет собой синтетическую последовательность ДНК по изобретению, кодирующую белок PER1 *Xerophyta viscosa* с использованием кодонов, оптимизированных для кукурузы, и в которой удалены последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1.

SEQ ID NO:72 представляет собой последовательность белка PER1 *Xerophyta viscosa*.

ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

Изобретение предоставляет синтетические последовательности нуклеиновых кислот, кодирующие представляющие интерес белки. Синтетические кодирующие последовательности, в частности, приспособлены для применения в экспрессии

представляющих интерес белков в трансгенных растениях.

Представляющий интерес белок может быть любым белком или полипептидом, который встречается в природе, или любым встречающимся в природе вариантом, включая процессированные формы таких белков, но не ограничиваясь ими.

5 Представляющий интерес белок также может являться белком, образованным путем объединения частей или фрагментов более чем одного встречающегося в природе белка, например, путем смешивания и сопоставления функциональных белковых доменов.

Предпочтительная группа представляющих интерес белков представляет собой группу, в которой получаемый в результате фенотип является агрономическим
10 признаком или белком-репортером, применяемым для создания агрономических признаков. Это включает в себя резистентность к насекомым, устойчивость к гербицидам, устойчивость к недостатку воды и/или тепловому стрессу и изменению масляного профиля, но не ограничивается перечисленным.

Более предпочтительная группа представляющих интерес белков представляет собой
15 группу, в которой получаемый в результате фенотип является агрономическим признаком. Другой предпочтительной группой является группа, в которой получаемый в результате фенотип обеспечивает устойчивость к гербицидам. Другой предпочтительной группой является группа, в которой получаемый в результате фенотип обеспечивает устойчивость к стрессу. Другой предпочтительной группой является
20 группа, в которой получаемый в результате фенотип обеспечивает модифицированный масляный профиль для более здоровой пищи. Более предпочтительной группой является группа, в которой представляющим интерес белком является белок Cгу, который обеспечивает резистентность к насекомым.

Нативные последовательности ДНК/последовательности ДНК дикого типа,
25 кодирующие представляющий интерес белок, должны быть идентифицированы и проанализированы для определения присутствия последовательностей сигналов полиаденилирования, приведенных в таблицах 1 и 2 и/или 3. В соответствии с изобретением, для кодирующих последовательностей, предназначенных для применения в кукурузе, количество последовательностей сигналов полиаденилирования,
30 приведенных в таблице 2, снижается по сравнению с количеством, присутствующим в нативной последовательности. Для кодирующих последовательностей, предназначенных для применения в сое, снижается количество последовательностей сигналов полиаденилирования, приведенных в таблице 3. Очень важно удалить последовательности сигналов полиаденилирования, приведенные в таблицах 2 и 3, в
35 частности, когда они появляются во вложенной мультимерной форме.

В дополнение к удалению последовательностей сигналов полиаденилирования, приведенных в таблицах 2 и 3, может быть необходимым удаление появлений последовательности Шоу-Камен, АТТГА.

В дополнение к удалению последовательностей сигналов полиаденилирования и
40 последовательностей Шоу-Камен, является предпочтительным построение синтетических кодирующих последовательностей ДНК, в которых кодоны используются приблизительно с той же частотой, с которой они используются, в среднем, во встречающихся в природе генах в растениях тех видов, в которых будет применяться синтетическая последовательность ДНК. В таблице 4 приведены подходящие процентные
45 значения для использования кодонов в синтетических генах, предназначенных для применения в различных сельскохозяйственных растениях, а также для применения, в общем случае, в двудольных растениях, или, в общем случае, в растениях.

Целевые ремасштабированные составы кодонов синтетических генов растений							
Аминокислота	Кодон	% в кукурузе	% в сое	Аминокислота	Кодон	% в кукурузе	% в сое
5 ALA (A)	GCA	18,0	33,1	LEU (L)	CTA	0	0
	GCC	34,0	24,5		CTC	29,9	22,4
	GCG	24,0	0		CTG	33,3	16,3
	GCT	24,0	42,3		CTT	19,5	31,5
10 ARG (R)	AGA	18,8	36,0	LYS (K)	TTA	0	0
	AGG	32,5	32,2		TTG	17,2	29,9
	CGA	0	0	AAA	22,0	42,5	
	CGC	30,0	15	AAG	78,0	57,5	
	CGG	18,8	0	MET (M)	ATG	100	100
15 ASN (N)	CGT	0	16,9	PHE (F)	TTC	71,0	49,2
	AAC	68,0	50,0		TTT	29,0	50,8
ASP (D)	AAT	32,0	50,0	PRO (P)	CCA	26,0	39,8
	GAC	63,0	38,1		CCC	24,0	20,9
15 CYS (C)	GAT	37,0	61,9		CCG	28,0	0,0
	TGC	68,0	50,0		CCT	22,0	39,3
END	TGT	32,0	50,0	SER (S)	AGC	25,3	16,0
	TAA	0	0		AGT	0,0	18,2
	TAG	0	0		TCA	17,6	21,9
20 GLN (Q)	TGA	100	100		TCC	25,3	18,0
	CAA	38,0	55,5		TCG	15,4	0
GLU (E)	CAG	62,0	44,5		TCT	16,5	25,8
	GAA	29,0	50,5	THR (T)	ACA	21,0	32,4
GAG	71,0	49,5	ACC		37,0	30,2	
25 GLY (G)	GGA	19,0	31,9		ACG	22,0	0,0
	GGC	42,0	19,3		ACT	20,0	37,4
	GGG	19,0	18,4	TRP (W)	TGG	100	100
HIS (H)	GGT	20,0	30,4	TYR (Y)	TAC	73,0	48,2
	CAC	62,0	44,8		TAT	27,0	51,8
30 ILE (I)	CAT	38,0	55,2	VAL (V)	GTA	0	11,5
	ATA	14,0	23,4		GTC	34,8	17,8
	ATC	58,0	29,9		GTG	42,4	32,0
	ATT	28,0	46,7		GTT	22,8	38,7

30 ТРАНСГЕННЫЕ РАСТЕНИЯ

Предпочтительным вариантом осуществления настоящего изобретения является трансформация растений генами, кодирующими токсины против насекомых. Трансформированные растения, которые экспрессируют гены токсинов против насекомых, являются устойчивыми к нападению целевого насекомого-вредителя благодаря наличию контролируемых объемов соответствующего инсектицидного белка или его вариантов в клетках трансформированного растения. В результате включения генетического материала, который кодирует инсектицидные свойства инсектицидных токсинов *B.t.*, в геном растения, съеденного конкретным насекомым-вредителем, взрослая особь или личинка погибает после употребления кормового растения. Было трансформировано множество представителей односемядольных и двусемядольных. Трансгенные агрономические сельскохозяйственные культуры, а также фрукты и овощи, представляют коммерческий интерес. Такие сельскохозяйственные культуры включают кукурузу, рис, сою, канолу, подсолнечник, люцерну, сорго, пшеницу, хлопок, арахис, томат, картофель и т.п., но не ограничиваются перечисленным. Существует несколько методик внесения чужеродного генетического материала в клетки растений и для получения растений, которые стабильно поддерживают и экспрессируют внесенный ген. Такие методики включают запуск генетического материала, нанесенного на микрочастицы, непосредственно в клетки (патент США No. 4945050 и патент США No.

5141131). Растения могут трансформироваться с применением агробактериальной технологии, см. патент США No. 5177010, Европейский патент No. EP 131624B1, Европейский патент No. EP 159418B1, Европейский патент No. EP176112B1, патент США No. 5149645, EP120516B1, патент США No. 5464763, патент США No. 4693976, Европейский патент No. EP 116718B1, Европейский патент No. EP290799B1, Европейский патент No. EP320500B1, Европейский патент No. EP604662B1, патент США No. 7060876, патент США No. 6037526, патент США No. 6376234, Европейский патент No. EP292435B1, патент США No. 5231019, патент США No. 5463174, патент США No. 4762785, патент США No. 5608142 и патент США No. 5159135. Другой технологией трансформации является технология WHISKERS™, см. патент США No. 5302523 и патент США No. 5464765. Для трансформации растений также была применена технология электропорации, см. WO1987006614, патент США No. 5472869, патент США No. 5384253, WO199209696, патент США No. 6074877, WO1993021335 и патент США No. 5679558. В дополнение к множеству технологий для трансформирования растений, также может изменяться тип ткани, который контактирует с чужеродными генами. Такая ткань могла бы включать зародышевую ткань, ткань каллуса типа I и типа II, гипокотиль, меристему и т.п., но не ограничивается перечисленным. Практически все растительные ткани могут быть трансформированы во время дедифференциации с применением соответствующих методик, известных специалисту в данной области техники.

Известные методики вставки ДНК в растения включают трансформацию с помощью Т-ДНК, доставляемой *Agrobacterium tumefaciens* или *Agrobacterium rhizogenes* в качестве посредника трансформации. Применение Т-ДНК-содержащих векторов для трансформации растительных клеток тщательно исследовалось и в достаточном объеме описано в европейском патенте No. EP120516B1; Lee and Gelvin (2008) *Plant Physiol.* 146: 325-332; Fraley et al. (1986) *Crit. Rev. Plant Sci.* 4: 1-46; и An et al. (1985) *EMBO J.* 4:277-284; и является общепризнанным в данной области. Кроме того, может быть применено слияние растительных протопластов с липосомами, содержащими ДНК, которая должна быть доставлена, прямое инъектирование ДНК, биолистическая трансформация (бомбардировка микрочастицами) или электропорация, а также другие возможные способы.

После того, как вставленная ДНК была интегрирована в растительный геном, она остается сравнительно стабильной в последующих поколениях. Вектор, используемый для трансформации растительной клетки, обычно содержит ген селективируемого маркера, кодирующий белок, который придает трансформированным растительным клеткам устойчивость к гербициду или антибиотику, таким как, в том числе, биалафос, канамицин, G418, блеомицин, или гигромицин. Отдельно используемый ген селективируемого маркера должен позволять выбирать трансформированные клетки, при этом рост клеток, которые не содержат вставленную ДНК, подавляется посредством селективного компонента.

В предпочтительном варианте осуществления настоящего изобретения растения трансформируются генами, в которых использование кодонов в кодирующем белок участке было оптимизировано для растений. См., например, патент США No. 5380831. Также, предпочтительно, могут быть использованы растения, кодирующие усеченный токсин, например, функциональный белковый домен. Усеченный токсин обычно кодирует от около 55% до около 80% нативного полноразмерного токсина. Способы создания синтетических генов *in vitro* для применения в растениях известны в технике (Stewart 2007, *Frontiers in Drug Design and Discovery* 1:297-341).

Независимо от методики трансформации, ген предпочтительно встраивается в вектор

передачи генов, приспособленный для экспрессии представляющего интерес белка путем включения в вектор растительного промотора. В дополнение к растительным промоторам, промоторы из ряда источников могут быть эффективно использованы в растительных клетках для экспрессии чужеродных генов. Например, могут быть

5 использованы промоторы из бактериального источника, такие как промотор октопин-синтазы, промотор нопапин-синтазы, промотор маннопин-синтазы; промоторы из вирусного источника, такие как промоторы 35S и 19S вируса мозаики цветной капусты (CaMV) и т.п. Полученные из растений промоторы включают, но не ограничиваются

10 перечисленным, рибулозу-1, малую субъединицу (ssu) 6-бисфосфат (RUBP) карбоксилазы, промотор бета-конглицинина, промотор фазеолина, промотор ADH (алкогольдегидрогеназы), промоторы белков теплового шока, промотор ADF (фактор деполимеризации актина) и тканеспецифичные промоторы. Промоторы также могут

15 содержать определенные элементы последовательностей энхансеров, которые могут повышать эффективность транскрипции. Типичные энхансеры включают ADH1-интрон 1 и ADH1-интрон 6, но не ограничиваются указанным. Также могут быть использованы конститутивные промоторы. Конститутивные промоторы управляют непрерывной экспрессией генов практически во всех типах клеток практически в любые моменты

20 времени (например, актин, убиквитин, CaMV 35S). Тканеспецифичные промоторы отвечают за экспрессию генов в конкретных типах клеток или ткани, таких как листья или семена (например, зеин, олеозин, напин, ACP (ацилпереносящий белок)), и такие промоторы также могут использоваться. Также могут быть использованы промоторы, которые являются активными во время определенной стадии развития растений, а также которые являются активными в конкретных тканях и органах растений. Примеры

25 таких промоторов включают, но не ограничиваются перечисленным, специфичные для корней, специфичные для пыльцы, специфичные для зародышей, специфичные для кукурузных рылец, специфичные для хлопкового волокна, специфичные для эндосперма семени, специфичные для флоэмы и т.п.

В определенных обстоятельствах может потребоваться использование индуцируемого промотора. Индуцируемый промотор отвечает за экспрессию генов в ответ на

30 конкретный сигнал, такой как: физическое стимулирование (например, гены теплового шока); свет (например, RUBP-карбоксилаза); гормон (например, глюкокортикоид); антибиотик (например, тетрациклин); метаболиты; и стресс (например, засуху). Также могут быть использованы другие необходимые элементы транскрипции и трансляции, которые функционируют в растениях, такие как 5'-нетранслируемые лидерные

35 последовательности, последовательности терминации транскрипции РНК и сигнальные последовательности добавления полиаденилирования. Множество специфичных для растений векторов переноса генов известно в технике.

Трансгенные сельскохозяйственные растения, содержащие свойства устойчивости к насекомым (IR), преобладают в растениях кукурузы и хлопка в Северной Америке,

40 и использование данных свойств расширяется по всему миру. Коммерческие трансгенные сельскохозяйственные растения, объединяющие свойства IR и устойчивости к гербицидам (HT), были разработаны множеством компаний, производящих семена. Это включает комбинации свойств IR, придаваемых инсектицидными белками B.t., и свойства HT, такие как устойчивость к ингибиторам ацетолактатсинтазы (ALS), таким как

45 сульфонилмочевины, имидазолиноны, триазолопиримидин, сульфонилиды и т.п., ингибиторы глутаминсинтетазы (GS), такие как биалафос, глюфосинат и т.п., ингибиторы 4-гидроксифенилпируват диоксигеназы (HPPD), такие как мезотрион, изоксафлютол и т.п., ингибиторы 5-энолпирувилшикимат-3-фосфат синтазы (EPSPS), такие как глифосат

и т.п., и ингибиторы ацетил-кофермент А карбоксилазы (ACCase), такие как галоксифоп, квизалофоп, диклофоп и т.п. Известны другие примеры, в которых трансгенно полученные белки обеспечивали устойчивость растений к химическим классам гербицидов, таким как феноксикислотные гербициды и гербициды, основанные на пиридилоксиацетатах ауксина (см. WO2007053482), или феноксикислотные гербициды и арилоксифеноксипропионатные гербициды (см. заявку на патент США No. 20090093366). Способность решения проблем с множеством вредителей через свойства IR является ценной концепцией в производстве товарных продуктов, и удобство данной концепции товара повышается, если свойства борьбы с насекомыми и средства борьбы с сорняками объединяются в одном и том же растении. Кроме того, повышенная ценность может быть получена через объединение в одном растении свойств IR, придаваемых инсектицидным белком B.t., таким образом, как в настоящем изобретении, с одним или более дополнительными свойствами НТ, такими как были указаны выше, плюс одно или более дополнительных входных свойств (например, устойчивость к другому насекомому, придаваемая полученными из B.t. или другими инсектицидными белками, устойчивость к насекомому, придаваемая такими механизмами, как РНКi и т.п., устойчивость к нематодам, устойчивость к заболеваниям, устойчивость к стрессам, улучшенная утилизация азота и т.п.) или выходных свойств (например, высокое содержание масел, здоровый состав масел, улучшение питательных свойств и т.п.). Такие комбинации могут быть получены через обычное скрещивание (селекционные гибриды) или совместно в форме нового факта трансформации, включающего одновременное введение множества генов (молекулярные гибриды или совместная трансформация). Преимущества включают возможность борьбы с насекомыми-вредителями и улучшения борьбы с сорняками в сельскохозяйственном растении, которые обеспечивают вторичные преимущества для производителя и/или потребителя. Таким образом, настоящее изобретение может быть применено в связи с множеством свойств для предоставления полного агрономического комплекса улучшенного качества сельскохозяйственного растения с возможностью гибко и с оптимальными затратами бороться с любым количеством агрономических проблем.

Все патенты, заявки на патенты, предварительные заявки и публикации, цитируемые в настоящей заявке или на которые имеются ссылки в настоящей заявке, включены посредством ссылки во всей их полноте до той степени, в которой они не противоречат явно изложенному в настоящей спецификации. Если это конкретно не указано или не подразумевается, то термин "один" означает "по меньшей мере один" при использовании в настоящем описании. Под "изолированными" заявители понимают, что нуклеотидные или полипептидные молекулы были удалены из своего нативного окружения и были помещены в другое окружение с помощью человека.

Варианты осуществления настоящего изобретения подробнее определены в приведенных ниже примерах. Следует понимать, что данные примеры приведены только в качестве иллюстрации. На основании приведенного выше обсуждения и данных примеров специалист в данной области техники может установить важные характеристики настоящего изобретения и, без отклонения от его формы и объема, может вносить различные изменения и модификации в варианты осуществления изобретения с целью их настройки для различных применений и условий. Таким образом, различные модификации вариантов осуществления изобретения, в дополнение к показанным и описанным в настоящей заявке, будут очевидны специалистам в данной области техники на основании приведенного выше описания. Предполагается, что такие модификации также попадают в пределы объема прилагаемой формулы

изобретения.

Все проценты приведены по весу, и все пропорции смешанного растворителя приведены по объему, если не отмечено обратное. Все температуры приведены в градусах Цельсия.

5 ПРИМЕР 1

Синтетическая кодирующая область, кодирующая ядерный токсин Cry1Fa *Bacillus thuringiensis*

Последовательность для сравнения. Нативная последовательность ДНК, кодирующая ядерный токсин Cry1Fa, приведена под SEQ ID NO:1. Данная последовательность
10 анализировалась с целью определения того, какие из последовательностей, идентифицированных в таблице 1, присутствуют в SEQ ID NO:1, и их положений. Аминокислотная последовательность, кодируемая SEQ ID NO:1, затем подвергалась обратной трансляции с использованием целевых частот кодонов, приведенных в столбце таблицы 4 для синтетических генов, предназначенных для использования в кукурузе.
15 Полученная в результате последовательность ДНК анализировалась, и кодоны заменялись при необходимости для удаления нежелательных открытых рамок считывания и удаления нежелательных сайтов рестрикции, и восстановления последовательностей, идентифицированных в таблице 1. Аминокислотная последовательность, кодируемая SEQ ID NO:1, сохранялась. Полученная в результате
20 последовательность ДНК дана под SEQ ID NO:3.

SEQ ID NO:3 анализировалась, и кодоны были изменены для удаления потенциальных последовательностей сигналов полиаденилирования, идентифицированных в таблице 2, при этом поддерживалось то же самое количество последовательностей, идентифицированных в таблице 1. Полученная в результате последовательность, которая
25 является вариантом осуществления настоящего изобретения, дана под SEQ ID NO:5. В таблице 5 показано, что количество и положения последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, поддерживается в SEQ ID NO:5, за исключением того, что два вхождения ААТААА, одно в 426 н.п. и одно в 582 н.п., в SEQ ID NO:1 были заменены на ААТСАА, в результате чего поддерживается количество
30 и положение последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, но производится замена каждой из двух последовательностей ААТААА на менее проблемные последовательности. В таблице 6 показано, что количество последовательностей сигналов полиаденилирования, идентифицированных в таблице 2, снижается в SEQ ID NO:5. Поскольку имеется перекрытие последовательностей,
35 идентифицированных в таблицах 2 и 3 (последовательности 1, 2, 6, 7, 8, 9, 10, 14, 13 и 20 в таблице 2 соответствуют последовательностям 16, 15, 2, 5, 1, 3, 4, 6, 13 и 12, соответственно, в таблице 3), также получается, что количество последовательностей сигналов полиаденилирования, идентифицированных в таблице 3, снижается в SEQ ID NO:5.

40 Синтетическая кодирующая область SEQ ID NO:5 была оптимизирована для экспрессии в кукурузе.

Конструкция для применения в экспрессии синтетической кодирующей области SEQ ID NO:5 сделана путем объединения синтетической кодирующей области SEQ ID NO:5 с 5'-нетранслируемой областью, содержащей промотор, функциональный в растительных
45 клетках, и 3'-нетранслируемой областью, содержащей последовательность терминации транскрипции и полиаденилирования.

В одном из вариантов осуществления такой конструкции продуцирование первичного транскрипта мРНК, содержащего SEQ ID NO:5, управлялось копией промотора

убиквитина 1 кукурузы со своим нативным интроном 1 (патент США No. 5510474). Фрагмент, содержащий 3'-нетранслируемую область из гена пероксидазы 5 кукурузы (ZmPer5 3'UTR v2; патент США No. 6699984), был использован для терминации транскрипции. Бинарная плазмида для трансформации растений, pDAB111440, содержащая указанную выше кассету экспрессии генов, была сконструирована и применена в продуцировании трансгенных растений кукурузы. Плазмида pDAB111440 также содержит ген устойчивости к гербицидам, содержащий кодирующую область для арилоксиалконат диоксигеназы (AAD-1 v3; патент США No. 7838733(B2), и Wright et al. (2010) Proc. Natl. Acad. Sci. U.S.A. 107:20240-5) под транскрипционным контролем промотора палочковидного баднавируса сахарного тростника (ScBV) (Schenk et al. (1999) Plant Molec. Biol. 39: 1221-30). Фрагмент, содержащий 3'-нетранслируемую область из гена липазы кукурузы (ZmLip 3'UTR; патент США No. 7179902), был использован для терминации транскрипции.

15

Таблица 5

Последовательности таблицы 1, обнаруженные в нативной кодирующей области ядерного токсина Cry1Fa (SEQ ID NO:1) и в ее перестроенной версии (SEQ ID NO:5)

	Последовательность из таблицы 1	Число участков в нативной последовательности ядерного токсина Cry1Fa (SEQ ID NO:1)	Положение в нативной последовательности ядерного токсина Cry1Fa, н.п. (SEQ ID NO:1)	Число участков в перестроенной последовательности ядерного токсина Cry1Fa (SEQ ID NO:5)	Положение в перестроенной последовательности ядерного токсина Cry1Fa, н.п. (SEQ ID NO:5)	
20	1	ААТААА	2	426; 582	0	NA*
	2	ААТААГ	5	7; 46; 358; 430; 562	5	7; 46; 358; 430; 562
	3	ААССАА	0	NA	0	NA
	4	АТАТАА	1	1520	1	1520
	5	ААТСАА	2	19; 628	4	19; 426; 582; 628
25	6	АТАСТА	1	1508	1	1508
	7	АТАААА	0	NA	0	NA
	8	АТГААА	2	314; 1211	2	314; 1211
	9	ААГСАГ	0	NA	0	NA
	10	АТТААГ	2	579; 1690	2	579; 1690
	11	АТАСАГ	0	NA	0	NA
30	12	ААААТА	0	NA	0	NA
	13	АТТААА	2	66; 1266	2	66; 1266
	14	ААТТАА	2	368; 779	2	368; 779
	15	ААТАСА	3	400; 1369; 1693	3	400; 1369; 1693
	16	САТААА	0	NA	0	NA
		итого	22		22	

35 *NA = Не применимо

40

Таблица 6

Последовательности таблицы 2, обнаруженные в нативной кодирующей области ядерного токсина Cry1Fa (SEQ ID NO:1) и в ее перестроенной версии (SEQ ID NO:5)

	Последовательность из таблицы 2	Число участков в нативной последовательности ядерного токсина Cry1Fa (SEQ ID NO:1)	Положение в нативной последовательности ядерного токсина Cry1Fa, н.п. (SEQ ID NO:1)	Число участков в перестроенной последовательности ядерного токсина Cry1Fa (SEQ ID NO:5)	Положение в перестроенной последовательности ядерного токсина Cry1Fa, н.п. (SEQ ID NO:5)	
45	1	АТАТАГ	1	104	0	NA*
	2	ТТГТТТ	3	39; 612; 907	0	NA
	3	ТТТТГТ	1	1089	0	NA
	4	ТГТТТТ	2	1086; 1334	0	NA
	5	ТАТАТА	1	1771	0	NA
	6	ТАТТТТ	0	NA	0	NA
	7	ТТТТТТ	0	NA	0	NA
	8	АТТТТТ	1	1615	0	NA
	9	ТТАТТТ	2	172; 217	0	NA

10	ТТТАТТ	0	NA	0	NA
11	ТААТАА	4	357; 416; 561; 581	0	NA
12	АТТТАГ	3	319; 497; 793	0	NA
13	ТАТАТТ	1	322	0	NA
14	ТТТТАГ	3	192; 464; 1063	0	NA
15	АТАТТТ	0	NA	0	NA
16	ТАТТАГ	0	NA	0	NA
17	ТГТТТГ	2	613; 908	0	NA
18	ТТАТАГ	2	321; 1770	0	NA
19	ТГТААГ	0	NA	0	NA
20	АААТАА	2	45; 429	0	NA
	итого	28		0	NA

*NA = Не применимо

ПРИМЕР 2

Синтетическая кодирующая область, кодирующая токсин Cry34A *Bacillus thuringiensis*

Последовательности для сравнения. Нативная последовательность ДНК, кодирующая токсин Cry34A, дана под SEQ ID NO:7. Данная последовательность была проанализирована с целью определения того, какие последовательности из идентифицированных в таблице 1 присутствуют в SEQ ID NO:7, и их положений.

Нативная последовательность ДНК была протранслирована в соответствующую аминокислотную последовательность с использованием стандартного генетического кода. Аминокислотная последовательность, кодируемая SEQ ID NO:7, подвергалась обратной трансляции с использованием целевых частот кодонов, приведенных в столбце таблицы 7 для синтетических генов, предназначенных для использования в кукурузе.

Полученная в результате последовательность ДНК анализировалась, и кодоны заменялись при необходимости для удаления нежелательных открытых рамок считывания и удаления нежелательных сайтов рестрикции, и восстановления всех последовательностей, идентифицированных в таблице 1. Аминокислотная последовательность, кодируемая SEQ ID NO:7, сохранялась. Полученная в результате последовательность ДНК дана под SEQ ID NO:9. Была синтезирована ДНК, имеющая последовательность SEQ ID NO:9.

SEQ ID NO:9 была проанализирована, и кодоны были изменены с целью удаления потенциальных последовательностей сигналов полиаденилирования, идентифицированных в таблице 2, при этом поддерживалось то же самое количество последовательностей, идентифицированных в таблице 1. Полученная в результате последовательность, которая является вариантом осуществления настоящего изобретения, дана под SEQ ID NO:11. В таблице 7 показано, что количество и положения последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, поддерживается в SEQ ID NO:11. В таблице 8 показано, что количество последовательностей сигналов полиаденилирования, идентифицированных в таблицах 2 и 3, снижается в SEQ ID NO:5.

ДНК с SEQ ID NO:5 синтезируется, и уровни экспрессии, наблюдаемые в растительных клетках, трансформированных для экспрессии данной последовательности, сравниваются с уровнями экспрессии, наблюдаемыми в растительных клетках, трансформированных для экспрессии SEQ ID NO:1 и SEQ ID NO:3.

Синтетическая кодирующая область SEQ ID NO:5 была оптимизирована для экспрессии в кукурузе.

Конструкция для применения в экспрессии синтетической кодирующей области SEQ ID NO:5 сделана путем объединения синтетической кодирующей области SEQ ID NO:5 с 5'-нетранслируемой областью, содержащей промотор, функциональный в растительных

клетках, и 3'-нетранслируемой области, содержащей последовательность терминации транскрипции и полиаденилирования.

5

10

15

20

Таблица 7

Последовательности таблицы 1, обнаруженные в нативной кодирующей области Cry34Ab1 (SEQ ID NO:7), и ее перестроенной версии (SEQ ID NO:11)

	Последовательность из таблицы 1	Число участков в нативной последовательности Cry34Ab1 (SEQ ID NO:7)	Положение в нативной последовательности Cry34Ab1, н.п. (SEQ ID NO:7))	Число участков в перестроенной последовательности Cry34Ab1 (SEQ ID NO:11)	Положение в перестроенной последовательности Cry34Ab1, н.п. (SEQ ID NO:11)	
	1	ААТААА	2	247; 268	2	247; 268
	2	ААТААГ	1	31	1	31
	3	ААССАА	0	NA*	0	NA
	4	АТАГАА	0	NA	0	NA
	5	ААТСАА	2	146; 310	2	146; 310
	6	АТАСТА	1	329	1	329
	7	АТАААА	1	65	1	65
	8	АТГААА	1	281	1	281
	9	ААГСАТ	0	NA	0	NA
	10	АТТААГ	0	NA	0	NA
	11	АГАСАТ	1	47	1	47
	12	ААААТА	0	NA	0	NA
	13	АТТААА	1	127	1	127
	14	ААТГАА	1	126	1	126
	15	ААТАСА	0	NA	0	NA
	16	САТААА	1	361	1	361
		итого	12		12	

*NA = Не применимо

25

30

35

40

45

Таблица 8

Последовательности таблицы 2, обнаруженные в нативной кодирующей области Cry34Ab1 (SEQ ID NO:7) и ее перестроенной версии (SEQ ID NO:11)

	Последовательность из таблицы 2	Число участков в нативной последовательности Cry34Ab1 (SEQ ID NO:7)	Положение в нативной последовательности Cry34Ab1, н.п. (SEQ ID NO:7))	Число участков в перестроенной последовательности Cry34Ab1 (SEQ ID NO:11)	Положение в перестроенной последовательности Cry34Ab1, н.п. (SEQ ID NO:11)	
	1	АТАТАГ	1	181	0	NA*
	2	ТТГТТТ	0	NA	0	NA
	3	ТТТТГТ	0	NA	0	NA
	4	ТГТТТТ	0	NA	0	NA
	5	ТАТАГА	1	180	0	NA
	6	ТАТТТТ	1	220	0	NA
	7	ТТТТТТ	0	NA	0	NA
	8	АГТТТТ	0	NA	0	NA
	9	ТТАТТТ	0	NA	0	NA
	10	ТТТАТТ	0	NA	0	NA
	11	ТААТАА	2	33; 246	2	33; 246
	12	АГТТАГ	0	NA	0	NA
	13	ТАТААТ	2	182; 218	0	NA
	14	ТТТТАГ	1	156	0	NA
	15	АТАТТТ	1	219	0	NA
	16	ТАТТАГ	1	184	0	NA
	17	ТГТТТГ	0	NA	0	NA
	18	ТТАТАГ	1	217	0	NA
	19	ТГТААГ	0	NA	0	NA
	20	АААТАА	1	30	1	30
		итого	12		3	

*NA = Не применимо

ПРИМЕР 3

Синтетическая кодирующая область, кодирующая токсин Cry35Ab1 *Bacillus thuringiensis*

Последовательности для сравнения. Нативная последовательность ДНК, кодирующая токсин Cry35Ab1, дана под SEQ ID NO:13. Данная последовательность была проанализирована с целью определения того, какие последовательности из идентифицированных в таблице 1 присутствуют в SEQ ID NO:13, и их положений. Аминокислотная последовательность, кодируемая SEQ ID NO:13, подвергалась обратной трансляции с использованием целевых частот кодонов, приведенных в столбце таблицы 4 для синтетических генов, предназначенных для использования в кукурузе. Полученная в результате последовательность ДНК анализировалась, и кодоны заменялись при необходимости для удаления нежелательных открытых рамок считывания и удаления нежелательных сайтов рестрикции, при этом поддерживались все последовательности, идентифицированные в таблице 1. Аминокислотная последовательность, кодируемая SEQ ID NO:13, сохранялась. Полученная в результате последовательность ДНК дана под SEQ ID NO:15. Данная последовательность будет синтезирована и использована для сравнения с синтетической кодирующей областью, спроектированной по изобретению.

SEQ ID NO:15 была проанализирована, и кодоны были изменены с целью удаления потенциальных последовательностей сигналов полиаденилирования, идентифицированных в таблице 2, при этом поддерживалось то же самое количество последовательностей, идентифицированных в таблице 1, за исключением двух вхождений ААТААА, одного в 228 н.п. и одного в 276 н.п. в SEQ ID NO:8, которые были заменены на ААТСАА. Полученная в результате последовательность, которая является вариантом осуществления настоящего изобретения, дана под SEQ ID NO:17. В таблице 9 показано, что количество и положение последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, поддерживается в SEQ ID NO:17. В таблице 10 показано, что количество последовательностей сигналов полиаденилирования, идентифицированных в таблицах 2 и 3, снижается в SEQ ID NO:17 по сравнению с SEQ ID NO:13.

ДНК из SEQ ID NO:17 синтезируется, и уровни экспрессии, наблюдаемые в растительных клетках, трансформированных для экспрессии данной последовательности, сравниваются с уровнями экспрессии, наблюдаемыми в растительных клетках, трансформированных для экспрессии SEQ ID NO:13 и SEQ ID NO:15.

Синтетическая кодирующая область SEQ ID NO:17 была оптимизирована для экспрессии в кукурузе.

Конструкция для применения в экспрессии синтетической кодирующей области SEQ ID NO:17 сделана путем объединения синтетической кодирующей области SEQ ID NO:17 с 5'-нетранслируемой областью, содержащей промотор, функциональный в растительных клетках, и 3'-нетранслируемой области, содержащей последовательность терминации транскрипции и полиаденилирования.

Таблица 9

Последовательности таблицы 1, обнаруженные в нативной кодирующей области Cry35Ab1 (SEQ ID NO:13) и ее перестроенной версии (SEQ ID NO:17)

	Последовательность из таблицы 1	Число участков в нативной последовательности Cry35Ab1 (SEQ ID NO:13)	Положение в нативной последовательности Cry35Ab1, н.п. (SEQ ID NO:13)	Число участков в перестроенной последовательности Cry35Ab1 (SEQ ID NO:17)	Положение в перестроенной последовательности Cry35Ab1, н.п. (SEQ ID NO:17)
1	ААТААА	5	13; 100; 228; 276; 810	3	13; 100; 810

2	AATAAT	4	193; 217; 385; 864	4	193; 217; 385; 864
3	AACCAA	0	NA*	0	NA
4	ATATAA	1	966	1	966
5	AATCAA	3	394; 750; 914	5	228; 276; 394; 750; 914
6	ATACTA	1	8	1	8
7	ATAAAA	5	101; 224; 277; 575; 811	5	101; 224; 277; 575; 811
8	ATGAAA	5	23; 671; 769; 806; 854	5	23; 671; 769; 806; 854
9	AAGCAT	0	NA	0	NA
10	ATTAAT	1	522	1	522
11	ATACAT	1	734	1	734
12	AAAATA	7	226; 578; 618; 838; 862; 873; 1137	7	226; 578; 618; 838; 862; 873; 1137
13	ATGAAA	4	462; 589; 834; 1131	4	462; 589; 834; 1131
14	AATGAA	5	461; 521; 588; 833; 1130	5	461; 521; 588; 833; 1130
15	AATACA	3	261; 303; 733	3	261; 303; 733
16	CATAAA	0	NA	0	NA
	итого	45		45	
*NA = Не применимо					

Таблица 10					
Последовательности таблицы 2, обнаруженные в нативной кодирующей области Cry35Ab1 (SEQ ID NO:13) и ее перестроенной версии (SEQ ID NO:17)					
	Последовательность из таблицы 2	Число участков в нативной последовательности Cry35Ab1 (SEQ ID NO:13)	Положение в нативной последовательности Cry35Ab1, н.п. (SEQ ID NO:13)	Число участков в перестроенной последовательности Cry35Ab1 (SEQ ID NO:17)	Положение в перестроенной последовательности Cry35Ab1, н.п. (SEQ ID NO:17)
1	ATATAT	1	168	0	NA*
2	TTGTTT	0	NA	0	NA
3	TTTTGT	0	NA	0	NA
4	TGTTTT	0	NA	0	NA
5	TATATA	1	959	0	NA
6	TATTTT	2	609; 1144	0	NA
7	TTTTTT	0	NA	0	NA
8	ATTTTT	1	1145	0	NA
9	TTATTT	3	63; 145; 1143	1	1143
10	TTTATT	2	144; 1056	0	NA
11	TAATAA	2	12; 216	1	12
12	ATTTAT	0	NA	0	NA
13	TATATT	2	169; 607	0	NA
14	TTTTAT	1	143	0	NA
15	ATATTT	1	608	0	NA
16	TATTAT	4	171; 549; 604; 1141	1	1141
17	TGTTTG	0	NA	0	NA
18	TTATAT	2	606; 958	0	NA
19	TGTAAT	1	300	0	NA
20	AAATAA	8	26; 192; 227; 275; 384; 809; 863; 1097	2	809; 863
	итого	31		5	
*NA = Не применимо					

ПРИМЕР 4

Синтетическая кодирующая область, кодирующая ядерный токсин Cry1Ab *Bacillus thuringiensis*

Последовательности для сравнения. Нативная последовательность ДНК, кодирующая ядерный токсин Cry1Ab, дана под SEQ ID NO:19. Данная последовательность была проанализирована с целью определения того, какие последовательности из идентифицированных в таблице 1 присутствуют в SEQ ID NO:19, и их положений. Аминокислотная последовательность, кодируемая SEQ ID NO:19, подвергалась обратной

трансляции с использованием целевых частот кодонов, приведенных в столбце таблицы 4 для синтетических генов, предназначенных для использования в кукурузе. Полученная в результате последовательность ДНК анализировалась, и кодоны заменялись при необходимости для удаления нежелательных открытых рамок считывания и удаления нежелательных сайтов рестрикции, при этом поддерживались все последовательности, идентифицированные в таблице 1. Аминокислотная последовательность, кодируемая SEQ ID NO:19, сохранялась. Полученная в результате последовательность ДНК дана под SEQ ID NO:21.

SEQ ID NO:21 была проанализирована, и кодоны были изменены с целью удаления потенциальных последовательностей сигналов полиаденилирования, идентифицированных в таблице 2, при этом поддерживалось то же самое количество последовательностей, идентифицированных в таблице 1. Полученная в результате последовательность, которая является вариантом осуществления настоящего изобретения, дана под SEQ ID NO:23. В таблице 11 показано, что количество и положение последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, поддерживается в SEQ ID NO:23. В таблице 12 показано, что количество последовательностей сигналов полиаденилирования, идентифицированных в таблицах 2 и 3, снижается в SEQ ID NO:23 по сравнению с SEQ ID NO:19.

Синтетическая кодирующая область SEQ ID NO:23 была оптимизирована для экспрессии в кукурузе.

Конструкция для применения в экспрессии синтетической кодирующей области SEQ ID NO:23 была сделана путем объединения синтетической кодирующей области SEQ ID NO:23 с 5'-нетранслируемой областью, содержащей промотор, функциональный в растительных клетках, и 3'-нетранслируемой области, содержащей последовательность терминации транскрипции и полиаденилирования.

В одном из вариантов осуществления такой конструкции, продуцирование первичного транскрипта мРНК, содержащего SEQ ID NO:23, управлялось копией промотора убиквитина 1 кукурузы со своим нативным интроном 1 (патент США No. 5510474). Фрагмент, содержащий 3'-нетранслируемую область из гена пероксидазы 5 кукурузы (ZmPer5 3'UTR v2; патент США No. 6699984), был использован для терминации транскрипции. Бинарная плазмида для трансформации растений, pDAB111449, содержащая указанную выше кассету экспрессии генов, была сконструирована и применена в продуцировании трансгенных растений кукурузы. Плазмида pDAB111449 также содержит ген устойчивости к гербицидам, содержащий кодирующую область для арилоксиалконат диоксигеназы (AAD-1 v3; патент США No. 7838733(B2), and Wright et al. (2010) Proc. Natl. Acad. Sci. U.S.A. 107:20240-5) под транскрипционным контролем промотора палочковидного баднавируса сахарного тростника (ScBV) (Schenk et al. (1999) Plant Molec. Biol. 39: 1221-30). Фрагмент, содержащий 3'-нетранслируемую область из гена липазы кукурузы (ZmLip 3'UTR; патент США No. 7179902), был использован для терминации транскрипции.

Таблица 11

Последовательности таблицы 1, обнаруженные в нативной кодирующей ядерный токсин Cry1Ab области (SEQ ID NO:19) и ее перестроенной версии (SEQ ID NO:23)

	Последовательность из таблицы 1	Число участков в нативной последовательности ядерного токсина Cry1Ab (SEQ ID NO:19)	Положение в нативной последовательности ядерного токсина Cry1Ab, н.п. (SEQ ID NO:19)	Число участков в перестроенной последовательности ядерного токсина Cry1Ab (SEQ ID NO:23)	Положение в перестроенной последовательности ядерного токсина Cry1Ab, н.п. (SEQ ID NO:23)
1	AATAAA	0	NA*	0	NA
2	AATAAT	3	960, 1126, 1387	3	960, 1126, 1387

3	AACCAA	2	253, 280	2	253, 280
4	ATATAA	2	185, 1391	2	185, 1391
5	AATCAA	2	688, 1129	3	688, 1129, 1639
6	ATACTA	0	NA	0	NA
7	ATAAAA	0	NA	0	NA
8	ATGAAA	1	1232	1	1232
9	AAGCAT	0	NA	0	NA
10	ATTAAT	1	1636	1	1636
11	ATACAT	2	1366, 1613	2	1366, 1613
12	AAAATA	0	NA	0	NA
13	ATGAAA	3	249, 704, 785	3	249, 704, 785
13	AATGAA	0	NA	0	NA
15	AATACA	0	NA	0	NA
16	CATAAA	0	NA	0	NA
	итого	16	NA	17	NA

*NA = Не применимо

Таблица 12

Последовательности таблицы 2, обнаруженные в нативной кодирующей области Cry1Ab (SEQ ID NO:19) и ее перестроенной версии (SEQ ID NO:23)

	Последовательность из таблицы 2	Число участков в нативной последовательности ядерного токсина Cry1Ab (SEQ ID NO:19)	Положение в нативной последовательности ядерного токсина Cry1Ab, н.п. (SEQ ID NO:19)	Число участков в перестроенной последовательности ядерного токсина Cry1Ab (SEQ ID NO:23)	Положение в перестроенной последовательности ядерного токсина Cry1Ab, н.п. (SEQ ID NO:23)
1	ATATAT	0	NA*	0	NA
2	TGTTTT	1	42	0	NA
3	TTTTGT	0	NA	0	NA
4	TGTTTT	0	NA	0	NA
5	TATATA	2	1097, 1792	0	NA
6	TATTTT	0	NA	0	NA
7	TTTTTT	0	NA	0	NA
8	ATTTTT	2	199, 1649	0	NA
9	TTATTT	0	NA	0	NA
10	TTTATT	1	470	0	NA
11	TAATAA	2	1340, 1386	0	NA
12	ATTTAT	2	503, 799	0	NA
13	TATATT	0	NA	0	NA
14	TTTTAT	0	NA	0	NA
15	ATATTT	1	110	0	NA
16	TATTAT	2	937, 940	0	NA
17	TGTTTG	1	530	0	NA
18	TTATAT	2	1096, 1791	0	NA
19	TGTAAT	0	NA	0	NA
20	AAATAA	2	959, 1125	1	959
	итого	18		1	

*NA = Не применимо

ПРИМЕР 5

Синтетическая кодирующая область, кодирующая ядерный токсин Cry1Ca *Bacillus thuringiensis*

Последовательности для сравнения. Нативная последовательность ДНК, кодирующая ядерный токсин Cry35A, дана под SEQ ID NO:25. Данная последовательность была проанализирована с целью определения того, какие последовательности из идентифицированных в таблице 1 присутствуют в SEQ ID NO:25, и их положений. Аминокислотная последовательность, кодируемая SEQ ID NO:25 подвергалась обратной трансляции с использованием целевых частот кодонов, приведенных в столбце таблицы

4 для синтетических генов, предназначенных для использования в кукурузе. Полученная в результате последовательность ДНК анализировалась, и кодоны заменялись при необходимости для удаления нежелательных открытых рамок считывания и удаления нежелательных сайтов распознавания рестрикционных ферментов, при этом поддерживались все последовательности, идентифицированные в таблице 1. Аминокислотная последовательность, кодируемая SEQ ID NO:25, сохранялась. Полученная в результате последовательность ДНК дана под SEQ ID NO:27. Данная последовательность будет синтезирована и использована для сравнения с синтетическим геном, спроектированным по изобретению.

10 SEQ ID NO:27 была проанализирована, и кодоны были изменены с целью удаления потенциальных последовательностей сигналов полиаденилирования, идентифицированных в таблице 2, при этом поддерживалось то же самое количество последовательностей, идентифицированных в таблице 1. Полученная в результате последовательность, которая является вариантом осуществления настоящего изобретения, дана под SEQ ID NO:29. В таблице 13 показано, что количество и положение последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, поддерживается в SEQ ID NO:29. В таблице 14 показано, что количество последовательностей сигналов полиаденилирования, идентифицированных в таблицах 2 и 3, снижается в SEQ ID NO:29 по сравнению с SEQ ID NO:25.

15 Синтезируется ДНК из SEQ ID NO:29, и уровни экспрессии, наблюдаемые в растительных клетках, трансформированных для экспрессии данной последовательности, сравниваются с уровнями экспрессии, наблюдаемыми в растительных клетках, трансформированных для экспрессии SEQ ID NO:25 и SEQ ID NO:27.

20 Синтетический ген SEQ ID NO:29 был оптимизирован для экспрессии в кукурузе.

25 Конструкция для применения в экспрессии синтетического гена SEQ ID NO:29 была сделана путем объединения синтетического гена SEQ ID NO:29 с 5'-нетранслируемой областью, содержащей промотор, функциональный в растительных клетках, и 3'-нетранслируемой области, содержащей терминатор транскрипции и последовательность полиаденилирования.

Таблица 13

Последовательности таблицы 1, обнаруженные в нативной кодирующей Cry1Ca области (SEQ ID NO:25) и ее перестроенной версии (SEQ ID NO:29)						
	Последовательность из таблицы 1	Число участков в нативной последовательности ядерного токсина Cry1Ca (SEQ ID NO:25)	Положение в нативной последовательности ядерного токсина Cry1Ca, н.п. (SEQ ID NO:25)	Число участков в перестроенной последовательности ядерного токсина Cry1Ca (SEQ ID NO:29)	Положение в перестроенной последовательности ядерного токсина Cry1Ca, н.п. (SEQ ID NO:29)	
35	1	AATAAA	0	NA*	0	NA
	2	AATAAT	2	646, 916	2	646, 916
	3	AACCAA	0	NA	1	1042
40	4	ATATAA	2	684, 1757	2	684, 1757
	5	AATCAA	1	1405	1	1405
	6	ATACTA	0	NA	0	NA
	7	ATAAAA	1	1826	1	1826
	8	ATGAAA	2	254, 569	2	254, 569
	9	AAGCAT	1	335	1	335
45	10	ATTAAT	7	177, 246, 250, 813, 817, 1402, 1534	7	177, 246, 250, 813, 817, 1402, 1534
	11	ATACAT	0	NA	0	NA
	12	AAAATA	0	NA	0	NA
	13	ATTA AAA	4	245, 249, 816, 1401	4	245, 249, 816, 1401

13	ААТГАА	1	642	1	642
15	ААТАСА	1	1381	1	1381
16	САТААА	0	NA	0	NA
	итого	22		23	

*NA = Не применимо

5

Таблица 14

Последовательности таблицы 2, обнаруженные в нативной кодирующей ядерный токсин Cry1Ca области (SEQ ID NO:25) и ее перестроенной версии (SEQ ID NO:29)

10

	Последовательность из таблицы 2	Число участков в нативной последовательности ядерного токсина Cry1Ca (SEQ ID NO:25)	Положение в нативной последовательности ядерного токсина Cry1Ca, н.п. (SEQ ID NO:25)	Число участков в перестроенной последовательности ядерного токсина Cry1Ca (SEQ ID NO:29)	Положение в перестроенной последовательности ядерного токсина Cry1Ca, н.п. (SEQ ID NO:29)	
	1	АГАТАГ	4	323, 325, 908, 1024	0	NA*
	2	ТТГТТТ		NA	0	NA
	3	ТТТТГТ	3	186, 1302, 1512	0	NA
	4	ТГТТТТ	0	NA	0	NA
15	5	ТАТАГА	3	324, 1023, 1819	0	NA
	6	ТАТТТТ	1	1346	0	NA
	7	ТТТТТТ	1	1326	0	NA
	8	АГТТТТ	2	529, 959	0	NA
	9	ТТАТТТ	1	901	0	NA
20	10	ТТТАТГ	2	900, 962	0	NA
	11	ТААТАА	0	NA	0	NA
	12	АГТТАГ	1	899	0	NA
	13	ТАТАТГ	2	510, 909	0	NA
	14	ТТТТАГ	2	470, 961	0	NA
	15	АТАТТТ	1	110	0	NA
	16	ТАТТАГ	0	NA	0	NA
25	17	ТГТТТГ	0	NA	0	NA
	18	ТТАТАГ	1	1818	0	NA
	19	ТГТААТ	1	525	0	NA
	20	АААТАА	1	645	1	645
		итого	26		1	

*NA = Не применимо

30

ПРИМЕР 6

Синтетическая кодирующая область, кодирующая токсин Cry6Aa *Bacillus thuringiensis*

Последовательности для сравнения. Нативная последовательность ДНК, кодирующая токсин Cry6Aa, дана под SEQ ID NO:31. Данная последовательность была

35

проанализирована с целью определения того, какие последовательности из идентифицированных в таблице 1 присутствуют в SEQ ID NO:31, и их положений.

Аминокислотная последовательность, кодируемая SEQ ID NO:31 подвергалась обратной трансляции с использованием целевых частот кодонов, приведенных в столбце таблицы 4 для синтетических генов, предназначенных для использования в кукурузе. Полученная

40

в результате последовательность ДНК анализировалась, и кодоны заменялись при необходимости для удаления нежелательных открытых рамок считывания и удаления нежелательных сайтов распознавания рестрикционных ферментов, при этом поддерживались все последовательности, идентифицированные в таблице 1.

Аминокислотная последовательность, кодируемая SEQ ID NO:31, сохранялась.

45

Полученная в результате последовательность ДНК дана под SEQ ID NO:33. Данная последовательность была синтезирована и использована для сравнения с синтетическим геном, спроектированным по изобретению.

SEQ ID NO:33 была проанализирована, и кодоны были изменены с целью удаления потенциальных последовательностей сигналов полиаденилирования,

идентифицированных в таблице 2, при этом поддерживалось количество последовательностей, идентифицированных в таблице 1. Полученная в результате последовательность, которая является вариантом осуществления настоящего изобретения, дана под SEQ ID NO:35. В таблице 15 показано, что количество и положение последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, поддерживается в SEQ ID NO:35. В таблице 16 показано, что количество последовательностей сигналов полиаденилирования, идентифицированных в таблицах 2 и 3, снижается в SEQ ID NO:35 по сравнению с SEQ ID NO:31.

Синтезируется ДНК из SEQ ID NO:35, и уровни экспрессии, наблюдаемые в растительных клетках, трансформированных для экспрессии данной последовательности, сравниваются с уровнями экспрессии, наблюдаемыми в растительных клетках, трансформированных для экспрессии SEQ ID NO:31 и SEQ ID NO:33.

Синтетическая кодирующая область SEQ ID NO:35 была оптимизирована для экспрессии в кукурузе.

Конструкция для применения в экспрессии синтетической кодирующей области SEQ ID NO:35 сделана путем объединения синтетической кодирующей области SEQ ID NO:35 с 5'-нетранслируемой областью, содержащей промотор, функциональный в растительных клетках и 3'-нетранслируемой области, содержащей терминатор транскрипции и последовательность полиаденилирования.

Таблица 15

Последовательности таблицы 1, обнаруженные в нативной кодирующей области СгубАа (SEQ ID NO:31) и ее перестроенной версии (SEQ ID NO:35)

	Последовательность из таблицы 1	Число участков в нативной последовательности СгубАа (SEQ ID NO:31)	Положение в нативной последовательности СгубАа, н.п. (SEQ ID NO:31)	Число участков в перестроенной последовательности СгубАа (SEQ ID NO:35)	Положение в перестроенной последовательности СгубАа, н.п. (SEQ ID NO:35)
1	ААТААА	1	292	1	292
2	ААТААТ	6	430, 1309, 1360, 1384, 1402, 1420	6	430, 1309, 1360, 1384, 1402, 1420
3	ААССАА	0	NA*	0	NA
4	АТАТАА	2	824, 1344	2	824, 1344
5	ААТСАА	5	103, 634, 832, 1234, 1270	5	103, 634, 832, 1234, 1270
6	АТАСТА	0	NA	0	NA
7	АТАААА	3	269, 293, 826	3	269, 293, 826
8	АТГААА	1	794	1	794
9	ААГСАТ	0	NA	0	NA
10	АТТААТ	2	919, 1183	2	919, 1183
11	АТАСАТ	0	NA	1	1275
12	ААААТА	3	530, 806, 1358	3	530, 806, 1358
13	АТТААА	5	51, 56, 188, 495, 963	5	51, 56, 188, 495, 963
14	ААТГАА	7	52, 57, 316, 463, 496, 718, 964	7	52, 57, 316, 463, 496, 718, 964
15	ААТАСА	2	922, 1238	3	922, 1238, 1274
16	САТААА	1	664	1	664
	итого	38		40	

*NA = Не применимо

Таблица 16

Последовательности таблицы 2, обнаруженные в нативной кодирующей области СгубАа (SEQ ID NO:31) и ее перестроенной версии (SEQ ID NO:35)

	Последовательность из таблицы 2	Число участков в нативной последовательности СгубАа (SEQ ID NO:31)	Положение в нативной последовательности СгубАа, н.п. (SEQ ID NO:31)	Число участков в перестроенной последовательности СгубАа (SEQ ID NO:35)	Положение в перестроенной последовательности СгубАа, н.п. (SEQ ID NO:35)
1	АТАТАТ	4	147, 218, 1275, 1372	0	NA*

2	TTGTTT	1	788	0	NA
3	TTTTGT		NA	0	NA
4	TGTTTT	0	NA	0	NA
5	TATATA	1	941	0	NA
6	TATTTT	2	388, 489	0	NA
7	TTTTTT		NA	0	NA
8	ATTTTT	2	236, 555	0	NA
9	TGATTT	1	113	0	NA
10	TGTATT	1	109, 257	0	NA
11	TAATAA	5	66, 429, 1383, 1401, 1419	0	NA
12	ATTTAT	3	108, 299, 938	0	NA
13	TATATT	2	148, 1373	0	NA
14	TTTTAT	2	1314, 1365	0	NA
15	ATATTT	1	387	0	NA
16	TATPAT	1	111	0	NA
17	TGTTTG	0	NA	0	NA
18	TTATAT	4	247, 301, 940, 1190	0	NA
19	TGTAAT	1	1204	0	NA
20	AAATAA	2	1308, 1359	1	1359
	итого	33		1	

*NA = Не применимо

ПРИМЕР 7

Синтетическая кодирующая область, кодирующая AAD1 Sphingobium herbicidovorans

Последовательности для сравнения. Нативная последовательность ДНК, кодирующая белок AAD1, дана под SEQ ID NO:37. Данная последовательность была проанализирована с целью определения того, какие последовательности из идентифицированных в таблице 1 присутствуют в SEQ ID NO:37, и их положений.

Аминокислотная последовательность, кодируемая SEQ ID NO:37, подвергалась обратной трансляции с использованием целевых частот кодонов, приведенных в столбце таблицы 4 для синтетических генов, предназначенных для использования в кукурузе. Полученная в результате последовательность ДНК анализировалась, и кодоны заменялись при необходимости для удаления нежелательных открытых рамок считывания и удаления нежелательных сайтов распознавания рестрикционных ферментов, при этом поддерживались все последовательности, идентифицированные в таблице 1. Аминокислотная последовательность, кодируемая SEQ ID NO:37, сохранялась. Полученная в результате последовательность ДНК дана под SEQ ID NO:39. Данная последовательность была синтезирована и использована для сравнения с синтетическим геном, спроектированным по изобретению.

SEQ ID NO:39 была проанализирована, и кодоны были изменены с целью удаления потенциальных последовательностей сигналов полиаденилирования, идентифицированных в таблице 2, при этом поддерживалось количество последовательностей, идентифицированных в таблице 1. Полученная в результате последовательность, которая является вариантом осуществления настоящего изобретения, дана под SEQ ID NO:41. В таблице 17 показано, что количество и положение последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, поддерживается в SEQ ID NO:41. В таблице 18 показано, что количество последовательностей сигналов полиаденилирования, идентифицированных в таблицах 2 и 3, снижается в SEQ ID NO:41 по сравнению с SEQ ID NO:37.

Синтезируется ДНК из SEQ ID NO:41, и уровни экспрессии, наблюдаемые в растительных клетках, трансформированных для экспрессии данной последовательности, сравниваются с уровнями экспрессии, наблюдаемыми в

растительных клетках, трансформированных для экспрессии SEQ ID NO:37 и SEQ ID NO:39.

Синтетическая кодирующая область SEQ ID NO:41 была оптимизирована для экспрессии в кукурузе.

- 5 Конструкция для применения в экспрессии синтетической кодирующей области SEQ ID NO:41 сделана путем объединения синтетической кодирующей области SEQ ID NO:41 с 5'-нетранслируемой областью, содержащей промотор, функциональный в растительных клетках, и 3'-нетранслируемой области, содержащей терминатор транскрипции и последовательность полиаденилирования.

10

Таблица 17						
Последовательности таблицы 1, обнаруженные в нативной кодирующей области AAD1 (SEQ ID NO:37) и ее перестроенной версии (SEQ ID NO:41)						
	Последовательность из таблицы 1	Число участков в нативной последовательности AAD1 (SEQ ID NO:37)	Положение в нативной последовательности AAD1, н.п. (SEQ ID NO:37)	Число участков в перестроенной последовательности AAD1 (SEQ ID NO:41)	Положение в перестроенной последовательности AAD1, н.п. (SEQ ID NO:41)	
15	1	ААГААА	0	NA*	0	NA
	2	ААГААТ	0	NA	0	NA
	3	ААССАА	0	NA	1	652
	4	АГАГАА	0	NA	0	NA
	5	ААТСАА	0	NA	0	NA
20	6	АГАСТА	0	NA	0	NA
	7	АГАААА	0	NA	0	NA
	8	АТГААА	0	NA	0	NA
	9	ААГСАТ	0	NA	0	NA
	10	АТГААТ	0	NA	0	NA
	11	АТАСАТ	0	NA	0	NA
25	12	ААААТА	0	NA	0	NA
	13	АТГААА	0	NA	0	NA
	14	ААТГАА	0	NA	0	NA
	15	ААТАСА	0	NA	0	NA
	16	САГААА	0	NA	0	NA
		итого	0		1	
30	*NA = Не применимо					

Таблица 18						
Последовательности таблицы 2, обнаруженные в нативной кодирующей области AAD1 (SEQ ID NO:37) и ее перестроенной версии (SEQ ID NO:41)						
	Последовательность из таблицы 2	Число участков в нативной последовательности AAD1 (SEQ ID NO:37)	Положение в нативной последовательности AAD1, н.п. (SEQ ID NO:37)	Число участков в перестроенной последовательности AAD1 (SEQ ID NO:41)	Положение в перестроенной последовательности AAD1, н.п. (SEQ ID NO:41)	
35	1	АТАГАТ	0	NA*	0	NA
	2	ТТГТТТ	0	NA	0	NA
	3	ТТТТГТ	0	NA	0	NA
	4	ТГТТТТ	0	NA	0	NA
40	5	ТАГАТА	0	NA	0	NA
	6	ТАТТТТ	1	166	0	NA
	7	ТТТТТТ	0	NA	0	NA
	8	АТТТТТ	0	NA	0	NA
	9	ТТАТТТ	0	NA	0	NA
	10	ТТТАТТ	0	NA	0	NA
45	11	ТААГАА	0	NA	0	NA
	12	АТТТАТ	0	NA	0	NA
	13	ТАГАТТ	0	NA	0	NA
	14	ТТТТАТ	0	NA	0	NA
	15	АТАТТТ	0	NA	0	NA

16	ТАТТАТ	0	NA	0	NA
17	TGTTTG	0	NA	0	NA
18	ТТАТАТ	0	NA	0	NA
19	TGТААТ	0	NA	0	NA
20	АААТАА	0	NA	0	NA
	итого	1		0	

*NA = Не применимо

ПРИМЕР 8

Синтетическая кодирующая область, кодирующая дельта-9-десатуразу *Aspergillus nidulans*

Последовательности для сравнения. Нативная последовательность ДНК, кодирующая белок дельта-9-десатуразы *Aspergillus nidulans*, дана под SEQ ID NO:43. Данная последовательность была проанализирована с целью определения того, какие последовательности из идентифицированных в таблице 1 присутствуют в SEQ ID NO: 43, и их положений. Аминокислотная последовательность, кодируемая SEQ ID NO:43, подвергалась обратной трансляции с использованием целевых частот кодонов, приведенных в столбце таблицы 4 для синтетических генов, предназначенных для использования в кукурузе. Полученная в результате последовательность ДНК анализировалась, и кодоны заменялись при необходимости для удаления нежелательных открытых рамок считывания и удаления нежелательных сайтов рестрикции, при этом поддерживались все последовательности, идентифицированные в таблице 1. Аминокислотная последовательность, кодируемая SEQ ID NO:43, сохранялась. Полученная в результате последовательность ДНК дана под SEQ ID NO:45. Данная последовательность была синтезирована и использована для сравнения с синтетическим геном, спроектированным по изобретению.

SEQ ID NO:45 была проанализирована, и кодоны были изменены с целью удаления потенциальных последовательностей сигналов полиаденилирования, идентифицированных в таблице 2, при этом поддерживалось количество последовательностей, идентифицированных в таблице 1. Полученная в результате последовательность, которая является вариантом осуществления настоящего изобретения, дана под SEQ ID NO:47. В таблице 1 показано, что количество и положение последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, поддерживается в SEQ ID NO:47. В таблице 20 показано, что количество последовательностей сигналов полиаденилирования, идентифицированных в таблицах 2 и 3, снижается в SEQ ID NO:47 по сравнению с SEQ ID NO:43.

Синтезируется ДНК из SEQ ID NO:47, и уровни экспрессии, наблюдаемые в растительных клетках, трансформированных для экспрессии данной последовательности, сравниваются с уровнями экспрессии, наблюдаемыми в растительных клетках, трансформированных для экспрессии SEQ ID NO:43 и SEQ ID NO:45.

Синтетическая кодирующая область SEQ ID NO:47 была оптимизирована для экспрессии в кукурузе.

Конструкция для применения в экспрессии синтетической кодирующей области SEQ ID NO:47 сделана путем объединения синтетической кодирующей области SEQ ID NO: 47 с 5'-нетранслируемой областью, содержащей промотор, функциональный в растительных клетках, и 3'-нетранслируемой области, содержащей последовательность терминации транскрипции и полиаденилирования.

Таблица 19

Последовательности таблицы 1, обнаруженные в нативной кодирующей области дельта-9-десатуразы <i>Aspergillus nidulans</i> (SEQ ID NO:43) и ее перестроенной версии (SEQ ID NO:47)						
	Последовательность из таблицы 1	Число участков в нативной последовательности Asp-Δ9 (SEQ ID NO:43)	Положение в нативной последовательности Asp-Δ9, н.п. (SEQ ID NO: 43)	Число участков в перестроенной последовательности Asp-Δ9 (SEQ ID NO:47)	Положение в перестроенной последовательности Asp-Δ9, н.п. (SEQ ID NO:47)	
5	1	AATAAA	0	NA*	0	NA
	2	AATAAT	0	NA	0	NA
	3	AACCAA	1	1326	1	1326
	4	ATATAA	0	NA	0	NA
	5	AATCAA	0	NA	0	NA
10	6	ATACTA	0	NA	0	NA
	7	ATAAAA	0	NA	0	NA
	8	ATGAAA	0	NA	0	NA
	9	AAGCAT	1	94	1	94
	10	ATTAAT	0	NA	0	NA
	11	ATACAT	0	NA	0	NA
15	12	AAAATA	0	NA	0	NA
	13	ATGAAA	0	NA	0	NA
	14	AATGAA	0	NA	0	NA
	15	AATACA	0	NA	0	NA
	16	CATAAA	0	NA	0	NA
		итого	2		2	

*NA = Не применимо

Таблица 20						
Последовательности таблицы 2, обнаруженные в нативной кодирующей области дельта-9-десатуразы <i>Aspergillus nidulans</i> (SEQ ID NO:43) и ее перестроенной версии (SEQ ID NO:47)						
	Последовательность из таблицы 2	Число участков в нативной последовательности Asp-Δ9 (SEQ ID NO:43)	Положение в нативной последовательности Asp-Δ9, н.п. (SEQ ID NO:43)	Число участков в перестроенной последовательности Asp-Δ9 (SEQ ID NO: 47)	Положение в перестроенной последовательности Asp-Δ9, н.п. (SEQ ID NO:47)	
25	1	АТАГАТ	0	NA*	0	NA
	2	ТТГТТТ	0	NA	0	NA
	3	ТТТТГТ	0	NA	0	NA
	4	ТГТТТТ	0	NA	0	NA
30	5	ТАГАГА	0	NA	0	NA
	6	ТАТТТТ	1	166	0	NA
	7	ТТТТТТ	0	NA	0	NA
	8	АТТТТТ	0	NA	0	NA
	9	ТТАТТТ	0	NA	0	NA
	10	ТТТАТТ	0	NA	0	NA
35	11	ТААГАА	0	NA	0	NA
	12	АТТТАТ	0	NA	0	NA
	13	ТАТАТТ	0	NA	0	NA
	14	ТТТТАТ	0	NA	0	NA
	15	АТАТТТ	1	479	0	NA
	16	ТАТТАТ	0	NA	0	NA
40	17	ТГТТТГ	0	NA	0	NA
	18	ТТАТАТ	0	NA	0	NA
	19	ТГТААТ	0	NA	0	NA
	20	АААГАА	0	NA	0	NA
		итого	1		0	

*NA = Не применимо

ПРИМЕР 9

Синтетическая кодирующая область, кодирующая SAP1 *Xerophyta viscosa*
 Последовательности для сравнения. Нативная последовательность ДНК, кодирующая белок SAP1 *Xerophyta viscosa*, дана под SEQ ID NO:49. Данная последовательность была

проанализирована с целью определения того, какие последовательности из идентифицированных в таблице 1 присутствуют в SEQ ID NO:49, и их положений. Аминокислотная последовательность, кодируемая SEQ ID NO:49 подвергалась обратной трансляции с использованием целевых частот кодонов, приведенных в столбце таблицы 4 для синтетических генов, предназначенных для использования в кукурузе. Полученная в результате последовательность ДНК анализировалась, и кодоны заменялись при необходимости для удаления нежелательных открытых рамок считывания и удаления нежелательных сайтов рестрикции, при этом поддерживались все последовательности, идентифицированные в таблице 1. Аминокислотная последовательность, кодируемая SEQ ID NO:49, сохранялась. Полученная в результате последовательность ДНК дана под SEQ ID NO:51. Данная последовательность была синтезирована и использована для сравнения с синтетическим геном, спроектированным по изобретению.

SEQ ID NO:52 была проанализирована, и кодоны были изменены с целью удаления потенциальных последовательностей сигналов полиаденилирования, идентифицированных в таблице 2, при этом поддерживалось количество последовательностей, идентифицированных в таблице 1. Полученная в результате последовательность, которая является вариантом осуществления настоящего изобретения, дана под SEQ ID NO:53. В таблице 1 показано, что количество и положение последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, поддерживается в SEQ ID NO:53. В таблице 21 показано, что количество последовательностей сигналов полиаденилирования, идентифицированных в таблицах 2 и 3, снижается в SEQ ID NO:53 по сравнению с SEQ ID NO:49.

Синтезируется ДНК из SEQ ID NO:53, и уровни экспрессии, наблюдаемые в растительных клетках, трансформированных для экспрессии данной последовательности, сравниваются с уровнями экспрессии, наблюдаемыми в растительных клетках, трансформированных для экспрессии SEQ ID NO:49 и SEQ ID NO:51.

Синтетическая кодирующая область SEQ ID NO:53 была оптимизирована для экспрессии в кукурузе.

Конструкция для применения в экспрессии синтетической кодирующей области SEQ ID NO:53 сделана путем объединения синтетической кодирующей области SEQ ID NO:53 с 5'-нетранслируемой областью, содержащей промотор, функциональный в растительных клетках, и 3'-нетранслируемой области, содержащей терминатор транскрипции и последовательность полиаденилирования.

Таблица 21					
Последовательности таблицы 1, обнаруженные в нативной кодирующей области SAP1 <i>Xerophyta viscosa</i> (SEQ ID NO:49) и ее перестроенной версии (SEQ ID NO:53)					
	Последовательность из таблицы 1	Число участков в нативной последовательности XvSAP1 (SEQ ID NO:49)	Положение в нативной последовательности XvSAP1, н.п. (SEQ ID NO:49)	Число участков в перестроенной последовательности XvSAP1 (SEQ ID NO:53)	Положение в перестроенной последовательности XvSAP1, н.п. (SEQ ID NO:53)
1	AATAAA	0	NA*	0	NA
2	AATAAT	0	NA	0	NA
3	AACCAA	0	NA	0	NA
4	ATATAA	0	NA	0	NA
5	AATCAA	0	NA	0	NA
6	ATACAA	0	NA	0	NA
7	ATAAAA	0	NA	0	NA
8	ATGAAA	0	NA	1	25
9	AAGCAT	0	NA	0	NA
10	ATTAAT	0	NA	0	NA

11	ATACAT	0	NA	0	NA
12	AAAATA	0	NA	0	NA
13	ATTAAA	0	NA	0	NA
14	AATGAA	0	NA	0	NA
15	AATACA	0	NA	0	NA
16	CATAAA	0	NA	0	NA
	итого	0		1	

*NA = Не применимо

Таблица 22

Последовательности таблицы 2, обнаруженные в нативной кодирующей области SAP1 *Xerophyta viscosa* (SEQ ID NO:49) и ее перестроенной версии (SEQ ID NO:53)

	Последовательность из таблицы 2	Число участков в нативной последовательности XvSAP1 (SEQ ID NO:49)	Положение в нативной последовательности XvSAP1, н.п. (SEQ ID NO:49)	Число участков в перестроенной последовательности XvSAP1 (SEQ ID NO:53)	Положение в перестроенной последовательности XvSAP1, н.п. (SEQ ID NO:53)
1	ATATAT	0	NA*	0	NA
2	TGTGTT	0	NA	0	NA
3	TTTTGT	0	NA	0	NA
4	TGTGTT	0	NA	0	NA
5	TATATA	0	NA	0	NA
6	TATGTT	1	755	0	NA
7	TTTTTT	0	NA	0	NA
8	ATTTTT	1	756	0	NA
9	TTAGTT	0	NA	0	NA
10	TTGATG	0	NA	0	NA
11	TAATAA	0	NA	0	NA
12	ATTTAT	0	NA	0	NA
13	TAGATT	0	NA	0	NA
14	TTGATG	0	NA	0	NA
15	ATATTT	1	754	0	NA
16	TATGAT	1	665	0	NA
17	TGTTTG	1	696	0	NA
18	TTATAT	0	NA	0	NA
19	TGTAAT	0	NA	0	NA
20	AAATAA	0	NA	0	NA
	итого	5		0	

*NA = Не применимо

ПРИМЕР 10

Синтетическая кодирующая область, кодирующая GFP1 *Aequorea victoria*

Последовательности для сравнения. Нативная последовательность ДНК, кодирующая GFP1 *Aequorea victoria*, дана под SEQ ID NO:55. Данная последовательность была проанализирована с целью определения того, какие последовательности из идентифицированных в таблице 1 присутствуют в SEQ ID NO:55, и их положений. Аминокислотная последовательность, кодируемая SEQ ID NO:55, подвергалась обратной трансляции с использованием целевых частот кодонов, приведенных в столбце таблицы 4 для синтетических генов, предназначенных для использования в кукурузе. Полученная в результате последовательность ДНК анализировалась, и кодоны заменялись при необходимости для удаления нежелательных открытых рамок считывания и удаления нежелательных сайтов рестрикции, при этом поддерживались все последовательности, идентифицированные в таблице 1. Аминокислотная последовательность, кодируемая SEQ ID NO:55, сохранялась. Полученная в результате последовательность ДНК дана под SEQ ID NO:57. Данная последовательность была синтезирована и использована для сравнения с синтетическим геном, спроектированным по изобретению.

SEQ ID NO:57 была проанализирована, и кодоны были изменены с целью удаления

потенциальных последовательностей сигналов полиаденилирования, идентифицированных в таблице 2, при этом поддерживалось количество последовательностей, идентифицированных в таблице 1. Полученная в результате последовательность, которая является вариантом осуществления настоящего изобретения, дана под SEQ ID NO:59. В таблице 1 показано, что количество и положение последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, поддерживается в SEQ ID NO:59. В таблице 23 показано, что количество последовательностей сигналов полиаденилирования, идентифицированных в таблицах 2 и 3, снижается в SEQ ID NO:59 по сравнению с SEQ ID NO:55.

Синтезируется ДНК из ID NO:59, и уровни экспрессии, наблюдаемые в растительных клетках, трансформированных для экспрессии данной последовательности, сравниваются с уровнями экспрессии, наблюдаемыми в растительных клетках, трансформированных для экспрессии SEQ ID NO:55 и SEQ ID NO:57.

Синтетическая кодирующая область SEQ ID NO:59 была оптимизирована для экспрессии в кукурузе.

Конструкция для применения в экспрессии синтетической кодирующей области SEQ ID NO:59 сделана путем объединения синтетической кодирующей области SEQ ID NO:59 с 5'-нетранслируемой областью, содержащей промотор, функциональный в растительных клетках, и 3'-нетранслируемой области, содержащей терминатор транскрипции и последовательность полиаденилирования.

Таблица 23					
Последовательности таблицы 1, обнаруженные в нативной кодирующей области GFP1 <i>Aequorea victoria</i> (SEQ ID NO:55) и ее перестроенной версии (SEQ ID NO:59)					
	Последовательность из таблицы 1	Число участков в нативной последовательности GFP1 (SEQ ID NO:55)	Положение в нативной последовательности GFP1, н.п. (SEQ ID NO:55)	Число участков в перестроенной последовательности GFP1 (SEQ ID NO:59)	Положение в перестроенной последовательности GFP1, н.п. (SEQ ID NO:59)
1	ААТААА	0	NA*	0	NA
2	ААТААТ	0	NA	0	NA
3	ААССАА	1	467	1	467
4	АТАТАА	0	NA	0	NA
5	ААТСАА	0	NA	0	NA
6	АТАСТА	0	NA	0	NA
7	АТАААА	0	NA	0	NA
8	АТГААА	1	237	1	237
9	ААГСАТ	0	NA	0	NA
10	АТТААТ	0	NA	0	NA
11	АТАСАТ	1	450	1	450
12	ААААТА	1	551	1	551
13	АТТААА	1	511	1	511
14	ААТГАА	0	NA	0	NA
15	ААТАСА	1	425	1	425
16	САТААА	0	NA	1	480
	итого	6		7	

*NA = Не применимо

Таблица 24					
Последовательности таблицы 2, обнаруженные в нативной кодирующей области GFP1 <i>Aequorea victoria</i> (SEQ ID NO:55) и ее перестроенной версии (SEQ ID NO:59)					
	Последовательность из таблицы 2	Число участков в нативной последовательности GFP1 (SEQ ID NO:55)	Положение в нативной последовательности GFP1, н.п. (SEQ ID NO:55)	Число участков в перестроенной последовательности GFP1 (SEQ ID NO:59)	Положение в перестроенной последовательности GFP1, н.п. (SEQ ID NO:59)
1	АТАТАТ	0	NA*	0	NA
2	ТТГТТТ	0	NA	0	NA

3	TTTTGT	0	NA	0	NA
4	TGTTTT	0	NA	0	NA
5	TATATA	0	NA	0	NA
6	TATTTT	1	293	0	NA
7	TTTTTT	0	NA	0	NA
8	ATTTTT	0	NA	0	NA
9	TTATTT	1	137	0	NA
10	TTTATTT	1	136	0	NA
11	TAATAA	0	NA	0	NA
12	ATTTAT	0	NA	0	NA
13	TATATT	1	291	0	NA
14	TTTTAT	1	135	0	NA
15	ATATTT	1	292	0	NA
16	TATGAT	0	NA	0	NA
17	TGTTTG	0	NA	0	NA
18	TTATAT	0	NA	0	NA
19	TGTAAT	0	NA	0	NA
20	AAATAA	0	NA	0	NA
	итого	6		0	

*NA = Не применимо

ПРИМЕР 11

Синтетическая кодирующая область, кодирующая FAD9 *Leptosphaeria nodorum*

Последовательности для сравнения. Нативная последовательность ДНК, кодирующая белок FAD9 *Leptosphaeria nodorum*, дана под SEQ ID NO:61. Данная последовательность была проанализирована с целью определения того, какие последовательности из идентифицированных в таблице 1 присутствуют в SEQ ID NO:61, и их положений.

Аминокислотная последовательность, кодируемая SEQ ID NO:61, подвергалась обратной трансляции с использованием целевых частот кодонов, приведенных в столбце таблицы 4 для синтетических генов, предназначенных для использования в кукурузе. Полученная в результате последовательность ДНК анализировалась, и кодоны заменялись при необходимости для удаления нежелательных открытых рамок считывания и удаления нежелательных сайтов рестрикции, при этом поддерживались все последовательности, идентифицированные в таблице 1. Аминокислотная последовательность, кодируемая SEQ ID NO:61, сохранялась. Полученная в результате последовательность ДНК дана под SEQ ID NO:63. Данная последовательность была синтезирована и использована для сравнения с синтетическим геном, спроектированным по изобретению.

SEQ ID NO:63 была проанализирована, и кодоны были изменены с целью удаления потенциальных последовательностей сигналов полиаденилирования, идентифицированных в таблице 2, при этом поддерживалось количество последовательностей, идентифицированных в таблице 1. Полученная в результате последовательность, которая является вариантом осуществления настоящего изобретения, дана под SEQ ID NO:65. В таблице 1 показано, что количество и положение последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, поддерживается в SEQ ID NO:65. В таблице 25 показано, что количество последовательностей сигналов полиаденилирования, идентифицированных в таблицах 2 и 3, снижается в SEQ ID NO:65 по сравнению с SEQ ID NO:61.

Синтезируется ДНК из ID NO:65, и уровни экспрессии, наблюдаемые в растительных клетках, трансформированных для экспрессии данной последовательности, сравниваются с уровнями экспрессии, наблюдаемыми в растительных клетках, трансформированных для экспрессии SEQ ID NO:61 и SEQ ID NO:63.

Синтетическая кодирующая область SEQ ID NO:65 была оптимизирована для

экспрессии в кукурузе.

Конструкция для применения в экспрессии синтетической кодирующей области SEQ ID NO:65 сделана путем объединения синтетической кодирующей области SEQ ID NO: 65 с 5'-нетранслируемой областью, содержащей промотор, функциональный в растительных клетках, и 3'-нетранслируемой области, содержащей терминатор транскрипции и последовательность полиаденилирования.

Таблица 25					
Последовательности таблицы 1, обнаруженные в нативной кодирующей области FAD9 <i>Leptosphaeria nodorum</i> (SEQ ID NO:61) и ее перестроенной версии (SEQ ID NO:65)					
	Последовательность из таблицы 1	Число участков в нативной последовательности Ln FAD9 (SEQ ID NO:61)	Положение в нативной последовательности Ln FAD9, н.п. (SEQ ID NO:61)	Число участков в перестроенной последовательности Ln FAD9 (SEQ ID NO:65)	Положение в перестроенной последовательности Ln FAD9, н.п. (SEQ ID NO:65)
1	AATAAA	0	NA*	0	NA
2	AATAAT	0	NA	0	NA
3	AACCAA	0	NA	0	NA
4	ATATAA	0	NA	0	NA
5	AATCAA	0	NA	0	NA
6	ATACTA	0	NA	0	NA
7	ATAAAA	0	NA	0	NA
8	ATGAAA	0	NA	0	NA
9	AAGCAT	0	NA	0	NA
10	ATTAAT	0	NA	0	NA
11	ATACAT	0	NA	0	NA
12	AAAATA	0	NA	0	NA
13	ATGAAA	0	NA	0	NA
14	AATGAA	0	NA	0	NA
15	AATACA	0	NA	0	NA
16	CATAAA	0	NA	0	NA
	итого	0		0	

*NA = Не применимо

Таблица 26					
Последовательности таблицы 2, обнаруженные в нативной кодирующей области FAD9 <i>Leptosphaeria nodorum</i> (SEQ ID NO:61) и ее перестроенной версии (SEQ ID NO:65)					
	Последовательность из таблицы 2	Число участков в нативной последовательности Ln FAD9 (SEQ ID NO:61)	Положение в нативной последовательности Ln FAD9, н.п. (SEQ ID NO:61)	Число участков в перестроенной последовательности Ln FAD9 (SEQ ID NO:65)	Положение в перестроенной последовательности Ln FAD9, н.п. (SEQ ID NO:65)
1	ATATAT	0	NA*	0	NA
2	TTGTTT	0	NA	0	NA
3	TTTTGT	0	NA	0	NA
4	TGTTTT	1	1275	0	NA
5	TATAGA	0	NA	0	NA
6	TATTTT	0	NA	0	NA
7	TTTTTT	0	NA	0	NA
8	ATTTTT	0	NA	0	NA
9	TTAGTT	0	NA	0	NA
10	TTTAGT	1	1090	0	NA
11	TAAGAA	0	NA	0	NA
12	ATTTAT	0	NA	0	NA
13	TATATT	0	NA	0	NA
14	TTTTAT	0	NA	0	NA
15	ATATTT	0	NA	0	NA
16	TATTTT	1	416	0	NA
17	TGTTTG	0	NA	0	NA
18	TTATAT	0	NA	0	NA

19	TGTAAT	0	NA	0	NA
20	AAATAA	0	NA	0	NA
	итого	3		0	

*NA = Не применимо

5 ПРИМЕР 12

Синтетическая кодирующая область, кодирующая PER1 *Xerophyta viscosa*

Последовательности для сравнения. Нативная последовательность ДНК, кодирующая белок PER1 *Xerophyta viscosa*, дана под SEQ ID NO:67. Данная последовательность была проанализирована с целью определения того, какие последовательности из

10 идентифицированных в таблице 1 присутствуют в SEQ ID NO:67, и их положений. Аминокислотная последовательность, кодируемая SEQ ID NO:67, подвергалась обратной трансляции с использованием целевых частот кодонов, приведенных в столбце таблицы 4 для синтетических генов, предназначенных для использования в кукурузе. Полученная в результате последовательность ДНК анализировалась, и кодоны заменялись при

15 необходимости для удаления нежелательных открытых рамок считывания и удаления нежелательных сайтов рестрикции, при этом поддерживались все последовательности, идентифицированные в таблице 1. Аминокислотная последовательность, кодируемая SEQ ID NO:67, сохранялась. Полученная в результате последовательность ДНК дана

20 для сравнения с синтетическим геном, спроектированным по изобретению.

SEQ ID NO:69 была проанализирована, и кодоны были изменены с целью удаления потенциальных последовательностей сигналов полиаденилирования, идентифицированных в таблице 2, при этом поддерживалось количество последовательностей, идентифицированных в таблице 1. Полученная в результате

25 последовательность, которая является вариантом осуществления настоящего изобретения, дана под SEQ ID NO:71. В таблице 1 показано, что количество и положение последовательностей сигналов полиаденилирования, идентифицированных в таблице 1, поддерживается в SEQ ID NO:71. В таблице 27 показано, что количество последовательностей сигналов полиаденилирования, идентифицированных в таблицах

30 2 и 3, снижается в SEQ ID NO:71 по сравнению с SEQ ID NO:67.

Синтезируется ДНК из ID NO:71, и уровни экспрессии, наблюдаемые в растительных клетках, трансформированных для экспрессии данной последовательности, сравниваются с уровнями экспрессии, наблюдаемыми в растительных клетках, трансформированных для экспрессии SEQ ID NO:67 и SEQ ID NO:69.

35 Синтетическая кодирующая область SEQ ID NO:71 была оптимизирована для экспрессии в кукурузе.

Конструкция для применения в экспрессии синтетической кодирующей области SEQ ID NO:71 сделана путем объединения синтетической кодирующей области SEQ ID NO:71 с 5'-нетранслируемой областью, содержащей промотор, функциональный в

40 растительных клетках, и 3'-нетранслируемой области, содержащей терминатор транскрипции и последовательность полиаденилирования.

Таблица 27

Последовательности таблицы 1, обнаруженные в нативной кодирующей области PER1 *Xerophyta viscosa* (SEQ ID NO:67) и ее перестроенной версии (SEQ ID NO:71)

	Последовательность из таблицы 1	Число участков в нативной последовательности XvPER1 (SEQ ID NO:67)	Положение в нативной последовательности XvPER1, н.п. (SEQ ID NO:67)	Число участков в перестроенной последовательности XvPER1 (SEQ ID NO:71)	Положение в перестроенной последовательности XvPER1, н.п. (SEQ ID NO:71)
1	AATAAA	0	NA*	0	NA
2	AATAAT	0	NA	0	NA

3	AACCAA	0	NA	0	NA
4	ATATAA	0	NA	0	NA
5	AATCAA	0	NA	0	NA
6	ATACTA	0	NA	0	NA
7	ATAAAA	1	605	1	605
8	ATGAAA	0	NA	0	NA
9	AAGCAT	0	NA	0	NA
10	ATTAAT	0	NA	0	NA
11	ATACAT	0	NA	0	NA
12	AAAATA	1	282	1	282
13	ATGAAA	0	NA	0	NA
14	AATGAA	0	NA	0	NA
15	AATACA	0	NA	0	NA
16	CATAAA	0	NA	0	NA
	итого	2		2	

*NA = Не применимо

Таблица 28
Последовательности таблицы 2, обнаруженные в нативной кодирующей области PER1 *Xerophyta viscosa* (SEQ ID NO:67) и ее перестроенной версии (SEQ ID NO:71)

	Последовательность из таблицы 2	Число участков в нативной последовательности XvPER1 (SEQ ID NO:67)	Положение в нативной последовательности XvPER1, н.п. (SEQ ID NO:67)	Число участков в перестроенной последовательности XvPER1 (SEQ ID NO:71)	Положение в перестроенной последовательности XvPER1, н.п. (SEQ ID NO:71)
1	ATATAT	0	NA*	0	NA
2	TGTTTT	0	NA	0	NA
3	TTTTGT	0	NA	0	NA
4	TGTTTT	0	NA	0	NA
5	TAATGA	0	NA	0	NA
6	TAATTT	0	NA	0	NA
7	TTTTTT	0	NA	0	NA
8	ATTTTT	0	NA	0	NA
9	TTATTT	0	NA	0	NA
10	TTTATT	0	NA	0	NA
11	TAATAA	0	NA	0	NA
12	ATTTAT	0	NA	0	NA
13	TATATT	0	NA	0	NA
14	TTTTAT	0	NA	0	NA
15	ATATTT	0	NA	0	NA
16	TATTAT	0	NA	0	NA
17	TGTTTG	0	NA	0	NA
18	TTATAT	0	NA	0	NA
19	TGTAAT	0	NA	0	NA
20	AAATAA	0	NA	0	NA
	итого	0		0	

*NA = Не применимо

ПРИМЕР 13

WHISKERS[®]-трансформация кукурузы с применением XvSAP1

Был сконструирован стандартный вектор трансформации WHISKERS, в котором промотор *Arabidopsis thaliana*, Rd29A, был помещен в 5'-направлении от перестроенной последовательности кодирующей области XvSAP1 по изобретению (SEQ ID NO:53).

Данные последовательности были фланкированы PER5 *Zea mays*, 3' и 5'-нетранслируемыми областями, с целью стабилизации экспрессии перестроенной кодирующей области. Выборочная кассета pat (см., например, US 5648477), управляемая промотором актина1 риса, была помещена в 3'-направлении от экспрессионной кассеты

XvSAP1.

Векторная ДНК расщеплялась соответствующими рестрикционными ферментами с целью высвобождения фрагмента, содержащего бактериальный ген устойчивости к ампициллину, присутствующий в остове вектора, и для продуцирования линейного фрагмента ДНК, подходящего для WHISKERS™-опосредованной трансформации. Очистка линейного фрагмента, содержащего XvSAP1 и экспрессионные кассеты rat, достигалась на уровне подготовки посредством высокоэффективной жидкостной хроматографии (HPLC). Данная ДНК для трансформации растения доставлялась в клеточные культуры суспензии Hi-II кукурузы посредством WHISKERS™-опосредованной трансформации (по существу, в соответствии с описанным в патентах США No. 5302523 и 5464765; публикации патента США No. 2008/0182332; и работе Petolino and Arnold (2009), (Methods Molec. Biol. 526:59-67).

Трансформанты были помещены в селективную среду, после чего трансформированные изоляты были получены в течение приблизительно 8 недель. Селективная среда представляла собой основанную на LS среду (минимальная среда LS, витамины N6, 1,5 мг/л 2,4-D, 0,5 г/л MES моногидрат (2-(N-морфолино) этансульфоновой кислоты; PhytoTechnologies Labr.), 30,0 г/л сукрозы, 6 мМ L-пролина, 1,0 мг/л AgNO₃, 250 мг/л цефотаксима, 2,5 г/л геллановой камеди, pH=5,7), содержащую биалафос (Gold BioTechnology). Зародыши были перенесены в селективную среду, содержащую 3 мг/л биалафоса, пока не были получены зародышевые изоляты. Восстановленные изоляты насыщались путем переноса в свежую селективную среду через 2-недельные интервалы для регенерации и дальнейшего анализа.

Для регенерации культуры были перенесены в индукционную среду "28" (MS-соли и витамины, 30 г/л сукрозы, 5 мг/л бензиламинопурина, 0,25 мг/л 2, 4-D, 3 мг/л биалафоса, 250 мг/л цефотаксима, 2,5 г/л геллановой камеди, pH 5,7) на одну неделю в условиях низкой освещенности (14 мкЕм⁻²с⁻¹), затем на 1 неделю в условиях высокой освещенности (приблизительно 89 мкЕм⁻²с⁻¹). После этого ткани были перенесены в регенерационную среду "36" (такая же, как индукционная среда, за исключением отсутствия регуляторов роста растений). После того, как ростки достигали 3-5 см в длину, они переносились в стеклянные пробирки для культур, содержащие среду SHGA (соли и витамины Schenk и Hildebrandt (1972); PhytoTechnologies Labr.), 1,0 г/л мио-инозитола, 10 г/л сукрозы и 2,0 г/л геллановой камеди, pH 5,8) для обеспечения возможности дальнейшего развития побега и корней. Растения пересаживались в такую же почвенную смесь, как была описана выше в настоящем документе, и росли в теплице до цветения. Проводились контролируемые опыления для получения семян.

ПРИМЕР 14

Агробактериальная трансформация

Стандартные способы клонирования использовались при конструировании бинарных плазмид трансформации растений и экспрессии. Рестрикционные эндонуклеазы и ДНК-лигазу T4 получали из NEB. Подготовка плазмид выполнялась с использованием набора для подготовки плазмид NucleoSpin® или набора NucleoBond® AX Xtra Midi (оба поставляются компанией Macherey-Nagel) в соответствии с инструкциями производителей. Фрагменты ДНК очищались с использованием набора для очистки ПЦР QIAquick® или набора для экстракции из геля QIAEX II® (оба поставляются компанией Qiagen) после выделения геля.

Синтетические гены по изобретению могут быть синтезированы коммерческим поставщиком (например, DNA2.0, Menlo Park, CA) и поставлены в форме клонированных

фрагментов в стандартных плазмидных векторах, или могут быть получены посредством стандартных манипуляций молекулярной биологии с другими конструкциями, содержащими соответствующие нуклеотидные последовательности.

В неограничивающем примере базовая стратегия клонирования может состоять в субклонировании полноразмерных кодирующих последовательностей (CDS) в плазмиду для экспрессии в растениях в сайтах рестрикции NcoI и SacI. Получающиеся в результате растительные экспрессионные кассеты, содержащие соответствующую кодирующую область под управлением экспрессионных элементов растений (например, обеспечивающих экспрессию в растениях промоторов, детерминант 3'-концевой терминации транскрипции и добавления полиаденилирования и т.п.), субклонируются в бинарный плазмидный вектор с применением, например, технологии Gateway[®] или стандартных процедур клонирования фрагментов с применением рестрикционных ферментов. Например, LR Clonase[™] (Invitrogen) может быть использована для рекомбинации растительных экспрессионных кассет полноразмерного и модифицированного гена в бинарную плазмиду для трансформации растений, если применяется технология Gateway[®]. Удобно использовать бинарный вектор трансформации растений, который содержит бактериальный ген, придающий устойчивость к антибиотику спектиномицину, когда плазида присутствует в клетках *E. coli* и *Agrobacterium*. Также удобно использовать бинарный плазмидный вектор, который содержит имеющий возможность экспрессии в растениях селективируемый маркерный ген, который является функциональным в желаемых растениях-хозяевах. Примеры имеющих возможность экспрессии в растениях селективируемых маркерных генов включают, но не ограничиваются перечисленным, гены, которые кодируют ген аминокликозидной фосфотрансферазы (*aphII*) транспозона Tn5, который придает устойчивость к антибиотикам канамицину, неомицину и G418, а также гены, который кодируют устойчивость или резистентность к глифосату; гигромицину; метотрексату; фосфинотрицину (биалафосу), имидазолиновым, сульфонилмочевинным и триазолопиримидиновым гербицидам, таким как хлорсульфурон, бромоксинил, далапон и т.п.

Электро-компетентные клетки *Agrobacterium tumefaciens*, штамм Z707S (устойчивая к стрептомицину производная Z707; Hepburn et al., 1985, J. Gen. Microbiol. 131:2961-2969) готовятся и трансформируются с применением электропорации (Weigel and Glazebrook, 2002, *Arabidopsis: A Laboratory Manual*). После электропорации, 1 мл жидкой среды YEP (г/л: дрожжевой экстракт, 10; пептон, 10; NaCl, 5) добавляется в кювету и суспензия клеток и YEP переносится в 15 мл пробирку для культур для инкубирования при 28° на водяной бане при постоянном перемешивании в течение 4 часов. Клетки наносятся на чашки с YEP и агаром (25 г/л) со спектиномицином (200 мкг/мл) и стрептомицином (250 мкг/мл), и чашки инкубируются в течение 2-4 дней при 28°. Хорошо разделенные одиночные колонии отбираются и наносятся на свежие чашки с YEP и агаром со спектиномицином и стрептомицином, как и ранее, и инкубируются при 28° в течение 1-3 дней.

Присутствие синтетической геной вставки в бинарном векторе трансформации растений подтверждается посредством ПЦР-анализа с использованием специфичных для вектора праймеров с шаблонной плазмидной ДНК, приготовленной из выбранных колоний *Agrobacterium*. Клеточная масса из 4 мл аликвоты 15 мл культуры, выросшей в течение ночи в YEP со спектиномицином и стрептомицином, как и ранее, извлекалась с использованием Qiagen Spin[®] Mini Preps, в соответствии с инструкциями производителя. Плазмидная ДНК из бинарного вектора, использованного для электропорационной

трансформации *Agrobacterium*, использовалась в качестве контроля. Реакция ПЦР выполняется с использованием ДНК-полимеразы Taq, поставляемой Invitrogen, в соответствии с инструкциями производителя в концентрациях 0,5X. Реакции ПЦР проводятся в термоциклере MJ Research Peltier, запрограммированном на следующие условия: шаг 1) 94° в течение 3 минут; шаг 2) 94° в течение 45 секунд; шаг 3) 55° в течение 30 секунд; шаг 4) 72° в течение 1 минуты на т.н.п. ожидаемой длины продукта; шаг 5) 29 раз повторяется шаг 2; шаг 6) 72° в течение 10 минут. Реакция поддерживается на 4° после проведения циклов. Продукты амплификации анализируются посредством электрофореза в агарозном геле {например, от 0,7% до 1% агарозы, вес/объем} и визуализируются посредством окрашивания этидия бромидом. Выбирается колония, ПЦР-продукт которой идентичен плазмидному контролю.

Альтернативно, плазмидную структуру бинарного вектора трансформации растений, содержащего вставку синтетического гена, получают путем фингерпринтингового отображения фрагментов рестрикции плазмидной ДНК, приготовленной из кандидатов-изолятов *Agrobacterium* с помощью стандартных методов молекулярной биологии, хорошо известных специалистам в области работы с *Agrobacterium*.

Специалистам в области получения трансформированных растений с помощью способов трансформации, опосредованной *Agrobacterium*, будет понятно, что могут быть с выгодой использованы другие штаммы *Agrobacterium* помимо Z707S, и выбор штамма может зависеть от вида растений-хозяев, подлежащих трансформации.

ПРИМЕР 15

Продуцирование инсектицидных белков в двудольных растениях

Трансформация *Arabidopsis thaliana* Col-01 трансформировался с применением способа погружения цветочных почек в суспензию (Weigel and Glazebrook, см. выше). Выбранная колония *Agrobacterium* использовалась для инокуляции от 1 мл до 15 мл культур жидкой среды YEP, содержащей соответствующие антибиотики для отбора. Культура инкубировалась в течение ночи при 28° с постоянным помешиванием на 220 об/мин. Каждая культура использовалась для инокуляции двух 500 мл культур жидкой среды YEP, содержащих соответствующие антибиотики для отбора, и новые культуры инкубировались в течение ночи при 28° с постоянным помешиванием. Клетки осаждались центрифугированием приблизительно на 8700g в течение 10 минут при комнатной температуре, и получающийся в результате супернатант удалялся. Сгусток клеток мягко ресуспендировался в 500 мл инфльтрационной среды, содержащей: 1/2x солей Murashige и витамины B5 Skoog (Sigma-Aldrich)/Gamborg (Gold BioTechnology, St. Louis, MO), 10% (вес/объем) сахарозы, 0,044 мкМ бензиламинопурина (10 мкл/л исходного раствора 1 мг/мл в DMSO) и 300 мкл/л Silwet L-77. Растения с возрастом приблизительно 1 месяц окунались в среду на 15 секунд, при этом принимались меры по обеспечению погружения новейших соцветий. Затем растения укладывались на бок и накрывались (прозрачным или непрозрачным материалом) на 24 часа, промывались водой и размещались вертикально. Растения выращивались при 22°, со световым периодом 16 часов света/8 часов темноты. Приблизительно после 4 недель после погружения собирались семена.

Рост и отбор *Arabidopsis*. Свежесобранные семена T1 подсушивались в течение, по меньшей мере, 7 дней при комнатной температуре в присутствии влагопоглотителя. Семена суспендировались в 0,1% растворе агар/вода (Sigma-Aldrich) и затем стратифицировались при 4° в течение 2 дней. Для подготовки посадок смесь Sunshine Mix LP5 (Sun Gro Horticulture Inc., Bellevue, WA) в поддонах для проращивания размером 10,5 дюйм × 21 дюйм (T.O. Plastics Inc., Clearwater, MN) покрывалась мелким

вермикулитом, подвергалась подпочвенному орошению питательной смесью Хогланда (Hoagland and Arnon, 1950) до увлажнения, затем допускалось дренирование в течение 24 часов. Стратифицированные семена высевались на вермикулит и покрывались камерами для увлажнения (KORD Products, Bramalea, Ontario, Canada) на 7 дней. Семена проращивались и растения росли в Conviron (модели CMP4030 или CMP3244; Controlled Environments Limited, Winnipeg, Manitoba, Canada) в условиях длинного дня (16 часов света/8 часов темноты) при интенсивности света 120-150 мкмоль/м²с при постоянной температуре (22°) и влажности (40-50%). Растения сначала увлажнялись питательной смесью Хогланда и затем деионизированной водой для сохранения почвы влажной, но не сырой.

Камеры удалялись через 5-6 дней после посева, и растения опрыскивались веществом химического отбора для уничтожения растений, выросших из нетрансформированных семян. Например, если имеющий возможность экспрессии в растениях селективируемый маркерный ген, представленный в бинарном векторе трансформации растений, представляет собой ген *pat* или *bar* (Wehrmann et al., 1996, Nat. Biotech. 14: 1274-1278), то трансформированные растения могут быть отобраны путем распыления 1000X раствора Finale (5,78% гюфосинат аммония, Farnam Companies Inc., Phoenix, AZ). Два следующих опрыскивания выполняются через 5-7 дневные интервалы. Выжившие (продолжающие активно расти) растения идентифицировались через 7-19 дней после последнего опрыскивания и пересаживались в горшки, подготовленные с использованием Sunshine Mix LP5. Пересаженные растения накрывались камерами для увлажнения на 3-4 дня и помещались в Conviron при указанных выше условиях роста.

Специалистам в области трансформации двудольных растений будет понятно, что существуют другие способы отбора трансформированных растений, в которых используются другие имеющие возможность экспрессии в растениях селективируемые маркерные гены (например, гены устойчивости к гербицидам).

Определение биологической активности трансгенных *Arabidopsis* против насекомых. Было продемонстрировано, что трансгенные линии *Arabidopsis*, экспрессирующие белки *Cru*, имеют активность против чувствительных к ним видов насекомых, при анализе методом насаивания в питательной среде. Белки, извлеченные из трансгенных и нетрансгенных линий *Arabidopsis*, количественно оценивались посредством соответствующих способов, и объемы проб настраивались с целью нормализации концентрации белка. Определение биологической активности проводилось в искусственной питательной среде, как описано выше. Нетрансгенный *Arabidopsis* и/или буфер и вода включались в исследования в качестве контрольных проверочных обработок.

ПРИМЕР 16

Агробактериальная трансформация для создания супербинарных векторов

Супербинарную систему *Agrobacterium* удобно использовать для трансформации однодольных растений-хозяев. Способы конструирования и проверки супербинарных векторов подробно изложены и включены в настоящее описание посредством ссылки (руководство пользователя по плазмиде pSB1, версия 3.1, от компании Japan Tobacco, Inc., Токуо, Япония). Стандартные методы молекулярной биологии и микробиологии применяются для создания супербинарных плазмид. Верификация/проверка структуры супербинарной плазмиды выполняется с применением способов, описанных выше для бинарных векторов, и может быть модифицирована в соответствии с предложенным в руководстве пользователя по плазмиде pSB1.

ПРИМЕР 17

Продуцирование инсектицидных белков в однодольных растениях

Опосредованная *Agrobacterium* трансформация кукурузы. Семена скрещивания F₁ High II (Armstrong et al., 1991, Maize Genet. Coop. Newsletter 65:92-93) высаживались в 5-галлонные горшки, содержащие смесь 95% беспочвенной среды Metro-Mix 360 (Sun Gro Horticulture, Bellevue, WA) и 5% глинистой/суглинистой почвы. Растения выращивались в теплице с применением комбинации натриевых ламп высокого давления и металлогалогеновых ламп со световым периодом 16:8 свет:темнота. Для получения незрелых зародышей F₂ для трансформации выполнялись контролируемые клональные опыления. Незрелые зародыши изолировались через 8-10 дней после опыления, когда зародыши имели размер приблизительно от 1,0 до 2,0 мм.

Инфицирование и совместное культивирование. Початки кукурузы подвергались поверхностной стерилизации путем чистки жидким мылом, погружения в 70% этанол на 2 минуты и последующего погружения в 20% промышленный отбеливатель (0,1% гипохлорит натрия) на 30 минут, и затем промывались стерильной водой. Суспензия клеток *Agrobacterium*, содержащих супербинарный вектор, готовилась путем переноса 1-2 поколений бактерий, выросших на твердой среде YEP, содержащей 100 мг/л спектиномицина, 10 мг/л тетрациклина и 250 мг/л стрептомицина при 28° на 2-3 дня в 5 мл жидкой инфекционной среды (базовая среда LS (Linsmaier and Skoog, 1965, Physiol. Plant. 18: 100-127), витамины N6 (Chu et al, 1975, Scientia Sinica 18:659-668), 1,5 мг/л 2,4-дихлорфеноксиуксусная кислота (2,4-D), 68,5 г/л сукрозы, 36,0 г/л глюкозы, 6 мМ L-пролина, рН 5,2), содержащей 100 мкМ ацетосирингона. Раствор перемешивался на вортексе до получения однородной суспензии, и концентрация может быть подобрана для конечной плотности, составляющей около 200 единиц Клетта, с применением колориметра Клетта-Саммерсона с фиолетовым фильтром, или оптической плотностью приблизительно 0,4 на 550 нм. Незрелые эмбрионы выделялись непосредственно в микроцентрифужную пробирку, содержащую 2 мл инфекционной среды. Среда удалялась и заменялась на 1 мл раствора *Agrobacterium* с плотностью, составляющей 200 единиц Клетта, и раствор, содержащий *Agrobacterium* и зародыши, инкубировался в течение 5 минут при комнатной температуре, и затем переносился в среду для совместного культивирования (базовая среда LS, витамины N6, 1,5 мг/л 2,4-D, 30,0 г/л сукрозы, 6 мМ L-пролина, 0,85 мг/л AgNO₃, 100 мкМ ацетосирингона, 3,0 г/л геллановой камеди (PhytoTechnology Laboratories, Lenexa, KS), рН 5,8) на 5 дней при 25° в темноте.

После совместного культивирования зародыши переносились на селективную среду, после чего трансформированные изоляты получали в течение приблизительно 8 недель. Для отбора тканей кукурузы, трансформированных с использованием супербинарной плазмиды, содержащей экспрессируемый в растениях селективируемый маркерный ген *pat* или *bar*, применялась основанная на LS среда (базовая среда LS, витамины N6, 1,5 мг/л 2,4-D, 0,5 г/л MES (моногидрат 2-(N-морфолино)этансульфоновой кислоты; PhytoTechnologies Labr.), 30,0 г/л сукрозы, 6 мМ L-пролина, 1,0 мг/л AgNO₃, 250 мг/л цефотаксима, 2,5 г/л геллановой камеди, рН 5,7) с биалафосом (Gold BioTechnology). Зародыши переносились на селективную среду, содержащую 3 мг/л биалафоса, до тех пор, пока не были получены изоляты зародышей. Восстановленные изоляты объединялись путем переноса на свежую селективную среду в 2-недельные интервалы для регенерации и дальнейшего анализа.

Специалистам в области трансформации кукурузы будет понятно, что существуют другие способы отбора трансформированных растений, в которых применяются другие имеющие возможность экспрессии в растениях селективируемые маркерные гены

(например, гены устойчивости к гербицидам).

Регенерация и получение семян. Для регенерации культуры переносились в индукционную среду "28" (MS-соли и витамины, 30 г/л сукрозы, 5 мг/л бензиламинопурина, 0,25 мг/л 2, 4-D, 3 мг/л биалафоса, 250 мг/л цефотаксима, 2,5 г/л геллановой камеди, рН 5,7) на одну неделю в условиях низкой освещенности (14 мкЕм⁻²с⁻¹), затем на 1 неделю в условиях высокой освещенности (приблизительно 89 мкЕм⁻²с⁻¹). После этого ткани были перенесены в регенерационную среду "36" (такая же, как индукционная среда, за исключением отсутствия регуляторов роста растений). После того, как ростки достигали 3-5 см в длину, они переносились в стеклянные пробирки для культур, содержащие среду SHGA (соли и витамины Schenk и Hildebrandt (1972); PhytoTechnologies Labr.), 1,0 г/л мио-инозитола, 10 г/л сукрозы и 2,0 г/л геллановой камеди, рН 5,8) для обеспечения возможности дальнейшего развития побега и корней. Растения пересаживались в такую же почвенную смесь, как была описана выше в настоящем описании, и росли в теплице до цветения. Проводились контролируемые опыления для получения семян.

Альтернативно, бинарные векторы могут применяться для получения трансгенных растений кукурузы, который содержат один или более химерных генов, стабильно встроенных в геном растения и содержащих кодирующую область, описанную в настоящей заявке. Например, растения, содержащие, по меньшей мере, одну кодирующую область из SEQ ID NO:5, 11, 17, 23, 29, 35, 41, 47, 53, 59, 65 или 71, получают после опосредованной *Agrobacterium* трансформации. Способы трансформации растений, в которых применяются бинарные векторы трансформации, известны в технике. В одном из вариантов осуществления трансформированные ткани отбираются по их способности к росту в содержащей галоксифоп среде, и исследуются на предмет продуцирования белков при необходимости.

Стерилизация початков и выделение зародышей. Незрелые зародыши кукурузы были получены из растений *Zea mays* инбредной линии В104, выращенных в теплице, и самоопыленных или клонально опыленных для получения початков. Початки собирались приблизительно через 9-12 дней после опыления. В день эксперимента очищенные от оберток початки подвергались поверхностной стерилизации посредством погружения в 20% раствор гипохлорита натрия (6,15%) и встряхивания в течение от 20 до 30 минут, после чего проводилось трехкратное ополаскивание в стерильной воде. После стерилизации незрелые зиготные зародыши (от 1,5 до 2,4 мм) стерильно вырезались из каждого початка и случайным образом распределялись по микроцентрифужным пробиркам, содержащим жидкую среду для инокуляции. Среда для инокуляции содержала: 2,2 г/л MS-солей (Frame et al, 2011, Genetic Transformation Using Maize Immature Zygotic Embryos. IN: Plant Embryo Culture Methods and Protocols: Methods in Molecular Biology. T. A. Thorpe and E. C. Yeung, (Eds), SPRINGER SCIENCE AND BUSINESS MEDIA, LLC. pp 327-341); 1X ISU модифицированные MS-витамины (Frame et al, 2011 см. выше); 68,4 г/л сукрозы; 36 г/л глюкозы; 115 мг/л L-пролина; 100 мг/л мио-инозитола; и 200 мкМ ацетосирингона (приготовленного в DMSO); при рН 5,4. В указанном множестве экспериментов зародыши из объединенных в группу початков использовались для каждой трансформации.

Стимулирование культур *Agrobacterium*. Содержащийся в глицерине штамм *Agrobacterium* DA13192 (международная РСТ-публикация No. WO2012016222(A2)), содержащий бинарный вектор трансформации pDAB111440 (пример 1) наносился на чашки с минимальной средой АВ (Watson, et al, (1975) J. Bacteriol. 123:255-264), содержащей соответствующие антибиотики, и выращивался при 20°C в течение от 3 до

4 дней. Единичная колония собиралась и наносилась на чашки с YEP (г/л: дрожжевой экстракт, 10; пептон, 10; NaCl 5), содержащие те же самые антибиотики, и инкубировалась при 20°C в течение 1-2 дней.

5 Культуры *Agrobacterium* и совместное культивирование. Культуры *Agrobacterium* брались с чашек YEP, суспендировались в 10 мл среды для инокуляции в 50 мл одноразовой пробирке, и плотность клеток подбиралась для значения OD₅₅₀ от 0,2 до 0,4 (оптическая плотность, измеренная при 550 нм; мера роста клеток) с применением спектрофотометра. Культуры *Agrobacterium* инкубировались на ротационном шейкере при 125 об/мин (комнатная температура) во время выполнения разрезания зародышей.
10 Незрелые зиготные зародыши (ранее выделенные из стерилизованных зерен кукурузы и помещенные в 1 мл среды для инокуляции) однократно промывались в той же среде. Два мл суспензии *Agrobacterium* было добавлено к каждой пробирке с зародышами, и пробирки были помещены на платформу шейкера на время от 10 до 15 минут. Зародыши были перенесены в среду для совместного культивирования, ориентированы так, чтобы
15 щиток зародыша смотрел вверх, и инкубировались при 25°C, при 24-часовом свете с интенсивностью 50 мкЕм⁻²с⁻¹ в течение 3 дней. Среда для совместного культивирования содержала 4,33 г/л MS-солей; 1X ISU модифицированные MS-витамины; 30 г/л сукрозы; 700 мг/л L-пролина; 3,3 мг/л дикамбы в KOH (3,6-дихлор-о-анисовая кислота или 3,6-дихлор-2-метоксибензойная кислота); 100 мг/л мио-инозитола; 100 мг/л ферментного
20 гидролизата казеина; 15 мг/л AgNO₃; 100 мкМ ацетосирингона в DMSO; и 3 г/л GELZAN™ (SIGMA-ALDRICH); при pH 5,8.

Отбор каллуса и регенерация для предполагаемых случаев. После периода совместного культивирования зародыши переносились на среду покоя и инкубировались
25 при 24-часовом свете с интенсивностью 50 мкЕм⁻²с⁻¹ в течение 3 дней. Среда покоя содержала 4,33 г/л MS-солей; 1X ISU модифицированные MS-витамины; 30 г/л сукрозы; 700 мг/л L-пролина; 3,3 мг/л дикамбы в KOH; 100 мг/л мио-инозитола; 100 мг/л ферментного гидролизата казеина; 15 мг/л AgNO₃; 0,5 г/л MES - моногидрата (2-(N-морфолино)этансульфоновой кислоты; PHYTOTECNOLOGIES LABR.; Lenexa, KS);
30 250 мг/л карбенициллина; и 2,3 г/л GELZAN™; при pH 5,8. Зародыши переносились в селективную среду 1 (которая состояла из среды покоя (см. выше) с 100 нМ R-галоксифоповой кислоты (0,0362 мг/л)), и инкубировались в темноте и/или при 24-часовом свете с интенсивностью 50 мкЕм⁻²с⁻¹ в течение от 7 до 14 дней при 28°C.
35 Размножающиеся зародышевые каллусы переносились на селективную среду 2 (которая состояла из среды покоя (см. выше) с 500 нМ R-галоксифоповой кислоты (0,1810 мг/л)), и инкубировались при 24-часовом свете с интенсивностью 50 мкЕм⁻²с⁻¹ в течение от 14 до 21 дней при 28°C. Данный этап отбора позволял осуществляться дальнейшему размножению и дифференциации трансгенного каллуса.

40 Размножающиеся зародышевые каллусы переносились на пререгенерационную среду и культивировались при 24-часовом свете с интенсивностью 50 мкЕм⁻²с⁻¹ в течение 7 дней при 28°C. Пререгенерационная среда содержала 4,33 г/л MS-солей; 1X ISU модифицированные MS-витамины; 45 г/л сукрозы; 350 мг/л L-пролина; 100 мг/л мио-инозитола; 50 мг/л ферментного гидролизата казеина; 1,0 мг/л AgNO₃; 0,25 г/л MES;
45 0,5 мг/л нафталинуксусной кислоты в NaOH; 2,5 мг/л абсцизовой кислоты в этаноле; 1 мг/л 6-бензиламинопурина; 250 мг/л карбенициллина; 2,5 г/л GELZAN™; и 500 нМ R-галоксифоповой кислоты; при pH 5,8. Зародышевые каллусы с похожими на побеги почками переносились на регенерационную среду и культивировались при 24-часовом

5 свете с интенсивностью $50 \text{ мкЕм}^{-2}\text{с}^{-1}$ в течение 7 дней. Регенерационная среда I содержала 4,33 г/л MS-солей; 1X ISU модифицированные MS-витамины; 60 г/л сукрозы; 100 мг/л мио-инозитола; 125 мг/л карбенициллина; 3,0 г/л GELZAN™; и 500 нМ R-галоксифоповой кислоты; при pH 5,8. Небольшие побеги с первичными корнями переносились в среду побегов/корней в PHYTATRAYS (PHYTOTECNOLOGIES LABR.; Lenexa, KS) и инкубировались в условиях 16:8 ч свет:темнота при интенсивности света от 140 до 190 $\text{мкЕм}^{-2}\text{с}^{-1}$ в течение 7 дней при 27°C. Среда побегов/корней содержала 4,33 г/л MS-солей; 1X ISU модифицированные MS-витамины; 30 г/л сукрозы; 100 мг/л мио-инозитола; 3,5 г/л GELZAN™; при pH 5,8. Предполагаемые трансгенные проростки анализировались относительно числа копий трансгена посредством количественной ПЦР в реальном времени или других стандартных методик молекулярного анализа, и переносились в почву.

Перенос и укоренение растений T_0 в теплице для получения семян.

15 Трансформированные растительные ткани, отобранные по их способности к росту в среде, содержащей 500 нМ R-галоксифоповой кислоты, пересаживались в беспочвенную среду для роста METRO-MIX 360 (SUN GRO HORTICULTURE) и закачивались в фитотроне. Растения пересаживались в почвенную смесь SUNSHINE CUSTOM BLEND 160 и выращивались в теплице до цветения. Проводились контролируемыми опыления для получения семян.

20 Образцы тканей листьев отобранных растений T_0 собирались на стадиях от V-3 до V-5. Два образца листьев диаметром 6 мм хранилось в 96-луночной подставке для кассетных пробирок при -80°C до дня анализа. Два стальных буферных основания DAISY™ и 200 мкл экстракционного буфера (раствор PBS, содержащий 0,05% Tween 20 и 5 мкл/мл смеси ингибитора протеазы SIGMA (каталожный номер 9599)) добавлялись к каждой пробирке. Образцы перемалывались в шаровой мельнице KLECO (Visalia, CA) в течение 3 минут на максимальной скорости. Образцы центрифугировались на 3000g в течение 5 минут, затем 100 мкл супернатанта переносилось в пустую пробирку для образцов. Еще 100 мкл экстракционного буфера добавлялось к растительному образцу, и осуществлялось перемалывание на шаровой мельнице в течение 30 дополнительных 3 минут. После повторного центрифугирования 100 мкл данного экстракта объединялось с первыми 100 мкл. Объединенные супернатанты смешивались и анализировались в день экстракции.

35 Белки, экстрагированные из измеренных областей ткани листа, анализировались на предмет экспрессии белка Cry1Fa и белка AAD-1 посредством стандартного ELISA (твердофазный иммуноферментный анализ) или иммуноблоттинга белков (вестерн-блоттинг). Для обнаружения белка Cry1Fa посредством ELISA использовались реагенты из набора ENVIROLOGIX для ELISA (каталожный номер AP 016 NW V10; Portland, ME) в соответствии с инструкциями производителя. Обнаружение AAD-1 выполнялось 40 посредством стандартных методик ELISA (например, в соответствии с изложенным в Ausubel et al. (1995 и обновления), Current Protocols in Molecular Biology, John Wiley and Sons, New York) с использованием антител кролика, приготовленных против очищенного белка AAD-1.

45 Результаты ELISA, полученные для экстрактов трансформированных pDAB111440 растений, изложены в таблице 29. Уровни белков выражены в нг искомого белка, обнаруженного на 1 квадратный сантиметр собранной области листа.

Таблица 29

Уровни экспрессии белков Cry1Fa и AAD-1, экстрагированных из растений кукурузы, трансформированных с помощью плазмиды

pDAB111440, в соответствии с определенным посредством методов ELISA		
Идентификатор образца	Cry1Fa, нг/см ³	AAD-1, нг/см ³
111440[3]-001.001	2,30	14,0
111440[3]-015.001	3,80	0,0
111440[3]-023.001	3,80	320,0
111440[3]-020.001	5,40	190,0
111440[3]-011.001	17,00	0,0

Экстракты белков пяти трансформированных с pDAB111440 растений, приведенные в таблице 29 (а также экстракт из нетрансформированного растения, использованного в качестве отрицательного контроля), были приготовлены в соответствии с описанным выше и зондировались с помощью антитела к Cry1Fa на иммуноблотах (вестерн-блоты). Процедуры иммуноблоттинга, по существу, соответствовали описанному в работе Gallagher et al. (2008; Immunoblotting and Immunodetection. Current Protocols in Immunology 8.10.1 - 8.10.28). Образы белка (80 мкл) смешивались с 20 мкл буфера для образцов INVITROGEN NuPAGE LDS, нагревались при >90°C в течение пяти минут, загружались в 4-12% бис-трис гель INVITROGEN NuPAGE и запускались в подвижном буфере MOPS SDS (200 вольт в течение 45 минут). Двухцветные стандарты BIORAD PRECISION PLUS загружались на отдельную дорожку. Белки переносились на 0,2 мкМ нитроцеллюлозную мембрану посредством системы переноса геля INVITROGEN iBLOT в соответствии с инструкциями производителя. Мембрана блокировалась посредством смеси для блокирования INVITROGEN WESTERN BREEZE, затем осуществлялась реакция с первичным антителом (очищенное антитело кролика против Cry1F No. D0609RA07-A0; Strategic Diagnostics Inc., Newark, DE), и затем со вторым антителом (биотинилированное козье антитело против кролика, производства INVITROGEN). За этим следовал конъюгат HRP-стрептавидин производства INVITROGEN, и реагирующие полоски определялись посредством PIERCE SUPERSIGNAL WEST PICO LUMINOL ENHANCER AND STABLE PEROXIDE (No. 34080).

Дорожки с положительным контролем содержали 0,5 нг или 1,0 нг очищенного белка ядерного токсина Cry1Fa, продуцированного посредством экспрессии полноразмерной кодирующей области Cry1Fa в экспрессионной системе *Pseudomonas fluorescens* (см., например, заявку на патент США No.20100269223A1). Полноразмерный белок Cry1Fa обрабатывался трипсином с целью высвобождения сегмента ядерного токсина Cry1Fa с рассчитанным молекулярным размером 68 кДа, который использовался в качестве стандарта положительного контроля на иммуноблоте. Не было обнаружено полос реагирования с антителом в экстракте из растения отрицательного контроля, тогда как во всех пяти экстрактах из трансгенных растений содержалась единственная преобладающая полоса (приблизительно соответствующая по интенсивности контрольным белкам Cry1Fa) с оцененным размером несколько более 75 кДа.

Способы борьбы с насекомыми-вредителями. Когда насекомое взаимодействует с эффективным количеством токсина, доставляемым посредством экспрессии в трансгенном растении, результатом обычно является смерть насекомого, или насекомые не питаются в источнике, в котором токсины доступны насекомым.

(57) Формула изобретения

1. Синтетическая ДНК для экспрессии белковых токсинов Cry в клетках кукурузы, которая содержит:

- оптимизированную по кодонам последовательность ДНК, кодирующую белковые токсины Cry,
- по меньшей мере одну последовательность сигнала полиаденилирования,

выбранную из группы, состоящей из класса I и класса II, где

класс I выбирают из группы, состоящей из ААТААА, ААТААТ, ААССАА, АТАТАА, ААТСАА, АТАСТА, АТАААА, АТГААА, ААГСАТ, АТТААТ, АТАСАТ, ААААТА, АТТААА, ААТТАА, ААТАСА и САТААА; и

5 класс II выбирают из группы, состоящей из АТАТАТ, ТТГТТТ, ТТТТГТ, ТГТТТТ, ТАТАТА, ТАТТТТ, ТТТТТТ, АТТТТТ, ТТАТТТ, ТТТАТТ, ТААТАА, АТТТАТ, ТАТАТТ, ТТТТАТ, АТАТТТ, ТАТТАТ, ТГТТТГ, ТТАТАТ, ТГТААТ и АААТАА; и

где указанная оптимизированная по кодонам последовательность ДНК содержит по меньшей мере одну последовательность сигнала полиаденилирования из класса II, и где указанная синтетическая последовательность ДНК содержит меньше последовательностей сигналов полиаденилирования класса II, чем нативная последовательность ДНК белка, и содержит такое же количество последовательностей сигналов полиаденилирования класса I по сравнению с указанной нативной последовательностью ДНК, где синтетическую ДНК выбирают из группы, состоящей
10 из SEQ ID NO: 5, SEQ ID NO: 11, SEQ ID NO: 17, SEQ ID NO: 23, SEQ ID NO: 29, SEQ ID NO: 35.

2. ДНК-конструкция для экспрессии белковых токсинов Cry, содержащая 5'-нетранслируемую последовательность, последовательность, кодирующую белковые токсины Cry, и 3'-нетранслируемую область, где указанная 5'-нетранслируемая последовательность содержит промотор, функциональный в растительной клетке, указанная кодирующая последовательность представляет собой синтетическую ДНК по п. 1 и указанная 3'-нетранслируемая последовательность содержит последовательность терминации транскрипции и сигнал полиаденилирования, где синтетическую ДНК выбирают из группы, состоящей из SEQ ID NO: 5, SEQ ID NO: 11, SEQ ID NO: 17, SEQ ID NO: 23, SEQ ID NO: 29, SEQ ID NO: 35.
25

3. Трансгенное растение, имеющее устойчивость к насекомым-вредителям, чувствительным к белковым токсинам Cry, содержащее синтетическую ДНК по п. 1, где синтетическую ДНК выбирают из группы, состоящей из SEQ ID NO: 5, SEQ ID NO: 11, SEQ ID NO: 17, SEQ ID NO: 23, SEQ ID NO: 29, SEQ ID NO: 35.
30

4. Способ борьбы с насекомыми-вредителями зерна или семян, чувствительными к белковым токсинам Cry, который включает выращивание растений, содержащих синтетическую ДНК по п. 1, с возможностью употребления указанными насекомыми-вредителями по крайней мере части растения и получение указанных зерен или семян из указанных растений, где синтетическую ДНК выбирают из группы, состоящей из SEQ ID NO: 5, SEQ ID NO: 11, SEQ ID NO: 17, SEQ ID NO: 23, SEQ ID NO: 29, SEQ ID NO: 35.
35

40

45

СПИСОК ПОСЛЕДОВАТЕЛЬНОСТЕЙ

<110> Dow AgroSciences LLC

<120> СИНТЕТИЧЕСКИЕ ГЕНЫ

<130> DAS-P0273-01-WO

<160> 72

<170> PatentIn версия 3.5

<210> 1

<211> 1818

<212> ДНК

<213> *Bacillus thuringiensis*

<220>

<221> CDS

<222> (1)..(1818)

<223> Нативная последовательность ДНК, кодирующая ядерный токсин Cry1Fa
Bacillus thuringiensis

<400> 1

atg gag aat aat att caa aat caa tgc gta cct tac aat tgt tta aat	48
Met Glu Asn Asn Ile Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn	
1 5 10 15	
aat cct gaa gta gaa ata tta aat gaa gaa aga agt act ggc aga tta	96
Asn Pro Glu Val Glu Ile Leu Asn Glu Glu Arg Ser Thr Gly Arg Leu	
20 25 30	
ccg tta gat ata tcc tta tcg ctt aca cgt ttc ctt ttg agt gaa ttt	144
Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe	
35 40 45	
gtt cca ggt gtg gga gtt gcg ttt gga tta ttt gat tta ata tgg ggt	192
Val Pro Gly Val Gly Val Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly	
50 55 60	
ttt ata act cct tct gat tgg agc tta ttt ctt tta cag att gaa caa	240
Phe Ile Thr Pro Ser Asp Trp Ser Leu Phe Leu Leu Gln Ile Glu Gln	
65 70 75 80	
ttg att gag caa aga ata gaa aca ttg gaa agg aac cgg gca att act	288
Leu Ile Glu Gln Arg Ile Glu Thr Leu Glu Arg Asn Arg Ala Ile Thr	
85 90 95	
aca tta cga ggg tta gca gat agc tat gaa att tat att gaa gca cta	336
Thr Leu Arg Gly Leu Ala Asp Ser Tyr Glu Ile Tyr Ile Glu Ala Leu	
100 105 110	
aga gag tgg gaa gca aat cct aat aat gca caa tta agg gaa gat gtg	384
Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp Val	
115 120 125	
cgt att cga ttt gct aat aca gac gac gct tta ata aca gca ata aat	432
Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala Ile Asn	
130 135 140	
aat ttt aca ctt aca agt ttt gaa atc cct ctt tta tcg gtc tat gtt	480
Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val Tyr Val	
145 150 155 160	

caa gcg gcg aat tta cat tta tca cta tta aga gac gct gta tcg ttt	528
Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Ala Val Ser Phe	
165 170 175	
ggg cag ggt tgg gga ctg gat ata gct act gtt aat aat cat tat aat	576
Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His Tyr Asn	
180 185 190	
aga tta ata aat ctt att cat aga tat acg aaa cat tgt ttg gac aca	624
Arg Leu Ile Asn Leu Ile His Arg Tyr Thr Lys His Cys Leu Asp Thr	
195 200 205	
tac aat caa gga tta gaa aac tta aga ggt act aat act cga caa tgg	672
Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg Gln Trp	
210 215 220	
gca aga ttc aat cag ttt agg aga gat tta aca ctt act gta tta gat	720
Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val Leu Asp	
225 230 235 240	
atc gtt gct ctt ttt ccg aac tac gat gtt aga aca tat cca att caa	768
Ile Val Ala Leu Phe Pro Asn Tyr Asp Val Arg Thr Tyr Pro Ile Gln	
245 250 255	
acg tca tcc caa tta aca agg gaa att tat aca agt tca gta att gag	816
Thr Ser Ser Gln Leu Thr Arg Glu Ile Tyr Thr Ser Ser Val Ile Glu	
260 265 270	
gat tct cca gtt tct gct aat ata cct aat ggt ttt aat agg gcg gaa	864
Asp Ser Pro Val Ser Ala Asn Ile Pro Asn Gly Phe Asn Arg Ala Glu	
275 280 285	
ttt gga gtt aga ccg ccc cat ctt atg gac ttt atg aat tct ttg ttt	912
Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser Leu Phe	
290 295 300	
gta act gca gag act gtt aga agt caa act gtg tgg gga gga cac tta	960
Val Thr Ala Glu Thr Val Arg Ser Gln Thr Val Trp Gly Gly His Leu	
305 310 315 320	
gtt agt tca cga aat acg gct ggt aac cgt ata aat ttc cct agt tac	1008
Val Ser Ser Arg Asn Thr Ala Gly Asn Arg Ile Asn Phe Pro Ser Tyr	
325 330 335	
ggg gtc ttc aat cct ggt ggc gcc att tgg att gca gat gag gat cca	1056
Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu Asp Pro	
340 345 350	
cgt cct ttt tat cgg aca tta tca gat cct gtt ttt gtc cga gga gga	1104
Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg Gly Gly	
355 360 365	
ttt ggg aat cct cat tat gta ctg ggg ctt agg gga gta gca ttt caa	1152
Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala Phe Gln	
370 375 380	
caa act ggt acg aac cac acc cga aca ttt aga aat agt ggg acc ata	1200
Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile	
385 390 395 400	
gat tct cta gat gaa atc cca cct cag gat aat agt ggg gca cct tgg	1248
Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp	
405 410 415	

aat gat tat agt cat gta tta aat cat gtt aca ttt gta cga tgg cca 1296
 Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp Pro
 420 425 430

ggt gag att tca gga agt gat tca tgg aga gct cca atg ttt tct tgg 1344
 Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser Trp
 435 440 445

acg cac cgt agt gca acc cct aca aat aca att gat ccg gag agg att 1392
 Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Glu Arg Ile
 450 455 460

act caa ata cca ttg gta aaa gca cat aca ctt cag tca ggt act act 1440
 Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly Thr Thr
 465 470 475 480

ggt gta aga ggg ccc ggg ttt acg gga gga gat att ctt cga cga aca 1488
 Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr
 485 490 495

agt gga gga cca ttt gct tat act att gtt aat ata aat ggg caa tta 1536
 Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val Asn Ile Asn Gly Gln Leu
 500 505 510

ccc caa agg tat cgt gca aga ata cgc tat gcc tct act aca aat cta 1584
 Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu
 515 520 525

aga att tac gta acg gtt gca ggt gaa cgg att ttt gct ggt caa ttt 1632
 Arg Ile Tyr Val Thr Val Ala Gly Glu Arg Ile Phe Ala Gly Gln Phe
 530 535 540

aac aaa aca atg gat acc ggt gac cca tta aca ttc caa tct ttt agt 1680
 Asn Lys Thr Met Asp Thr Gly Asp Pro Leu Thr Phe Gln Ser Phe Ser
 545 550 555 560

tac gca act att aat aca gct ttt aca ttc cca atg agc cag agt agt 1728
 Tyr Ala Thr Ile Asn Thr Ala Phe Thr Phe Pro Met Ser Gln Ser Ser
 565 570 575

ttc aca gta ggt gct gat act ttt agt tca ggg aat gaa gtt tat ata 1776
 Phe Thr Val Gly Ala Asp Thr Phe Ser Ser Gly Asn Glu Val Tyr Ile
 580 585 590

gac aga ttt gaa ttg att cca gtt act gca aca ttt gaa tag 1818
 Asp Arg Phe Glu Leu Ile Pro Val Thr Ala Thr Phe Glu
 595 600 605

<210> 2
 <211> 605
 <212> PRT
 <213> Bacillus thuringiensis

<400> 2

Met Glu Asn Asn Ile Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
 1 5 10 15

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

Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe
 35 40 45
 Val Pro Gly Val Gly Val Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly
 50 55 60
 Phe Ile Thr Pro Ser Asp Trp Ser Leu Phe Leu Leu Gln Ile Glu Gln
 65 70 75 80
 Leu Ile Glu Gln Arg Ile Glu Thr Leu Glu Arg Asn Arg Ala Ile Thr
 85 90 95
 Thr Leu Arg Gly Leu Ala Asp Ser Tyr Glu Ile Tyr Ile Glu Ala Leu
 100 105 110
 Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp Val
 115 120 125
 Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala Ile Asn
 130 135 140
 Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val Tyr Val
 145 150 155 160
 Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Ala Val Ser Phe
 165 170 175
 Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His Tyr Asn
 180 185 190
 Arg Leu Ile Asn Leu Ile His Arg Tyr Thr Lys His Cys Leu Asp Thr
 195 200 205
 Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg Gln Trp
 210 215 220
 Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val Leu Asp
 225 230 235 240
 Ile Val Ala Leu Phe Pro Asn Tyr Asp Val Arg Thr Tyr Pro Ile Gln
 245 250 255
 Thr Ser Ser Gln Leu Thr Arg Glu Ile Tyr Thr Ser Ser Val Ile Glu
 260 265 270
 Asp Ser Pro Val Ser Ala Asn Ile Pro Asn Gly Phe Asn Arg Ala Glu
 275 280 285

Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser Leu Phe
 290 295 300

Val Thr Ala Glu Thr Val Arg Ser Gln Thr Val Trp Gly Gly His Leu
 305 310 315 320

Val Ser Ser Arg Asn Thr Ala Gly Asn Arg Ile Asn Phe Pro Ser Tyr
 325 330 335

Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu Asp Pro
 340 345 350

Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg Gly Gly
 355 360 365

Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala Phe Gln
 370 375 380

Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile
 385 390 395 400

Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp
 405 410 415

Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp Pro
 420 425 430

Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser Trp
 435 440 445

Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Glu Arg Ile
 450 455 460

Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly Thr Thr
 465 470 475 480

Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr
 485 490 495

Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val Asn Ile Asn Gly Gln Leu
 500 505 510

Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu
 515 520 525

Arg Ile Tyr Val Thr Val Ala Gly Glu Arg Ile Phe Ala Gly Gln Phe
 530 535 540

Asn Lys Thr Met Asp Thr Gly Asp Pro Leu Thr Phe Gln Ser Phe Ser
545 550 555 560

Tyr Ala Thr Ile Asn Thr Ala Phe Thr Phe Pro Met Ser Gln Ser Ser
565 570 575

Phe Thr Val Gly Ala Asp Thr Phe Ser Ser Gly Asn Glu Val Tyr Ile
580 585 590

Asp Arg Phe Glu Leu Ile Pro Val Thr Ala Thr Phe Glu
595 600 605

<210> 3

<211> 1818

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Синтетическая последовательность ДНК, кодирующая ядерный токсин Cry1Fa *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1

<220>

<221> CDS

<222> (1)..(1818)

<400> 3

atg gag aat aat atc cag aat caa tgc gtg cct tac aat tgt tta aat 48

Met Glu Asn Asn Ile Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
1 5 10 15

aat ccc gag gtg gag ata tta aac gag gag aga tcc act ggc aga ctg 96

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

cca ctc gat ata tcc ttg tcc ctt acc cgt ttc ctt ttg agc gaa ttt 144

Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe
35 40 45

gtt cct ggt gtg gga gtg gct ttc gga tta ttt gat ctg ata tgg ggt 192

Val Pro Gly Val Gly Val Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly
50 55 60

ttt atc act cct tct gat tgg agc tta ttt ctt ctc cag att gag caa 240

Phe Ile Thr Pro Ser Asp Trp Ser Leu Phe Leu Leu Gln Ile Glu Gln
65 70 75 80

ttg att gag cag aga ata gaa acc ttg gaa agg aac cgt gca atc acg 288

Leu Ile Glu Gln Arg Ile Glu Thr Leu Glu Arg Asn Arg Ala Ile Thr
85 90 95

acc ttg cgc ggt ctc gcc gat agc tat gaa att tat att gaa gca ctg 336

Thr Leu Arg Gly Leu Ala Asp Ser Tyr Glu Ile Tyr Ile Glu Ala Leu
100 105 110

agg gag tgg gag gcc aac cct aat aat gct caa tta agg gaa gat gtg 384

Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp Val
115 120 125

cg	att	cg	ttt	gct	aat	aca	gac	gac	gct	tta	ata	aca	gca	ata	aat	432
Arg	Ile	Arg	Phe	Ala	Asn	Thr	Asp	Asp	Ala	Leu	Ile	Thr	Ala	Ile	Asn	
	130						135				140					
aat	ttc	aca	ctt	aca	tcc	ttt	gaa	atc	ccg	ctt	tta	tca	gtg	tac	gtt	480
Asn	Phe	Thr	Leu	Thr	Ser	Phe	Glu	Ile	Pro	Leu	Leu	Ser	Val	Tyr	Val	
	145				150				155						160	
caa	gcc	gcc	aat	ctc	cat	tta	tca	ctt	ctg	agg	gac	gct	gtc	tcc	ttt	528
Gln	Ala	Ala	Asn	Leu	His	Leu	Ser	Leu	Leu	Arg	Asp	Ala	Val	Ser	Phe	
			165						170					175		
ggg	caa	ggt	tgg	gga	ctg	gat	atc	gct	act	gtt	aat	aat	cac	tac	aat	576
Gly	Gln	Gly	Trp	Gly	Leu	Asp	Ile	Ala	Thr	Val	Asn	Asn	His	Tyr	Asn	
			180				185							190		
aga	tta	ata	aac	ctg	att	cat	aga	tat	acg	aag	cat	tgt	ttg	gac	aca	624
Arg	Leu	Ile	Asn	Leu	Ile	His	Arg	Tyr	Thr	Lys	His	Cys	Leu	Asp	Thr	
	195						200					205				
tac	aat	caa	gga	ctg	gag	aac	ctt	agg	gga	act	aac	act	agg	cag	tgg	672
Tyr	Asn	Gln	Gly	Leu	Glu	Asn	Leu	Arg	Gly	Thr	Asn	Thr	Arg	Gln	Trp	
	210					215					220					
gca	agg	ttc	aac	cag	ttc	aga	cg	gat	ctc	aca	ctt	act	gtg	ctg	gat	720
Ala	Arg	Phe	Asn	Gln	Phe	Arg	Arg	Asp	Leu	Thr	Leu	Thr	Val	Leu	Asp	
	225			230					235					240		
atc	g	g	ctc	ttt	ccg	aac	tac	gat	g	cg	acc	tac	cca	atc	cag	768
Ile	Val	Ala	Leu	Phe	Pro	Asn	Tyr	Asp	Val	Arg	Thr	Tyr	Pro	Ile	Gln	
			245					250						255		
acg	tca	tcc	caa	tta	aca	agg	gaa	att	tat	acc	tcc	tca	gtg	att	gag	816
Thr	Ser	Ser	Gln	Leu	Thr	Arg	Glu	Ile	Tyr	Thr	Ser	Ser	Val	Ile	Glu	
			260					265					270			
gac	tct	ccc	g	tct	gct	aac	ata	cct	aac	ggc	ttc	aac	cg	gcc	gag	864
Asp	Ser	Pro	Val	Ser	Ala	Asn	Ile	Pro	Asn	Gly	Phe	Asn	Arg	Ala	Glu	
		275					280					285				
ttc	gga	g	aga	ccg	ccc	cac	ctt	atg	gac	ttt	atg	aat	agc	ttg	ttt	912
Phe	Gly	Val	Arg	Pro	Pro	His	Leu	Met	Asp	Phe	Met	Asn	Ser	Leu	Phe	
	290					295					300					
gtg	act	gct	gag	act	g	aga	agc	caa	act	gtg	tgg	ggc	ggc	cac	ttg	960
Val	Thr	Ala	Glu	Thr	Val	Arg	Ser	Gln	Thr	Val	Trp	Gly	Gly	His	Leu	
	305				310					315					320	
gtc	agc	tca	cg	aac	acg	gct	ggc	aac	cg	atc	aac	ttc	ccg	tct	tac	1008
Val	Ser	Ser	Arg	Asn	Thr	Ala	Gly	Asn	Arg	Ile	Asn	Phe	Pro	Ser	Tyr	
			325					330					335			
ggg	gtc	ttt	aac	cct	ggt	ggc	gcc	att	tgg	att	gca	gac	gag	gac	cca	1056
Gly	Val	Phe	Asn	Pro	Gly	Gly	Ala	Ile	Trp	Ile	Ala	Asp	Glu	Asp	Pro	
			340				345						350			
cg	cct	ttt	tat	cg	acc	ctg	tca	gat	cct	g	ttt	gtc	aga	ggc	gga	1104
Arg	Pro	Phe	Tyr	Arg	Thr	Leu	Ser	Asp	Pro	Val	Phe	Val	Arg	Gly	Gly	
		355				360						365				
ttt	ggg	aat	cct	cat	tat	gtc	ctg	ggc	ctt	agg	gga	gtg	gct	ttc	caa	1152
Phe	Gly	Asn	Pro	His	Tyr	Val	Leu	Gly	Leu	Arg	Gly	Val	Ala	Phe	Gln	
	370					375					380					

cag act ggc acc aac cac acc cgt acg ttt cgc aat agc ggc acc ata 1200
 Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile
 385 390 395 400

gat tct ctt gat gaa atc cca cct caa gat aac agc ggc gca cct tgg 1248
 Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp
 405 410 415

aac gat tat tcc cac gta tta aat cac gtt acg ttc gtc cgc tgg ccg 1296
 Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp Pro
 420 425 430

ggg gag atc agc ggc agc gat tca tgg aga gca cca atg ttt tct tgg 1344
 Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser Trp
 435 440 445

acg cac cgt tca gcc acc cct aca aat aca att gac ccg gag agg att 1392
 Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Glu Arg Ile
 450 455 460

act caa atc cca ttg gtc aaa gca cat aca ctt cag tct ggg acc acc 1440
 Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly Thr Thr
 465 470 475 480

gtg gtc aga ggg cct ggg ttc acg gga gga gac att ctt agg cgc aca 1488
 Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr
 485 490 495

tcc gga gga ccc ttc gct tat act atc gtt aat ata aat ggg cag ctc 1536
 Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val Asn Ile Asn Gly Gln Leu
 500 505 510

ccc cag cgc tat cgt gcc aga atc cgt tac gcc tct act aca aat ctc 1584
 Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu
 515 520 525

aga atc tac gtg acg gtt gcc ggt gag cgc att ttt gct ggt cag ttt 1632
 Arg Ile Tyr Val Thr Val Ala Gly Glu Arg Ile Phe Ala Gly Gln Phe
 530 535 540

aac aag acg atg gat act ggc gac cca ctg aca ttc caa tct ttc tca 1680
 Asn Lys Thr Met Asp Thr Gly Asp Pro Leu Thr Phe Gln Ser Phe Ser
 545 550 555 560

tac gca act att aat aca gct ttc aca ttc cca atg agc cag tca tct 1728
 Tyr Ala Thr Ile Asn Thr Ala Phe Thr Phe Pro Met Ser Gln Ser Ser
 565 570 575

ttc acc gtc ggt gct gat acc ttc agc tct ggc aac gaa gtt tat ata 1776
 Phe Thr Val Gly Ala Asp Thr Phe Ser Ser Gly Asn Glu Val Tyr Ile
 580 585 590

gac aga ttt gag ttg att cca gtt act gca acg ttt gag tga 1818
 Asp Arg Phe Glu Leu Ile Pro Val Thr Ala Thr Phe Glu
 595 600 605

<210> 4
 <211> 605
 <212> PRT
 <213> Искусственная последовательность
 <220>

<223> Синтетическая конструкция

<400> 4

Met Glu Asn Asn Ile Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
 1 5 10 15

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

Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe
 35 40 45

Val Pro Gly Val Gly Val Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly
 50 55 60

Phe Ile Thr Pro Ser Asp Trp Ser Leu Phe Leu Leu Gln Ile Glu Gln
 65 70 75 80

Leu Ile Glu Gln Arg Ile Glu Thr Leu Glu Arg Asn Arg Ala Ile Thr
 85 90 95

Thr Leu Arg Gly Leu Ala Asp Ser Tyr Glu Ile Tyr Ile Glu Ala Leu
 100 105 110

Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp Val
 115 120 125

Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala Ile Asn
 130 135 140

Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val Tyr Val
 145 150 155 160

Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Ala Val Ser Phe
 165 170 175

Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His Tyr Asn
 180 185 190

Arg Leu Ile Asn Leu Ile His Arg Tyr Thr Lys His Cys Leu Asp Thr
 195 200 205

Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg Gln Trp
 210 215 220

Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val Leu Asp
 225 230 235 240

10

Ile Val Ala Leu Phe Pro Asn Tyr Asp Val Arg Thr Tyr Pro Ile Gln
 245 250 255

Thr Ser Ser Gln Leu Thr Arg Glu Ile Tyr Thr Ser Ser Val Ile Glu
 260 265 270

Asp Ser Pro Val Ser Ala Asn Ile Pro Asn Gly Phe Asn Arg Ala Glu
 275 280 285

Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser Leu Phe
 290 295 300

Val Thr Ala Glu Thr Val Arg Ser Gln Thr Val Trp Gly Gly His Leu
 305 310 315 320

Val Ser Ser Arg Asn Thr Ala Gly Asn Arg Ile Asn Phe Pro Ser Tyr
 325 330 335

Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu Asp Pro
 340 345 350

Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg Gly Gly
 355 360 365

Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala Phe Gln
 370 375 380

Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile
 385 390 395 400

Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp
 405 410 415

Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp Pro
 420 425 430

Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser Trp
 435 440 445

Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Glu Arg Ile
 450 455 460

Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly Thr Thr
 465 470 475 480

Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr
 485 490 495

11

Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val Asn Ile Asn Gly Gln Leu
500 505 510

Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu
515 520 525

Arg Ile Tyr Val Thr Val Ala Gly Glu Arg Ile Phe Ala Gly Gln Phe
530 535 540

Asn Lys Thr Met Asp Thr Gly Asp Pro Leu Thr Phe Gln Ser Phe Ser
545 550 555 560

Tyr Ala Thr Ile Asn Thr Ala Phe Thr Phe Pro Met Ser Gln Ser Ser
565 570 575

Phe Thr Val Gly Ala Asp Thr Phe Ser Ser Gly Asn Glu Val Tyr Ile
580 585 590

Asp Arg Phe Glu Leu Ile Pro Val Thr Ala Thr Phe Glu
595 600 605

<210> 5
<211> 1818
<212> ДНК
<213> Искусственная последовательность

<220>
<223> Синтетическая последовательность ДНК по изобретению, кодирующая ядерный токсин Cry1Fa *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой удаляются последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1

<220>
<221> CDS
<222> (1)..(1818)

<400> 5
atg gag aat aat atc cag aat caa tgc gtg cct tac aat tgt ctc aat 48
Met Glu Asn Asn Ile Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
1 5 10 15
aat ccc gag gtg gag ata tta aac gag gag aga tcc act ggc aga ctg 96
Asn Pro Glu Val Glu Ile Leu Asn Glu Glu Arg Ser Thr Gly Arg Leu
20 25 30
cca ctc gac ata tcc ttg tcc ctt acc cgt ttc ctt ttg agc gaa ttt 144
Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe
35 40 45
gtt cct ggt gtg gga gtg gct ttc gga ctg ttc gat ctg ata tgg ggc 192
Val Pro Gly Val Gly Val Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly
50 55 60
ttt atc act cct tct gat tgg agc ctc ttc ctt ctc cag att gag caa 240
Phe Ile Thr Pro Ser Asp Trp Ser Leu Phe Leu Leu Gln Ile Glu Gln
65 70 75 80

ttg att gag cag aga ata gaa acc ttg gaa agg aac cgt gca atc acg Leu Ile Glu Gln Arg Ile Glu Thr Leu Glu Arg Asn Arg Ala Ile Thr	288
85 90 95	
acc ttg cgc ggt ctc gcc gat agc tat gaa atc tac att gaa gca ctg Thr Leu Arg Gly Leu Ala Asp Ser Tyr Glu Ile Tyr Ile Glu Ala Leu	336
100 105 110	
agg gag tgg gag gcc aac ccc aat aat gct caa tta agg gaa gat gtg Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp Val	384
115 120 125	
cgt att cgt ttt gct aat aca gac gac gct ctc atc aca gca atc aat Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala Ile Asn	432
130 135 140	
aat ttc aca ctt aca tcc ttt gaa atc ccg ctt ttg agc gtg tac gtt Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val Tyr Val	480
145 150 155 160	
caa gcc gcc aat ctc cac ctc tca ctt ctg agg gac gct gtc tcc ttt Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Ala Val Ser Phe	528
165 170 175	
ggg caa ggt tgg gga ctg gat atc gct act gtg aat aat cac tac aat Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His Tyr Asn	576
180 185 190	
aga tta atc aac ctg att cat aga tat acg aag cac tgc ttg gac aca Arg Leu Ile Asn Leu Ile His Arg Tyr Thr Lys His Cys Leu Asp Thr	624
195 200 205	
tac aat caa gga ctg gag aac ctt agg gga act aac act agg cag tgg Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg Gln Trp	672
210 215 220	
gca agg ttc aac cag ttc aga cgt gat ctc aca ctt act gtg ctg gat Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val Leu Asp	720
225 230 235 240	
atc gtt gct ctc ttt ccg aac tac gat gtt cgc acc tac cca atc cag Ile Val Ala Leu Phe Pro Asn Tyr Asp Val Arg Thr Tyr Pro Ile Gln	768
245 250 255	
acg tca tcc caa tta aca agg gaa atc tac acc tcc tca gtg att gag Thr Ser Ser Gln Leu Thr Arg Glu Ile Tyr Thr Ser Ser Val Ile Glu	816
260 265 270	
gac tct ccc gtt tct gct aac ata cct aac ggc ttc aac cgc gcc gag Asp Ser Pro Val Ser Ala Asn Ile Pro Asn Gly Phe Asn Arg Ala Glu	864
275 280 285	
ttc gga gtt aga ccg ccc cac ctt atg gac ttt atg aat agc ttg ttc Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser Leu Phe	912
290 295 300	
gtg act gct gag act gtt aga agc caa act gtg tgg ggc ggc cac ttg Val Thr Ala Glu Thr Val Arg Ser Gln Thr Val Trp Gly Gly His Leu	960
305 310 315 320	
gtc agc tca cgc aac acg gct ggc aac cgt atc aac ttc ccg tct tac Val Ser Ser Arg Asn Thr Ala Gly Asn Arg Ile Asn Phe Pro Ser Tyr	1008
325 330 335	

ggg gtc ttt aac cct ggt ggc gcc att tgg att gca gac gag gac cca Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu Asp Pro 340 345 350	1056
cgt cct ttt tac cgc acc ctg tca gat ccg gtt ttc gtc aga ggc gga Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg Gly Gly 355 360 365	1104
ttt ggg aat cct cat tat gtc ctg ggc ctt agg gga gtg gct ttc caa Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala Phe Gln 370 375 380	1152
cag act ggc acc aac cac acc cgt acg ttt cgc aat agc ggg acc ata Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile 385 390 395 400	1200
gat tct ctt gat gaa atc cca cct caa gat aac agc ggc gca cct tgg Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp 405 410 415	1248
aac gat tat tcc cac gta tta aat cac gtt acg ttc gtc cgc tgg ccg Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp Pro 420 425 430	1296
ggg gag atc agc ggc agc gat tca tgg aga gca cca atg ttc tct tgg Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser Trp 435 440 445	1344
acg cac cgt tca gcc acc cct aca aat aca att gac ccg gag agg att Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Glu Arg Ile 450 455 460	1392
act caa atc cca ttg gtc aaa gca cat aca ctt cag tct ggg acc acc Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly Thr Thr 465 470 475 480	1440
gtg gtc aga ggg cct ggg ttc acg gga gga gac att ctt agg cgc aca Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr 485 490 495	1488
tcc gga gga ccc ttc gct tat act atc gtt aat ata aat ggg cag ctc Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val Asn Ile Asn Gly Gln Leu 500 505 510	1536
ccc cag cgc tat cgt gcc aga atc cgt tac gcc tct act aca aat ctc Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu 515 520 525	1584
aga atc tac gtg acg gtt gcc ggt gag cgc atc ttt gct ggt cag ttt Arg Ile Tyr Val Thr Val Ala Gly Glu Arg Ile Phe Ala Gly Gln Phe 530 535 540	1632
aac aag acg atg gat act ggc gac cca ctg aca ttc caa tct ttc tca Asn Lys Thr Met Asp Thr Gly Asp Pro Leu Thr Phe Gln Ser Phe Ser 545 550 555 560	1680
tac gca act att aat aca gct ttc aca ttc cca atg agc cag tca tct Tyr Ala Thr Ile Asn Thr Ala Phe Thr Phe Pro Met Ser Gln Ser Ser 565 570 575	1728
ttc acc gtc ggt gct gat acc ttc agc tct ggc aac gaa gtc tat atc Phe Thr Val Gly Ala Asp Thr Phe Ser Ser Gly Asn Glu Val Tyr Ile 580 585 590	1776

gac aga ttt gag ttg att cca gtt act gca acg ttt gag tga 1818
 Asp Arg Phe Glu Leu Ile Pro Val Thr Ala Thr Phe Glu
 595 600 605

<210> 6
 <211> 605
 <212> PRT
 <213> Искусственная последовательность

<220>
 <223> Синтетическая конструкция

<400> 6

Met Glu Asn Asn Ile Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
 1 5 10 15

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

Pro Leu Asp Ile Ser Leu Ser Leu Thr Arg Phe Leu Leu Ser Glu Phe
 35 40 45

Val Pro Gly Val Gly Val Ala Phe Gly Leu Phe Asp Leu Ile Trp Gly
 50 55 60

Phe Ile Thr Pro Ser Asp Trp Ser Leu Phe Leu Leu Gln Ile Glu Gln
 65 70 75 80

Leu Ile Glu Gln Arg Ile Glu Thr Leu Glu Arg Asn Arg Ala Ile Thr
 85 90 95

Thr Leu Arg Gly Leu Ala Asp Ser Tyr Glu Ile Tyr Ile Glu Ala Leu
 100 105 110

Arg Glu Trp Glu Ala Asn Pro Asn Asn Ala Gln Leu Arg Glu Asp Val
 115 120 125

Arg Ile Arg Phe Ala Asn Thr Asp Asp Ala Leu Ile Thr Ala Ile Asn
 130 135 140

Asn Phe Thr Leu Thr Ser Phe Glu Ile Pro Leu Leu Ser Val Tyr Val
 145 150 155 160

Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Ala Val Ser Phe
 165 170 175

Gly Gln Gly Trp Gly Leu Asp Ile Ala Thr Val Asn Asn His Tyr Asn
 180 185 190

15

Arg Leu Ile Asn Leu Ile His Arg Tyr Thr Lys His Cys Leu Asp Thr
 195 200 205

Tyr Asn Gln Gly Leu Glu Asn Leu Arg Gly Thr Asn Thr Arg Gln Trp
 210 215 220

Ala Arg Phe Asn Gln Phe Arg Arg Asp Leu Thr Leu Thr Val Leu Asp
 225 230 235 240

Ile Val Ala Leu Phe Pro Asn Tyr Asp Val Arg Thr Tyr Pro Ile Gln
 245 250 255

Thr Ser Ser Gln Leu Thr Arg Glu Ile Tyr Thr Ser Ser Val Ile Glu
 260 265 270

Asp Ser Pro Val Ser Ala Asn Ile Pro Asn Gly Phe Asn Arg Ala Glu
 275 280 285

Phe Gly Val Arg Pro Pro His Leu Met Asp Phe Met Asn Ser Leu Phe
 290 295 300

Val Thr Ala Glu Thr Val Arg Ser Gln Thr Val Trp Gly Gly His Leu
 305 310 315 320

Val Ser Ser Arg Asn Thr Ala Gly Asn Arg Ile Asn Phe Pro Ser Tyr
 325 330 335

Gly Val Phe Asn Pro Gly Gly Ala Ile Trp Ile Ala Asp Glu Asp Pro
 340 345 350

Arg Pro Phe Tyr Arg Thr Leu Ser Asp Pro Val Phe Val Arg Gly Gly
 355 360 365

Phe Gly Asn Pro His Tyr Val Leu Gly Leu Arg Gly Val Ala Phe Gln
 370 375 380

Gln Thr Gly Thr Asn His Thr Arg Thr Phe Arg Asn Ser Gly Thr Ile
 385 390 395 400

Asp Ser Leu Asp Glu Ile Pro Pro Gln Asp Asn Ser Gly Ala Pro Trp
 405 410 415

Asn Asp Tyr Ser His Val Leu Asn His Val Thr Phe Val Arg Trp Pro
 420 425 430

Gly Glu Ile Ser Gly Ser Asp Ser Trp Arg Ala Pro Met Phe Ser Trp
 435 440 445

16

Thr His Arg Ser Ala Thr Pro Thr Asn Thr Ile Asp Pro Glu Arg Ile
 450 455 460

Thr Gln Ile Pro Leu Val Lys Ala His Thr Leu Gln Ser Gly Thr Thr
 465 470 475 480

Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr
 485 490 495

Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val Asn Ile Asn Gly Gln Leu
 500 505 510

Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu
 515 520 525

Arg Ile Tyr Val Thr Val Ala Gly Glu Arg Ile Phe Ala Gly Gln Phe
 530 535 540

Asn Lys Thr Met Asp Thr Gly Asp Pro Leu Thr Phe Gln Ser Phe Ser
 545 550 555 560

Tyr Ala Thr Ile Asn Thr Ala Phe Thr Phe Pro Met Ser Gln Ser Ser
 565 570 575

Phe Thr Val Gly Ala Asp Thr Phe Ser Ser Gly Asn Glu Val Tyr Ile
 580 585 590

Asp Arg Phe Glu Leu Ile Pro Val Thr Ala Thr Phe Glu
 595 600 605

<210> 7
 <211> 372
 <212> ДНК
 <213> *Bacillus thuringiensis*

<220>
 <221> CDS
 <222> (1)..(372)
 <223> Нативная последовательность ДНК, кодирующая токсин Cry34Ab1 *Bacillus thuringiensis*

<400> 7
 atg tca gca cgt gaa gta cac att gat gta aat aat aag aca ggt cat 48
 Met Ser Ala Arg Glu Val His Ile Asp Val Asn Asn Lys Thr Gly His
 1 5 10 15

aca tta caa tta gaa gat aaa aca aaa ctt gat ggt ggt aga tgg cga 96
 Thr Leu Gln Leu Glu Asp Lys Thr Lys Leu Asp Gly Gly Arg Trp Arg
 20 25 30

aca tca cct aca aat gtt gct aat gat caa att aaa aca ttt gta gca 144
 Thr Ser Pro Thr Asn Val Ala Asn Asp Gln Ile Lys Thr Phe Val Ala
 35 40 45

gaa tca aat ggt ttt atg aca ggt aca gaa ggt act ata tat tat agt 192
 Glu Ser Asn Gly Phe Met Thr Gly Thr Glu Gly Thr Ile Tyr Tyr Ser
 50 55 60

ata aat gga gaa gca gaa att agt tta tat ttt gac aat cct ttt gca 240
 Ile Asn Gly Glu Ala Glu Ile Ser Leu Tyr Phe Asp Asn Pro Phe Ala
 65 70 75 80

ggt tct aat aaa tat gat gga cat tcc aat aaa tct caa tat gaa att 288
 Gly Ser Asn Lys Tyr Asp Gly His Ser Asn Lys Ser Gln Tyr Glu Ile
 85 90 95

att acc caa gga gga tca gga aat caa tct cat gtt acg tat act att 336
 Ile Thr Gln Gly Gly Ser Gly Asn Gln Ser His Val Thr Tyr Thr Ile
 100 105 110

caa acc aca tcc tca cga tat ggg cat aaa tca taa 372
 Gln Thr Thr Ser Ser Arg Tyr Gly His Lys Ser
 115 120

<210> 8
 <211> 123
 <212> PRT
 <213> Bacillus thuringiensis
 <400> 8

Met Ser Ala Arg Glu Val His Ile Asp Val Asn Asn Lys Thr Gly His
 1 5 10 15

Thr Leu Gln Leu Glu Asp Lys Thr Lys Leu Asp Gly Gly Arg Trp Arg
 20 25 30

Thr Ser Pro Thr Asn Val Ala Asn Asp Gln Ile Lys Thr Phe Val Ala
 35 40 45

Glu Ser Asn Gly Phe Met Thr Gly Thr Glu Gly Thr Ile Tyr Tyr Ser
 50 55 60

Ile Asn Gly Glu Ala Glu Ile Ser Leu Tyr Phe Asp Asn Pro Phe Ala
 65 70 75 80

Gly Ser Asn Lys Tyr Asp Gly His Ser Asn Lys Ser Gln Tyr Glu Ile
 85 90 95

Ile Thr Gln Gly Gly Ser Gly Asn Gln Ser His Val Thr Tyr Thr Ile
 100 105 110

Gln Thr Thr Ser Ser Arg Tyr Gly His Lys Ser
 115 120

<210> 9
 <211> 372
 <212> ДНК

<213> Искусственная последовательность

<220>

<223> Синтетическая последовательность ДНК, кодирующая токсин Cry34Ab1 *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1

<220>

<221> CDS

<222> (1)..(372)

<400> 9

atg tca gca cgg gag gtc cac atc gat gta aat aat aag acg ggt cat	48
Met Ser Ala Arg Glu Val His Ile Asp Val Asn Asn Lys Thr Gly His	
1 5 10 15	
aca tta cag ttg gag gat aaa aca aag cta gac ggt ggc aga tgg aga	96
Thr Leu Gln Leu Glu Asp Lys Thr Lys Leu Asp Gly Gly Arg Trp Arg	
20 25 30	
acc agt ccg acc aac gtt gct aac gat caa att aaa aca ttt gta gcc	144
Thr Ser Pro Thr Asn Val Ala Asn Asp Gln Ile Lys Thr Phe Val Ala	
35 40 45	
gaa tca aac ggt ttt atg act ggc acg gag ggg act ata tat tat tcc	192
Glu Ser Asn Gly Phe Met Thr Gly Thr Glu Gly Thr Ile Tyr Tyr Ser	
50 55 60	
atc aac gga gaa gcc gag att tcg tta tat ttt gac aat cca ttc gcg	240
Ile Asn Gly Glu Ala Glu Ile Ser Leu Tyr Phe Asp Asn Pro Phe Ala	
65 70 75 80	
ggg tct aat aaa tac gac gga cac tcc aat aaa tct caa tat gaa atc	288
Gly Ser Asn Lys Tyr Asp Gly His Ser Asn Lys Ser Gln Tyr Glu Ile	
85 90 95	
att aca caa ggc ggc agc gga aat caa agc cac gtc acg tat act atc	336
Ile Thr Gln Gly Gly Ser Gly Asn Gln Ser His Val Thr Tyr Thr Ile	
100 105 110	
cag acc act tca tcg cgc tac ggg cat aaa tca tag	372
Gln Thr Thr Ser Ser Arg Tyr Gly His Lys Ser	
115 120	

<210> 10

<211> 123

<212> PRT

<213> Искусственная последовательность

<220>

<223> Синтетическая конструкция

<400> 10

Met Ser Ala Arg Glu Val His Ile Asp Val Asn Asn Lys Thr Gly His	
1 5 10 15	
Thr Leu Gln Leu Glu Asp Lys Thr Lys Leu Asp Gly Gly Arg Trp Arg	
20 25 30	

19

Thr Ser Pro Thr Asn Val Ala Asn Asp Gln Ile Lys Thr Phe Val Ala
35 40 45

Glu Ser Asn Gly Phe Met Thr Gly Thr Glu Gly Thr Ile Tyr Tyr Ser
50 55 60

Ile Asn Gly Glu Ala Glu Ile Ser Leu Tyr Phe Asp Asn Pro Phe Ala
65 70 75 80

Gly Ser Asn Lys Tyr Asp Gly His Ser Asn Lys Ser Gln Tyr Glu Ile
85 90 95

Ile Thr Gln Gly Gly Ser Gly Asn Gln Ser His Val Thr Tyr Thr Ile
100 105 110

Gln Thr Thr Ser Ser Arg Tyr Gly His Lys Ser
115 120

<210> 11

<211> 372

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Синтетическая последовательность ДНК по изобретению, кодирующая токсин Cry34Ab1 *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой удаляются последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1

<220>

<221> CDS

<222> (1)..(372)

<400> 11

atg tca gca cgg gag gtc cac atc gat gta aat aat aag acg ggt cat 48
Met Ser Ala Arg Glu Val His Ile Asp Val Asn Asn Lys Thr Gly His
1 5 10 15

aca tta cag ttg gag gat aaa aca aag cta gac ggt ggc aga tgg aga 96
Thr Leu Gln Leu Glu Asp Lys Thr Lys Leu Asp Gly Gly Arg Trp Arg
20 25 30

acc agt ccg acc aac gtt gct aac gat caa att aaa aca ttt gta gcc 144
Thr Ser Pro Thr Asn Val Ala Asn Asp Gln Ile Lys Thr Phe Val Ala
35 40 45

gaa tca aac ggt ttc atg act ggc acg gag ggg act atc tac tac tcc 192
Glu Ser Asn Gly Phe Met Thr Gly Thr Glu Gly Thr Ile Tyr Tyr Ser
50 55 60

atc aac gga gaa gcc gag att tcg ctg tac ttc gac aat cca ttc gcg 240
Ile Asn Gly Glu Ala Glu Ile Ser Leu Tyr Phe Asp Asn Pro Phe Ala
65 70 75 80

ggg tct aat aaa tac gac gga cac tcc aat aaa tct caa tat gaa atc 288
Gly Ser Asn Lys Tyr Asp Gly His Ser Asn Lys Ser Gln Tyr Glu Ile
85 90 95

20

att aca caa ggc ggc agc gga aat caa agc cac gtc acg tat act atc 336
 ile Thr Gln Gly Gly Ser Gly Asn Gln Ser His Val Thr Tyr Thr Ile
 100 105 110

cag acc act tca tcc cgc tac ggg cat aaa tca tag 372
 Gln Thr Thr Ser Ser Arg Tyr Gly His Lys Ser
 115 120

<210> 12
 <211> 123
 <212> PRT
 <213> Искусственная последовательность

<220>
 <223> Синтетическая конструкция

<400> 12

Met Ser Ala Arg Glu Val His Ile Asp Val Asn Asn Lys Thr Gly His
 1 5 10 15

Thr Leu Gln Leu Glu Asp Lys Thr Lys Leu Asp Gly Gly Arg Trp Arg
 20 25 30

Thr Ser Pro Thr Asn Val Ala Asn Asp Gln Ile Lys Thr Phe Val Ala
 35 40 45

Glu Ser Asn Gly Phe Met Thr Gly Thr Glu Gly Thr Ile Tyr Tyr Ser
 50 55 60

Ile Asn Gly Glu Ala Glu Ile Ser Leu Tyr Phe Asp Asn Pro Phe Ala
 65 70 75 80

Gly Ser Asn Lys Tyr Asp Gly His Ser Asn Lys Ser Gln Tyr Glu Ile
 85 90 95

Ile Thr Gln Gly Gly Ser Gly Asn Gln Ser His Val Thr Tyr Thr Ile
 100 105 110

Gln Thr Thr Ser Ser Arg Tyr Gly His Lys Ser
 115 120

<210> 13
 <211> 1152
 <212> ДНК
 <213> *Bacillus thuringiensis*

<220>
 <221> CDS
 <222> (1)..(1152)
 <223> Нативная последовательность ДНК, кодирующая токсин Cry35Ab1 *Bacillus thuringiensis*

```

<400> 13
atg tta gat act aat aaa gtt tat gaa ata agc aat cat gct aat gga      48
Met Leu Asp Thr Asn Lys Val Tyr Glu Ile Ser Asn His Ala Asn Gly
1           5           10           15

cta tat gca gca act tat tta agt tta gat gat tca ggt gtt agt tta      96
Leu Tyr Ala Ala Thr Tyr Leu Ser Leu Asp Asp Ser Gly Val Ser Leu
           20           25           30

atg aat aaa aat gat gat gat att gat gat tat aac tta aaa tgg ttt      144
Met Asn Lys Asn Asp Asp Asp Ile Asp Asp Tyr Asn Leu Lys Trp Phe
           35           40           45

tta ttt cct att gat gat gat caa tat att att aca agc tat gca gca      192
Leu Phe Pro Ile Asp Asp Asp Gln Tyr Ile Ile Thr Ser Tyr Ala Ala
           50           55           60

aat aat tgt aaa gtt tgg aat gtt aat aat gat aaa ata aat gtt tcg      240
Asn Asn Cys Lys Val Trp Asn Val Asn Asn Asp Lys Ile Asn Val Ser
65           70           75           80

act tat tct tca aca aat tca ata caa aaa tgg caa ata aaa gct aat      288
Thr Tyr Ser Ser Thr Asn Ser Ile Gln Lys Trp Gln Ile Lys Ala Asn
           85           90           95

ggt tct tca tat gta ata caa agt gat aat gga aaa gtc tta aca gca      336
Gly Ser Ser Tyr Val Ile Gln Ser Asp Asn Gly Lys Val Leu Thr Ala
           100          105          110

gga acc ggt caa gct ctt gga ttg ata cgt tta act gat gaa tcc tca      384
Gly Thr Gly Gln Ala Leu Gly Leu Ile Arg Leu Thr Asp Glu Ser Ser
           115          120          125

aat aat ccc aat caa caa tgg aat tta act tct gta caa aca att caa      432
Asn Asn Pro Asn Gln Gln Trp Asn Leu Thr Ser Val Gln Thr Ile Gln
           130          135          140

ctt cca caa aaa cct ata ata gat aca aaa tta aaa gat tat ccc aaa      480
Leu Pro Gln Lys Pro Ile Ile Asp Thr Lys Leu Lys Asp Tyr Pro Lys
145           150          155          160

tat tca cca act gga aat ata gat aat gga aca tct cct caa tta atg      528
Tyr Ser Pro Thr Gly Asn Ile Asp Asn Gly Thr Ser Pro Gln Leu Met
           165          170          175

gga tgg aca tta gta cct tgt att atg gta aat gat cca aat ata gat      576
Gly Trp Thr Leu Val Pro Cys Ile Met Val Asn Asp Pro Asn Ile Asp
           180          185          190

aaa aat act caa att aaa act act cca tat tat att tta aaa aaa tat      624
Lys Asn Thr Gln Ile Lys Thr Thr Pro Tyr Tyr Ile Leu Lys Lys Tyr
           195          200          205

caa tat tgg caa cga gca gta gga agt aat gta gct tta cgt cca cat      672
Gln Tyr Trp Gln Arg Ala Val Gly Ser Asn Val Ala Leu Arg Pro His
           210          215          220

gaa aaa aaa tca tat act tat gaa tgg ggc aca gaa ata gat caa aaa      720
Glu Lys Lys Ser Tyr Thr Tyr Glu Trp Gly Thr Glu Ile Asp Gln Lys
225           230          235          240

aca aca att ata aat aca tta gga ttt caa atc aat ata gat tca gga      768
Thr Thr Ile Ile Asn Thr Leu Gly Phe Gln Ile Asn Ile Asp Ser Gly
           245          250          255

```

atg aaa ttt gat ata cca gaa gta ggt gga ggt aca gat gaa ata aaa 816
 Met Lys Phe Asp Ile Pro Glu Val Gly Gly Gly Thr Asp Glu Ile Lys
 260 265 270

aca caa cta aat gaa gaa tta aaa ata gaa tat agt cat gaa act aaa 864
 Thr Gln Leu Asn Glu Glu Leu Lys Ile Glu Tyr Ser His Glu Thr Lys
 275 280 285

ata atg gaa aaa tat caa gaa caa tct gaa ata gat aat cca act gat 912
 Ile Met Glu Lys Tyr Gln Glu Gln Ser Glu Ile Asp Asn Pro Thr Asp
 290 295 300

caa tca atg aat tct ata gga ttt ctt act att act tcc tta gaa tta 960
 Gln Ser Met Asn Ser Ile Gly Phe Leu Thr Ile Thr Ser Leu Glu Leu
 305 310 315 320

tat aga tat aat ggc tca gaa att cgt ata atg caa att caa acc tca 1008
 Tyr Arg Tyr Asn Gly Ser Glu Ile Arg Ile Met Gln Ile Gln Thr Ser
 325 330 335

gat aat gat act tat aat gtt act tct tat cca aat cat caa caa gct 1056
 Asp Asn Asp Thr Tyr Asn Val Thr Ser Tyr Pro Asn His Gln Gln Ala
 340 345 350

tta tta ctt ctt aca aat cat tca tat gaa gaa gta gaa gaa ata aca 1104
 Leu Leu Leu Leu Thr Asn His Ser Tyr Glu Glu Val Glu Glu Ile Thr
 355 360 365

aat att cct aaa agt aca cta aaa aaa tta aaa aaa tat tat ttt taa 1152
 Asn Ile Pro Lys Ser Thr Leu Lys Lys Leu Lys Lys Tyr Tyr Phe
 370 375 380

<210> 14
 <211> 383
 <212> PRT
 <213> Bacillus thuringiensis

<400> 14

Met Leu Asp Thr Asn Lys Val Tyr Glu Ile Ser Asn His Ala Asn Gly
 1 5 10 15

Leu Tyr Ala Ala Thr Tyr Leu Ser Leu Asp Asp Ser Gly Val Ser Leu
 20 25 30

Met Asn Lys Asn Asp Asp Asp Ile Asp Asp Tyr Asn Leu Lys Trp Phe
 35 40 45

Leu Phe Pro Ile Asp Asp Asp Gln Tyr Ile Ile Thr Ser Tyr Ala Ala
 50 55 60

Asn Asn Cys Lys Val Trp Asn Val Asn Asn Asp Lys Ile Asn Val Ser
 65 70 75 80

Thr Tyr Ser Ser Thr Asn Ser Ile Gln Lys Trp Gln Ile Lys Ala Asn
 85 90 95

Gly Ser Ser Tyr Val Ile Gln Ser Asp Asn Gly Lys Val Leu Thr Ala
 100 105 110

Gly Thr Gly Gln Ala Leu Gly Leu Ile Arg Leu Thr Asp Glu Ser Ser
 115 120 125

Asn Asn Pro Asn Gln Gln Trp Asn Leu Thr Ser Val Gln Thr Ile Gln
 130 135 140

Leu Pro Gln Lys Pro Ile Ile Asp Thr Lys Leu Lys Asp Tyr Pro Lys
 145 150 155 160

Tyr Ser Pro Thr Gly Asn Ile Asp Asn Gly Thr Ser Pro Gln Leu Met
 165 170 175

Gly Trp Thr Leu Val Pro Cys Ile Met Val Asn Asp Pro Asn Ile Asp
 180 185 190

Lys Asn Thr Gln Ile Lys Thr Thr Pro Tyr Tyr Ile Leu Lys Lys Tyr
 195 200 205

Gln Tyr Trp Gln Arg Ala Val Gly Ser Asn Val Ala Leu Arg Pro His
 210 215 220

Glu Lys Lys Ser Tyr Thr Tyr Glu Trp Gly Thr Glu Ile Asp Gln Lys
 225 230 235 240

Thr Thr Ile Ile Asn Thr Leu Gly Phe Gln Ile Asn Ile Asp Ser Gly
 245 250 255

Met Lys Phe Asp Ile Pro Glu Val Gly Gly Gly Thr Asp Glu Ile Lys
 260 265 270

Thr Gln Leu Asn Glu Glu Leu Lys Ile Glu Tyr Ser His Glu Thr Lys
 275 280 285

Ile Met Glu Lys Tyr Gln Glu Gln Ser Glu Ile Asp Asn Pro Thr Asp
 290 295 300

Gln Ser Met Asn Ser Ile Gly Phe Leu Thr Ile Thr Ser Leu Glu Leu
 305 310 315 320

Tyr Arg Tyr Asn Gly Ser Glu Ile Arg Ile Met Gln Ile Gln Thr Ser
 325 330 335

Asp Asn Asp Thr Tyr Asn Val Thr Ser Tyr Pro Asn His Gln Gln Ala
 340 345 350

Leu Leu Leu Leu Thr Asn His Ser Tyr Glu Glu Val Glu Glu Ile Thr
 355 360 365

Asn Ile Pro Lys Ser Thr Leu Lys Lys Leu Lys Lys Tyr Tyr Phe
 370 375 380

<210> 15
 <211> 1152
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Синтетическая последовательность ДНК, кодирующая токсин Cry35Ab1 *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1

<220>
 <221> CDS
 <222> (1)..(1152)

<400> 15
 atg ctc gat act aat aaa gtg tat gaa ata agc aac cat gcc aac ggg 48
 Met Leu Asp Thr Asn Lys Val Tyr Glu Ile Ser Asn His Ala Asn Gly
 1 5 10 15
 cta tat gcc gca act tat ttg agt ctg gac gac agc ggt gtg agc tta 96
 Leu Tyr Ala Ala Thr Tyr Leu Ser Leu Asp Asp Ser Gly Val Ser Leu
 20 25 30
 atg aat aaa aac gac gac gac att gac gac tac aac ctc aag tgg ttt 144
 Met Asn Lys Asn Asp Asp Asp Ile Asp Asp Tyr Asn Leu Lys Trp Phe
 35 40 45
 tta ttt cct att gac gac gat caa tat att att aca agc tac gca gca 192
 Leu Phe Pro Ile Asp Asp Asp Gln Tyr Ile Ile Thr Ser Tyr Ala Ala
 50 55 60
 aat aat tgc aaa gtc tgg aac gtt aat aat gat aaa ata aat gtt tcg 240
 Asn Asn Cys Lys Val Trp Asn Val Asn Asn Asp Lys Ile Asn Val Ser
 65 70 75 80
 acc tac agc tcc acc aac tca ata caa aag tgg caa ata aaa gct aat 288
 Thr Tyr Ser Ser Thr Asn Ser Ile Gln Lys Trp Gln Ile Lys Ala Asn
 85 90 95
 gga tcg tcg tat gta ata cag agt gac aat ggg aag gtc ttg aca gcg 336
 Gly Ser Ser Tyr Val Ile Gln Ser Asp Asn Gly Lys Val Leu Thr Ala
 100 105 110
 ggc act ggt caa gct ctt gga ctc ata agg ctc act gac gag tcc tca 384
 Gly Thr Gly Gln Ala Leu Gly Leu Ile Arg Leu Thr Asp Glu Ser Ser
 115 120 125
 aat aat ccc aat caa cag tgg aac ttg act tcc gtg cag acg atc caa 432
 Asn Asn Pro Asn Gln Gln Trp Asn Leu Thr Ser Val Gln Thr Ile Gln
 130 135 140
 ctt cca cag aaa cct atc atc gat aca aaa tta aaa gat tac ccc aag 480
 Leu Pro Gln Lys Pro Ile Ile Asp Thr Lys Leu Lys Asp Tyr Pro Lys
 145 150 155 160

25

tac tcg cca acc ggc aac atc gat aat gga acg tct cct caa tta atg 528
 Tyr Ser Pro Thr Gly Asn Ile Asp Asn Gly Thr Ser Pro Gln Leu Met
 165 170 175

ggc tgg acc ctc gta ccc tgt att atg gtg aac gac ccg aat atc gat 576
 Gly Trp Thr Leu Val Pro Cys Ile Met Val Asn Asp Pro Asn Ile Asp
 180 185 190

aaa aat act caa att aaa acc acg ccg tat tat ata ttg aaa aaa tac 624
 Lys Asn Thr Gln Ile Lys Thr Thr Pro Tyr Tyr Ile Leu Lys Lys Tyr
 195 200 205

caa tac tgg cag cgc gcg gtt ggc tca aac gtc gct ctg cgg cca cat 672
 Gln Tyr Trp Gln Arg Ala Val Gly Ser Asn Val Ala Leu Arg Pro His
 210 215 220

gaa aag aag tcc tac act tac gaa tgg ggc aca gag atc gat cag aaa 720
 Glu Lys Lys Ser Tyr Thr Tyr Glu Trp Gly Thr Glu Ile Asp Gln Lys
 225 230 235 240

acg acc att ata aat aca tta gga ttc caa atc aat atc gac agc gga 768
 Thr Thr Ile Ile Asn Thr Leu Gly Phe Gln Ile Asn Ile Asp Ser Gly
 245 250 255

atg aaa ttt gac atc ccg gaa gtg ggg ggc ggg acc gat gaa ata aaa 816
 Met Lys Phe Asp Ile Pro Glu Val Gly Gly Gly Thr Asp Glu Ile Lys
 260 265 270

acg cag ctc aac gaa gaa tta aaa ata gag tac agt cat gaa act aaa 864
 Thr Gln Leu Asn Glu Glu Leu Lys Ile Glu Tyr Ser His Glu Thr Lys
 275 280 285

ata atg gaa aaa tat caa gag caa tct gaa atc gat aac ccg acc gac 912
 Ile Met Glu Lys Tyr Gln Glu Gln Ser Glu Ile Asp Asn Pro Thr Asp
 290 295 300

caa tca atg aac tct atc ggt ttc ctt act att acc tcc ctg gag tta 960
 Gln Ser Met Asn Ser Ile Gly Phe Leu Thr Ile Thr Ser Leu Glu Leu
 305 310 315 320

tat aga tat aac ggc tct gag atc cgt ata atg cag att caa acc tca 1008
 Tyr Arg Tyr Asn Gly Ser Glu Ile Arg Ile Met Gln Ile Gln Thr Ser
 325 330 335

gac aat gac act tat aac gtc acc tct tac ccg aat cat caa caa gct 1056
 Asp Asn Asp Thr Tyr Asn Val Thr Ser Tyr Pro Asn His Gln Gln Ala
 340 345 350

tta ttg ctt ctt aca aac cac agt tat gaa gag gtg gaa gaa ata acg 1104
 Leu Leu Leu Leu Thr Asn His Ser Tyr Glu Glu Val Glu Glu Ile Thr
 355 360 365

aac att cct aaa tcc aca cta aag aaa tta aaa aaa tat tat ttt tga 1152
 Asn Ile Pro Lys Ser Thr Leu Lys Lys Leu Lys Lys Tyr Tyr Phe
 370 375 380

<210> 16
 <211> 383
 <212> PRT
 <213> Искусственная последовательность
 <220>

<223> Синтетическая конструкция

<400> 16

Met Leu Asp Thr Asn Lys Val Tyr Glu Ile Ser Asn His Ala Asn Gly
 1 5 10 15

Leu Tyr Ala Ala Thr Tyr Leu Ser Leu Asp Asp Ser Gly Val Ser Leu
 20 25 30

Met Asn Lys Asn Asp Asp Asp Ile Asp Asp Tyr Asn Leu Lys Trp Phe
 35 40 45

Leu Phe Pro Ile Asp Asp Asp Gln Tyr Ile Ile Thr Ser Tyr Ala Ala
 50 55 60

Asn Asn Cys Lys Val Trp Asn Val Asn Asn Asp Lys Ile Asn Val Ser
 65 70 75 80

Thr Tyr Ser Ser Thr Asn Ser Ile Gln Lys Trp Gln Ile Lys Ala Asn
 85 90 95

Gly Ser Ser Tyr Val Ile Gln Ser Asp Asn Gly Lys Val Leu Thr Ala
 100 105 110

Gly Thr Gly Gln Ala Leu Gly Leu Ile Arg Leu Thr Asp Glu Ser Ser
 115 120 125

Asn Asn Pro Asn Gln Gln Trp Asn Leu Thr Ser Val Gln Thr Ile Gln
 130 135 140

Leu Pro Gln Lys Pro Ile Ile Asp Thr Lys Leu Lys Asp Tyr Pro Lys
 145 150 155 160

Tyr Ser Pro Thr Gly Asn Ile Asp Asn Gly Thr Ser Pro Gln Leu Met
 165 170 175

Gly Trp Thr Leu Val Pro Cys Ile Met Val Asn Asp Pro Asn Ile Asp
 180 185 190

Lys Asn Thr Gln Ile Lys Thr Thr Pro Tyr Tyr Ile Leu Lys Lys Tyr
 195 200 205

Gln Tyr Trp Gln Arg Ala Val Gly Ser Asn Val Ala Leu Arg Pro His
 210 215 220

Glu Lys Lys Ser Tyr Thr Tyr Glu Trp Gly Thr Glu Ile Asp Gln Lys
 225 230 235 240

27

Thr Thr Ile Ile Asn Thr Leu Gly Phe Gln Ile Asn Ile Asp Ser Gly
245 250 255

Met Lys Phe Asp Ile Pro Glu Val Gly Gly Gly Thr Asp Glu Ile Lys
260 265 270

Thr Gln Leu Asn Glu Glu Leu Lys Ile Glu Tyr Ser His Glu Thr Lys
275 280 285

Ile Met Glu Lys Tyr Gln Glu Gln Ser Glu Ile Asp Asn Pro Thr Asp
290 295 300

Gln Ser Met Asn Ser Ile Gly Phe Leu Thr Ile Thr Ser Leu Glu Leu
305 310 315 320

Tyr Arg Tyr Asn Gly Ser Glu Ile Arg Ile Met Gln Ile Gln Thr Ser
325 330 335

Asp Asn Asp Thr Tyr Asn Val Thr Ser Tyr Pro Asn His Gln Gln Ala
340 345 350

Leu Leu Leu Leu Thr Asn His Ser Tyr Glu Glu Val Glu Glu Ile Thr
355 360 365

Asn Ile Pro Lys Ser Thr Leu Lys Lys Leu Lys Lys Tyr Tyr Phe
370 375 380

<210> 17

<211> 1152

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Синтетическая последовательность ДНК по изобретению, кодирующая токсин Cry35Ab1 *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой удаляются последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1

<220>

<221> CDS

<222> (1)..(1152)

<400> 17

atg ctc gat act aat aaa gtg tat gaa ata tcg aac cat gcc aac ggg 48
Met Leu Asp Thr Asn Lys Val Tyr Glu Ile Ser Asn His Ala Asn Gly
1 5 10 15

cta tat gcc gca act tac ctg agt ctg gac gat agt ggt gtg agc tta 96
Leu Tyr Ala Ala Thr Tyr Leu Ser Leu Asp Asp Ser Gly Val Ser Leu
20 25 30

atg aat aaa aac gac gac gac att gac gac tac aac ctc aag tgg ttc 144
Met Asn Lys Asn Asp Asp Asp Ile Asp Asp Tyr Asn Leu Lys Trp Phe
35 40 45

ctg ttt cct att gac gac gat cag tat atc att aca agc tac gca gcg Leu Phe Pro Ile Asp Asp Asp Gln Tyr Ile Ile Thr Ser Tyr Ala Ala 50 55 60	192
aat aat tgc aaa gtc tgg aac gtc aat aat gat aaa atc aat gtt tcg Asn Asn Cys Lys Val Trp Asn Val Asn Asn Asp Lys Ile Asn Val Ser 65 70 75 80	240
acc tac agc tcc acc aac tca ata caa aag tgg caa atc aaa gct aat Thr Tyr Ser Ser Thr Asn Ser Ile Gln Lys Trp Gln Ile Lys Ala Asn 85 90 95	288
ggc agc tcg tac gta ata cag agt gac aat ggg aag gtc ttg aca gcg Gly Ser Ser Tyr Val Ile Gln Ser Asp Asn Gly Lys Val Leu Thr Ala 100 105 110	336
ggc act ggt caa gct ctt gga ctc ata agg ctc act gac gag tcc tcg Gly Thr Gly Gln Ala Leu Gly Leu Ile Arg Leu Thr Asp Glu Ser Ser 115 120 125	384
aat aat ccc aat caa cag tgg aac ttg act tcc gtg cag acg atc caa Asn Asn Pro Asn Gln Gln Trp Asn Leu Thr Ser Val Gln Thr Ile Gln 130 135 140	432
ctt cca cag aaa cct atc atc gat aca aaa tta aaa gat tac ccc aag Leu Pro Gln Lys Pro Ile Ile Asp Thr Lys Leu Lys Asp Tyr Pro Lys 145 150 155 160	480
tac tcg cca acc ggc aac atc gat aat gga acg tct cct caa tta atg Tyr Ser Pro Thr Gly Asn Ile Asp Asn Gly Thr Ser Pro Gln Leu Met 165 170 175	528
ggc tgg acc ctc gta ccc tgt ata atg gtg aac gac ccg aat atc gat Gly Trp Thr Leu Val Pro Cys Ile Met Val Asn Asp Pro Asn Ile Asp 180 185 190	576
aaa aat act caa att aaa acc acg ccg tac tac ata ctc aaa aaa tac Lys Asn Thr Gln Ile Lys Thr Thr Pro Tyr Tyr Ile Leu Lys Lys Tyr 195 200 205	624
caa tac tgg cag cgc gcg gtt ggc tca aac gtc gct ctg cgg cca cat Gln Tyr Trp Gln Arg Ala Val Gly Ser Asn Val Ala Leu Arg Pro His 210 215 220	672
gaa aag aag tcc tac act tac gaa tgg ggc aca gag atc gat cag aaa Glu Lys Lys Ser Tyr Thr Tyr Glu Trp Gly Thr Glu Ile Asp Gln Lys 225 230 235 240	720
acg acc att ata aat aca tta gga ttc caa atc aat atc gac agc gga Thr Thr Ile Ile Asn Thr Leu Gly Phe Gln Ile Asn Ile Asp Ser Gly 245 250 255	768
atg aaa ttt gac atc ccg gaa gtg ggg ggc ggg acc gat gaa ata aaa Met Lys Phe Asp Ile Pro Glu Val Gly Gly Thr Asp Glu Ile Lys 260 265 270	816
acg cag ctc aac gaa gaa tta aaa ata gag tac agt cat gaa act aaa Thr Gln Leu Asn Glu Glu Leu Lys Ile Glu Tyr Ser His Glu Thr Lys 275 280 285	864
ata atg gaa aaa tat caa gag caa tct gaa atc gat aac ccg acc gac Ile Met Glu Lys Tyr Gln Glu Gln Ser Glu Ile Asp Asn Pro Thr Asp 290 295 300	912

caa tca atg aac tct atc ggt ttc ctt act att acc tcc ctg gag ttg 960
 Gln Ser Met Asn Ser Ile Gly Phe Leu Thr Ile Thr Ser Leu Glu Leu
 305 310 315 320

tac aga tat aac ggc tct gag atc cgt ata atg cag att caa acc tca 1008
 Tyr Arg Tyr Asn Gly Ser Glu Ile Arg Ile Met Gln Ile Gln Thr Ser
 325 330 335

gac aat gac act tat aac gtc acc tct tac ccg aat cat cag caa gcc 1056
 Asp Asn Asp Thr Tyr Asn Val Thr Ser Tyr Pro Asn His Gln Gln Ala
 340 345 350

ctg ctg ctt ctt aca aac cac agt tat gaa gag gtg gaa gag ata acg 1104
 Leu Leu Leu Leu Thr Asn His Ser Tyr Glu Glu Val Glu Glu Ile Thr
 355 360 365

aac att cct aaa tcc aca cta aag aaa tta aaa aaa tat tat ttc tga 1152
 Asn Ile Pro Lys Ser Thr Leu Lys Lys Leu Lys Lys Tyr Tyr Phe
 370 375 380

<210> 18

<211> 383

<212> PRT

<213> Искусственная последовательность

<220>

<223> Синтетическая конструкция

<400> 18

Met Leu Asp Thr Asn Lys Val Tyr Glu Ile Ser Asn His Ala Asn Gly
 1 5 10 15

Leu Tyr Ala Ala Thr Tyr Leu Ser Leu Asp Asp Ser Gly Val Ser Leu
 20 25 30

Met Asn Lys Asn Asp Asp Asp Ile Asp Asp Tyr Asn Leu Lys Trp Phe
 35 40 45

Leu Phe Pro Ile Asp Asp Asp Gln Tyr Ile Ile Thr Ser Tyr Ala Ala
 50 55 60

Asn Asn Cys Lys Val Trp Asn Val Asn Asn Asp Lys Ile Asn Val Ser
 65 70 75 80

Thr Tyr Ser Ser Thr Asn Ser Ile Gln Lys Trp Gln Ile Lys Ala Asn
 85 90 95

Gly Ser Ser Tyr Val Ile Gln Ser Asp Asn Gly Lys Val Leu Thr Ala
 100 105 110

Gly Thr Gly Gln Ala Leu Gly Leu Ile Arg Leu Thr Asp Glu Ser Ser
 115 120 125

30

Asn Asn Pro Asn Gln Gln Trp Asn Leu Thr Ser Val Gln Thr Ile Gln
 130 135 140

Leu Pro Gln Lys Pro Ile Ile Asp Thr Lys Leu Lys Asp Tyr Pro Lys
 145 150 155 160

Tyr Ser Pro Thr Gly Asn Ile Asp Asn Gly Thr Ser Pro Gln Leu Met
 165 170 175

Gly Trp Thr Leu Val Pro Cys Ile Met Val Asn Asp Pro Asn Ile Asp
 180 185 190

Lys Asn Thr Gln Ile Lys Thr Thr Pro Tyr Tyr Ile Leu Lys Lys Tyr
 195 200 205

Gln Tyr Trp Gln Arg Ala Val Gly Ser Asn Val Ala Leu Arg Pro His
 210 215 220

Glu Lys Lys Ser Tyr Thr Tyr Glu Trp Gly Thr Glu Ile Asp Gln Lys
 225 230 235 240

Thr Thr Ile Ile Asn Thr Leu Gly Phe Gln Ile Asn Ile Asp Ser Gly
 245 250 255

Met Lys Phe Asp Ile Pro Glu Val Gly Gly Gly Thr Asp Glu Ile Lys
 260 265 270

Thr Gln Leu Asn Glu Glu Leu Lys Ile Glu Tyr Ser His Glu Thr Lys
 275 280 285

Ile Met Glu Lys Tyr Gln Glu Gln Ser Glu Ile Asp Asn Pro Thr Asp
 290 295 300

Gln Ser Met Asn Ser Ile Gly Phe Leu Thr Ile Thr Ser Leu Glu Leu
 305 310 315 320

Tyr Arg Tyr Asn Gly Ser Glu Ile Arg Ile Met Gln Ile Gln Thr Ser
 325 330 335

Asp Asn Asp Thr Tyr Asn Val Thr Ser Tyr Pro Asn His Gln Gln Ala
 340 345 350

Leu Leu Leu Leu Thr Asn His Ser Tyr Glu Glu Val Glu Glu Ile Thr
 355 360 365

Asn Ile Pro Lys Ser Thr Leu Lys Lys Leu Lys Lys Tyr Tyr Phe
 370 375 380

<210> 19
 <211> 1830
 <212> ДНК
 <213> *Bacillus thuringiensis*

<220>
 <221> CDS
 <222> (1)..(1830)
 <223> Нативная последовательность ДНК, кодирующая ядерный токсин Cry1Ab1
Bacillus thuringiensis

<400> 19
 atg gat aac aat ccg aac atc aat gaa tgc att cct tat aat tgt tta 48
 Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
 1 5 10 15

agt aac cct gaa gta gaa gta tta ggt gga gaa aga ata gaa act ggt 96
 Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly
 20 25 30

tac acc cca atc gat att tcc ttg tgc cta acg caa ttt ctt ttg agt 144
 Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
 35 40 45

gaa ttt gtt ccc ggt gct gga ttt gtg tta gga cta gtt gat ata ata 192
 Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
 50 55 60

tgg gga att ttt ggt ccc tct caa tgg gac gca ttt ctt gta caa att 240
 Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
 65 70 75 80

gaa cag tta att aac caa aga ata gaa gaa ttc gct agg aac caa gcc 288
 Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
 85 90 95

att tct aga tta gaa gga cta agc aat ctt tat caa att tac gca gaa 336
 Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu
 100 105 110

tct ttt aga gag tgg gaa gca gat cct act aat cca gca tta aga gaa 384
 Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu
 115 120 125

gag atg cgt att caa ttc aat gac atg aac agt gcc ctt aca acc gct 432
 Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala
 130 135 140

att cct ctt ttt gca gtt caa aat tat caa gtt cct ctt tta tca gta 480
 Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val
 145 150 155 160

tat gtt caa gct gca aat tta cat tta tca gtt ttg aga gat gtt tca 528
 Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser
 165 170 175

gtg ttt gga caa agg tgg gga ttt gat gcc gcg act atc aat agt cgt 576
 Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg
 180 185 190

tat aat gat tta act agg ctt att ggc aac tat aca gat tat gct gta 624
 Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val
 195 200 205

cgc tgg tac aat acg gga tta gaa cgt gta tgg gga ccg gat tct aga Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg 210 215 220	672
gat tgg gta agg tat aat caa ttt aga aga gaa tta aca cta act gta Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val 225 230 235 240	720
tta gat atc gtt gct ctg ttc ccg aat tat gat agt aga aga tat cca Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro 245 250 255	768
att cga aca gtt tcc caa tta aca aga gaa att tat aca aac cca gta Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val 260 265 270	816
tta gaa aat ttt gat ggt agt ttt cga ggc tcg gct cag ggc ata gaa Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu 275 280 285	864
aga agt att agg agt cca cat ttg atg gat ata ctt aac agt ata acc Arg Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr 290 295 300	912
atc tat acg gat gct cat agg ggt tat tat tat tgg tca ggg cat caa Ile Tyr Thr Asp Ala His Arg Gly Tyr Tyr Tyr Trp Ser Gly His Gln 305 310 315 320	960
ata atg gct tct cct gta ggg ttt tcg ggg cca gaa ttc act ttt ccg Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro 325 330 335	1008
cta tat gga act atg gga aat gca gct cca caa caa cgt att gtt gct Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala 340 345 350	1056
caa cta ggt cag ggc gtg tat aga aca tta tcg tcc act tta tat aga Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg 355 360 365	1104
aga cct ttt aat ata ggg ata aat aat caa caa cta tct gtt ctt gac Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp 370 375 380	1152
ggg aca gaa ttt gct tat gga acc tcc tca aat ttg cca tcc gct gta Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val 385 390 395 400	1200
tac aga aaa agc gga acg gta gat tcg ctg gat gaa ata ccg cca cag Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln 405 410 415	1248
aat aac aac gtg cca cct agg caa gga ttt agt cat cga tta agc cat Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His 420 425 430	1296
gtt tca atg ttt cgt tca ggc ttt agt aat agt agt gta agt ata ata Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile 435 440 445	1344
aga gct cct atg ttc tct tgg ata cat cgt agt gct gaa ttt aat aat Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn 450 455 460	1392

ata att cct tca tca caa att aca caa ata cct tta aca aaa tct act 1440
 Ile Ile Pro Ser Ser Gln Ile Thr Gln Ile Pro Leu Thr Lys Ser Thr
 465 470 475 480

aat ctt ggc tct gga act tct gtc gtt aaa gga cca gga ttt aca gga 1488
 Asn Leu Gly Ser Gly Thr Ser Val Val Lys Gly Pro Gly Phe Thr Gly
 485 490

gga gat att ctt cga aga act tca cct ggc cag att tca acc tta aga 1536
 Gly Asp Ile Leu Arg Arg Thr Ser Pro Gly Gln Ile Ser Thr Leu Arg
 500 505 510

gta aat att act gca cca tta tca caa aga tat cgg gta aga att cgc 1584
 Val Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg Val Arg Ile Arg
 515 520 525

tac gct tct acc aca aat tta caa ttc cat aca tca att gac gga aga 1632
 Tyr Ala Ser Thr Thr Asn Leu Gln Phe His Thr Ser Ile Asp Gly Arg
 530 535 540

cct att aat cag ggg aat ttt tca gca act atg agt agt ggg agt aat 1680
 Pro Ile Asn Gln Gly Asn Phe Ser Ala Thr Met Ser Ser Gly Ser Asn
 545 550 555 560

tta cag tcc gga agc ttt agg act gta ggt ttt act act ccg ttt aac 1728
 Leu Gln Ser Gly Ser Phe Arg Thr Val Gly Phe Thr Thr Pro Phe Asn
 565 570 575

ttt tca aat gga tca agt gta ttt acg tta agt gct cat gtc ttc aat 1776
 Phe Ser Asn Gly Ser Ser Val Phe Thr Leu Ser Ala His Val Phe Asn
 580 585 590

tca ggc aat gaa gtt tat ata gat cga att gaa ttt gtt ccg gca gaa 1824
 Ser Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala Glu
 59 600 605

gta acc 1830
 Val Thr
 610

<210> 20
 <211> 610
 <212> PRT
 <213> Bacillus thuringiensis

<400> 20

Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
 1 5 10 15

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

Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
 35 40 45

Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
 50 55 60

Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
65 70 75 80

Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
85 90 95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu
100 105 110

Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu
115 120 125

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala
130 135 140

Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val
145 150 155 160

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser
165 170 175

Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg
180 185 190

Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val
195 200 205

Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg
210 215 220

Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val
225 230 235 240

Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro
245 250 255

Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val
260 265 270

Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu
275 280 285

Arg Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr
290 295 300

Ile Tyr Thr Asp Ala His Arg Gly Tyr Tyr Tyr Trp Ser Gly His Gln
305 310 315 320

Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro
 325 330 335

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
 340 345 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
 355 360 365

Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp
 370 375 380

Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val
 385 390 395 400

Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln
 405 410 415

Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His
 420 425 430

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
 435 440 445

Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn
 450 455 460

Ile Ile Pro Ser Ser Gln Ile Thr Gln Ile Pro Leu Thr Lys Ser Thr
 465 470 475 480

Asn Leu Gly Ser Gly Thr Ser Val Val Lys Gly Pro Gly Phe Thr Gly
 485 490 495

Gly Asp Ile Leu Arg Arg Thr Ser Pro Gly Gln Ile Ser Thr Leu Arg
 500 505 510

Val Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg Val Arg Ile Arg
 515 520 525

Tyr Ala Ser Thr Thr Asn Leu Gln Phe His Thr Ser Ile Asp Gly Arg
 530 535 540

Pro Ile Asn Gln Gly Asn Phe Ser Ala Thr Met Ser Ser Gly Ser Asn
 545 550 555 560

Leu Gln Ser Gly Ser Phe Arg Thr Val Gly Phe Thr Thr Pro Phe Asn
 565 570 575

Phe Ser Asn Gly Ser Ser Val Phe Thr Leu Ser Ala His Val Phe Asn
580 585 590

Ser Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala Glu
595 600 605

Val Thr
610

<210> 21

<211> 1830

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Синтетическая последовательность ДНК, кодирующая ядерный токсин Cry1Ab1 *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1

<220>

<221> CDS

<222> (1)..(1830)

<400> 21

atg gat aac aac ccg aac atc aat gag tgc atc ccg tat aac tgt ctc 48
Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
1 5 10 15

agt aac cct gaa gtg gag gtc tta ggt ggc gaa cgc atc gaa act ggt 96
Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly
20 25 30

tac acc cca atc gac att agc ttg tgc ttg acg cag ttc ctt ttg tcc 144
Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
35 40 45

gag ttc gtg ccc ggt gcg ggt ttc gtg ctg ggg cta gtt gat ata atc 192
Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
50 55 60

tgg gga atc ttt ggt ccc tct cag tgg gac gcc ttt ctt gtg caa att 240
Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
65 70 75 80

gag cag cta att aac caa aga ata gaa gag ttc gcg agg aac caa gcc 288
Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
85 90 95

att tcc aga ctg gag gga cta agc aac ctt tat caa atc tac gcg gag 336
Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu
100 105 110

tct ttt agg gag tgg gag gca gat cct acg aac ccg gca ctg cgc gaa 384
Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu
115 120 125

gag atg cgt att cag ttc aac gac atg aac agt gcc ctt aca acc gct 432
Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala
130 135 140

att ccc ctt ttc gca gtt caa aat tac caa gtt ccc ctt ctc tca gtg Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val 145 150 155 160	480
tac gtt caa gcc gca aat tta cac cta agc gtt ctc cgc gat gtg tca Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser 165 170 175	528
gtg ttt ggc cag agg tgg gga ttt gat gcc gcc act atc aat agt cgt Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg 180 185 190	576
tat aat gat ctg acg agg ctt atc gcc aac tat acc gac tat gct gtc Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val 195 200 205	624
cgc tgg tac aat acg gga tta gag cgg gtc tgg ggt ccg gat tcc cga Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg 210 215 220	672
gac tgg gtg cgc tac aat caa ttc cgc cgc gaa tta acc ctc act gtc Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val 225 230 235 240	720
ctc gac atc gtg gcg ctg ttc ccg aac tac gac agt agg aga tac cca Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro 245 250 255	768
atc cgc aca gtt tcc caa tta acg cgg gaa att tac acc aac cca gtc Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val 260 265 270	816
ctg gag aat ttt gac ggg agc ttc cga ggc tcg gct caa ggc ata gaa Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu 275 280 285	864
cgc agc att agg tcg cca cac ttg atg gat atc ctt aac agc atc acc Arg Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr 290 295 300	912
atc tac acg gat gcc cat agg ggt tac tac tac tgg tcg ggg cat caa Ile Tyr Thr Asp Ala His Arg Gly Tyr Tyr Tyr Trp Ser Gly His Gln 305 310 315 320	960
ata atg gct tct cct gtc ggg ttt tcg ggg cca gag ttc acc ttc ccg Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro 325 330 335	1008
ctc tac ggc act atg gga aat gcc gcg cca caa caa cgt atc gtc gct Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala 340 345 350	1056
caa cta ggt caa ggc gtg tac cgg aca ctg tcg tcc act ctc tat cgg Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg 355 360 365	1104
cgg cct ttc aat ata ggg ata aat aat caa cag ttg tct gtg ctg gac Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp 370 375 380	1152
ggg aca gag ttt gct tac gga acc tca agc aac ttg cca tcc gct gta Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val 385 390 395 400	1200

tac aga aaa agc ggc acg gtg gac tcg ctg gat gaa atc ccg ccc cag 1248
 Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln
 405 410 415

aat aac aac gtg ccc cct cgg caa ggc ttc agt cat cga ctg agc cac 1296
 Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His
 420 425 430

gtt agc atg ttc cgt tcg ggc ttc agc aac tcc tcc gta agt ata ata 1344
 Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
 435 440 445

aga gca cct atg ttc agc tgg ata cat cgt tcc gcc gag ttt aat aat 1392
 Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn
 450 455 460

ata att ccc tcc tct caa atc aca cag atc cct ctg aca aag tct act 1440
 Ile Ile Pro Ser Ser Gln Ile Thr Gln Ile Pro Leu Thr Lys Ser Thr
 465 470 475 480

aat ctt ggc tct ggg act tct gtc gtt aag ggg cct ggc ttt acg ggc 1488
 Asn Leu Gly Ser Gly Thr Ser Val Val Lys Gly Pro Gly Phe Thr Gly
 485 490 495

ggc gat att ctg cgg aga act tca cct ggc cag att tcc acc ctg cgc 1536
 Gly Asp Ile Leu Arg Arg Thr Ser Pro Gly Gln Ile Ser Thr Leu Arg
 500 505 510

gtg aat atc acc gcg cca ttg tca caa cgt tac cgc gtg cgg att cgc 1584
 Val Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg Val Arg Ile Arg
 515 520 525

tac gct tct acc aca aac ctc cag ttc cat aca tct att gac ggc aga 1632
 Tyr Ala Ser Thr Thr Asn Leu Gln Phe His Thr Ser Ile Asp Gly Arg
 530 535 540

ccc att aat caa ggg aat ttc tcc gcc acg atg tcg tcc ggc tcc aat 1680
 Pro Ile Asn Gln Gly Asn Phe Ser Ala Thr Met Ser Ser Gly Ser Asn
 545 550 555 560

ctc cag tcc gga agt ttc cgc acc gta ggt ttt act acc ccg ttc aac 1728
 Leu Gln Ser Gly Ser Phe Arg Thr Val Gly Phe Thr Thr Pro Phe Asn
 565 570 575

ttt tca aac ggc tca agt gtg ttt acg ctg tcc gct cat gtg ttc aac 1776
 Phe Ser Asn Gly Ser Ser Val Phe Thr Leu Ser Ala His Val Phe Asn
 580 585 590

tct ggc aat gag gtt tat atc gac cgg att gag ttc gtc ccg gca gaa 1824
 Ser Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala Glu
 595 600 605

gtc acc 1830
 Val Thr
 610

<210> 22
 <211> 610
 <212> PRT
 <213> Искусственная последовательность

<220>

<223> Синтетическая конструкция

<400> 22

Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
 1 5 10 15

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

Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
 35 40 45

Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
 50 55 60

Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
 65 70 75 80

Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
 85 90

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu
 100 105 110

Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu
 115 120 125

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala
 130 135 140

Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val
 145 150 155 160

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser
 165 170 175

Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg
 180 185 190

Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val
 195 200 205

Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg
 210 215 220

Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val
 225 230 235 240

40

Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro
 245 250 255

Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val
 260 265 270

Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu
 275 280 285

Arg Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr
 290 295 300

Ile Tyr Thr Asp Ala His Arg Gly Tyr Tyr Trp Ser Gly His Gln
 305 310 315 320

Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro
 325 330 335

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
 340 345 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
 355 360 365

Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp
 370 375 380

Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val
 385 390 395 400

Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln
 405 410 415

Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His
 420 425 430

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
 435 440 445

Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn
 450 455 460

Ile Ile Pro Ser Ser Gln Ile Thr Gln Ile Pro Leu Thr Lys Ser Thr
 465 470 475 480

Asn Leu Gly Ser Gly Thr Ser Val Val Lys Gly Pro Gly Phe Thr Gly
 485 490 495

41

Gly Asp Ile Leu Arg Arg Thr Ser Pro Gly Gln Ile Ser Thr Leu Arg
500 505 510

Val Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg Val Arg Ile Arg
515 520 525

Tyr Ala Ser Thr Thr Asn Leu Gln Phe His Thr Ser Ile Asp Gly Arg
530 535 540

Pro Ile Asn Gln Gly Asn Phe Ser Ala Thr Met Ser Ser Gly Ser Asn
545 550 555 560

Leu Gln Ser Gly Ser Phe Arg Thr Val Gly Phe Thr Thr Pro Phe Asn
565 570 575

Phe Ser Asn Gly Ser Ser Val Phe Thr Leu Ser Ala His Val Phe Asn
580 585 590

Ser Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala Glu
595 600 605

Val Thr
610

<210> 23
<211> 1830
<212> ДНК
<213> Искусственная последовательность

<220>
<223> Синтетическая последовательность ДНК по изобретению, кодирующая ядерный токсин CryIAb1 *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой удаляются последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1

<220>
<221> CDS
<222> (1)..(1830)

<400> 23
atg gat aac aac ccg aac atc aat gag tgc atc ccg tat aac tgt ctc 48
Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
1 5 10 15

agt aac cct gaa gtg gag gtc tta ggt ggc gaa cgc atc gaa act ggt 96
Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly
20 25 30

tac acc cca atc gac att agc ttg tgc ttg acg cag ttc ctc ttg tcc 144
Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
35 40 45

gag ttc gtg ccc ggt gcg ggt ttc gtg ctg ggg cta gtt gat ata atc 192
Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
50 55 60

tgg gga atc ttt ggt ccc tct cag tgg gac gcc ttt ctt gtg caa att	240
Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile	
65 70 75 80	
gag cag cta att aac caa aga ata gaa gag ttc gcg agg aac caa gcc	288
Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala	
85 90 95	
att tcc aga ctg gag gga cta agc aac ctt tat caa atc tac gcg gag	336
Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu	
100 105 110	
tct ttt agg gag tgg gag gca gat cct acg aac ccg gca ctg cgc gaa	384
Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu	
115 120 125	
gag atg cgt att cag ttc aac gac atg aac agt gcc ctt aca acc gct	432
Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala	
130 135 140	
att ccc ctt ttc gca gtt caa aat tac caa gtt ccc ctt ctc tca gtg	480
Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val	
145 150 155 160	
tac gtt caa gcc gca aat tta cac cta agc gtt ctc cgc gat gtg tca	528
Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser	
165 170 175	
gtg ttc ggc cag agg tgg gga ttt gat gcc gcc act atc aat agt cgt	576
Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg	
180 185 190	
tat aat gat ctg acg agg ctt atc ggc aac tat acc gac tat gct gtc	624
Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val	
195 200 205	
cgc tgg tac aat acg gga tta gag cgg gtc tgg ggt ccg gat tcc cga	672
Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg	
210 215 220	
gac tgg gtg cgc tac aat caa ttc cgc cgc gaa tta acc ctc act gtc	720
Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val	
225 230 235 240	
ctc gac atc gtg gcg ctg ttc ccg aac tac gac agt agg aga tac cca	768
Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro	
245 250 255	
atc cgc aca gtt tcc caa tta acg cgg gaa att tac acc aac cca gtc	816
Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val	
260 265 270	
ctg gag aat ttt gac ggg agc ttc cga ggc tcg gct caa ggc ata gaa	864
Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu	
275 280 285	
cgc agc att agg tcg cca cac ttg atg gat atc ctt aac agc atc acc	912
Arg Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr	
290 295 300	
atc tac acg gat gcc cat agg ggt tac tac tac tgg tcg ggg cat caa	960
Ile Tyr Thr Asp Ala His Arg Gly Tyr Tyr Trp Ser Gly His Gln	
305 310 315 320	

ata atg gct tct cct gtc ggg ttt tcg ggg cca gag ttc acc ttc ccg Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro 325 330 335	1008
ctc tac ggc act atg gga aat gcc gcg cca caa caa cgt atc gtc gct Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala 340 345 350	1056
caa cta ggt caa ggc gtg tac cgg aca ctg tcg tcc act ctc tat cgg Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg 355 360 365	1104
cgg cct ttc aat ata ggg atc aat aat caa cag ttg tct gtg ctg gac Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp 370 375 380	1152
ggg aca gag ttt gct tac gga acc tca agc aac ttg cca tcc gct gta Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val 385 390 395 400	1200
tac aga aaa agc ggc acg gtg gac tcg ctg gat gaa atc ccg ccc cag Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln 405 410 415	1248
aat aac aac gtg ccc cct cgg caa gcc ttc agt cat cga ctg agc cac Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His 420 425 430	1296
gtt agc atg ttc cgt tcg ggc ttc agc aac tcc tcc gta agt atc ata Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile 435 440 445	1344
aga gca cct atg ttc agc tgg ata cat cgt tcc gcc gag ttc aat aat Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn 450 455 460	1392
ata att ccc tcc tct caa atc aca cag atc cct ctg aca aag tct act Ile Ile Pro Ser Ser Gln Ile Thr Gln Ile Pro Leu Thr Lys Ser Thr 465 470 475 480	1440
aat ctt ggc tct ggg act tct gtc gtt aag ggg cct ggc ttt acg ggc Asn Leu Gly Ser Gly Thr Ser Val Val Lys Gly Pro Gly Phe Thr Gly 485 490 495	1488
ggc gat att ctg cgg aga act tca cct ggc cag att tcc acc ctg cgc Gly Asp Ile Leu Arg Arg Thr Ser Pro Gly Gln Ile Ser Thr Leu Arg 500 505 510	1536
gtg aat atc acc gcg cca ttg tca caa cgt tac cgc gtg cgg att cgc Val Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg Val Arg Ile Arg 515 520 525	1584
tac gct tct acc aca aac ctc cag ttc cat aca tct att gac ggc aga Tyr Ala Ser Thr Thr Asn Leu Gln Phe His Thr Ser Ile Asp Gly Arg 530 535 540	1632
ccc att aat caa ggg aat ttc tcc gcc acg atg tcg tcc ggc tcc aat Pro Ile Asn Gln Gly Asn Phe Ser Ala Thr Met Ser Ser Gly Ser Asn 545 550 555 560	1680
ctc cag tcc gga agt ttc cgc acc gta ggt ttt act acc ccg ttc aac Leu Gln Ser Gly Ser Phe Arg Thr Val Gly Phe Thr Thr Pro Phe Asn 565 570 575	1728

ttt tca aac ggc tca agt gtg ttt acg ctg tcc gct cat gtg ttc aac 1776
 Phe Ser Asn Gly Ser Ser Val Phe Thr Leu Ser Ala His Val Phe Asn
 580 585 590

tct ggc aat gag gtt tac atc gac cgg att gag ttc gtc ccg gca gaa 1824
 Ser Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala Glu
 595 600 605

gtc acc 1830
 Val Thr
 610

<210> 24
 <211> 610
 <212> PRT
 <213> Искусственная последовательность

<220>
 <223> Синтетическая конструкция

<400> 24

Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
 1 5 10 15

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

Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
 35 40 45

Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
 50 55 60

Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
 65 70 75 80

Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
 85 90 95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu
 100 105 110

Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu
 115 120 125

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala
 130 135 140

Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val
 145 150 155 160

45

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser
 165 170 175

Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg
 180 185 190

Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val
 195 200 205

Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg
 210 215 220

Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val
 225 230 235 240

Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro
 245 250 255

Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val
 260 265 270

Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu
 275 280 285

Arg Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr
 290 295 300

Ile Tyr Thr Asp Ala His Arg Gly Tyr Tyr Tyr Trp Ser Gly His Gln
 305 310 315 320

Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro
 325 330 335

Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
 340 345 350

Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
 355 360 365

Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp
 370 375 380

Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val
 385 390 395 400

Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln
 405 410 415

46

Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His
 420 425 430

Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
 435 440 445

Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn
 450 455 460

Ile Ile Pro Ser Ser Gln Ile Thr Gln Ile Pro Leu Thr Lys Ser Thr
 465 470 475 480

Asn Leu Gly Ser Gly Thr Ser Val Val Lys Gly Pro Gly Phe Thr Gly
 485 490 495

Gly Asp Ile Leu Arg Arg Thr Ser Pro Gly Gln Ile Ser Thr Leu Arg
 500 505 510

Val Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg Val Arg Ile Arg
 515 520 525

Tyr Ala Ser Thr Thr Asn Leu Gln Phe His Thr Ser Ile Asp Gly Arg
 530 535 540

Pro Ile Asn Gln Gly Asn Phe Ser Ala Thr Met Ser Ser Gly Ser Asn
 545 550 555 560

Leu Gln Ser Gly Ser Phe Arg Thr Val Gly Phe Thr Thr Pro Phe Asn
 565 570 575

Phe Ser Asn Gly Ser Ser Val Phe Thr Leu Ser Ala His Val Phe Asn
 580 585 590

Ser Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala Glu
 595 600 605

Val Thr
 610

<210> 25
 <211> 1866
 <212> ДНК
 <213> *Bacillus thuringiensis*

<220>
 <221> CDS
 <222> (1)..(1866)
 <223> Нативная последовательность ДНК, кодирующая ядерный токсин Cry1Ca
Bacillus thuringiensis

```

<400> 25
atg gat aac aat ccg aac atc aat gaa tgc atc ccg tac aac tgc ctg      48
Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
1          5          10          15

agc aac ccg gaa gaa gtg ctg ttg gat gga gaa cgg ata tca act ggt      96
Ser Asn Pro Glu Glu Val Leu Leu Asp Gly Glu Arg Ile Ser Thr Gly
          20          25          30

aat tca tca att gat att tct ctg tca ctt gtt cag ttt ctg gta tct      144
Asn Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Phe Leu Val Ser
          35          40          45

aac ttc gtc cca ggc gga gga ttc ctg gtt gga tta ata gat ttt gta      192
Asn Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val
          50          55          60

tgg gga ata gtt ggc cct tct caa tgg gat gca ttt cta gta caa att      240
Trp Gly Ile Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
65          70          75          80

gaa caa tta att aat gaa aga ata gct gaa ttt gct agg aat gct gct      288
Glu Gln Leu Ile Asn Glu Arg Ile Ala Glu Phe Ala Arg Asn Ala Ala
          85          90          95

att gct aat tta gaa gga tta gga aac aat ttc aat ata tat gtg gaa      336
Ile Ala Asn Leu Glu Gly Leu Gly Asn Asn Phe Asn Ile Tyr Val Glu
          100          105          110

gca ttt aaa gaa tgg gaa gaa gat cct aag aat cca gca acc agg acc      384
Ala Phe Lys Glu Trp Glu Glu Asp Pro Lys Asn Pro Ala Thr Arg Thr
          115          120          125

aga gta att gat cgc ttt cgt ata ctt gat ggg cta ctt gaa agg gac      432
Arg Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp
          130          135          140

att cct tcg ttt cga att tct gga ttt gaa gta ccc ctt tta tcc gtt      480
Ile Pro Ser Phe Arg Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val
145          150          155          160

tat gct caa gcg gcc aat ctg cat cta gct ata tta aga gat tct gta      528
Tyr Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val
          165          170          175

att ttt gga gaa aga tgg gga ttg aca acg ata aat gtc aat gaa aac      576
Ile Phe Gly Glu Arg Trp Gly Leu Thr Thr Ile Asn Val Asn Glu Asn
          180          185          190

tat aat aga cta att agg cat att gat gaa tat gct gat cac tgt gca      624
Tyr Asn Arg Leu Ile Arg His Ile Asp Glu Tyr Ala Asp His Cys Ala
          195          200          205

aat acg tat aat cgg gga tta aat aat tta ccg aaa tct acg tat caa      672
Asn Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln
          210          215          220

gat tgg ata aca tat aat cga ttg cgg aga gac tta aca ttg act gta      720
Asp Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp Leu Thr Leu Thr Val
225          230          235          240

tta gat atc gcc gct ttc ttt cca aac tat gac aat agg aga tat cca      768
Leu Asp Ile Ala Ala Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro
          245          250          255
    
```

att cag cca gtt ggt caa cta aca agg gaa gtt tat acg gac cca tta Ile Gln Pro Val Gly Gln Leu Thr Arg Glu Val Tyr Thr Asp Pro Leu 260 265 270	816
att aat ttt aat cca cag tta cag tct gta gct caa tta cct act ttt Ile Asn Phe Asn Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe 275 280 285	864
aac gtt atg gag aac agc gca att aga aat cct cat tta ttt gat ata Asn Val Met Glu Asn Ser Ala Ile Arg Asn Pro His Leu Phe Asp Ile 290 295 300	912
ttg aat aat ctt aca atc ttt acg gat tgg ttt agt gtt gga cgc aat Leu Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe Ser Val Gly Arg Asn 305 310 315 320	960
ttt tat tgg gga gga cat cga gta ata tct agc ctt ata gga ggt ggt Phe Tyr Trp Gly Gly His Arg Val Ile Ser Ser Leu Ile Gly Gly Gly 325 330 335	1008
aac ata aca tct cct ata tat gga aga gag gcg aac cag gag cct cca Asn Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala Asn Gln Glu Pro Pro 340 345 350	1056
aga tcc ttt act ttt aat gga ccg gta ttt agg act tta tca aat cct Arg Ser Phe Thr Phe Asn Gly Pro Val Phe Arg Thr Leu Ser Asn Pro 355 360 365	1104
act tta cga tta tta cag caa cct tgg cca gcg cca cca ttt aat tta Thr Leu Arg Leu Leu Gln Gln Pro Trp Pro Ala Pro Pro Phe Asn Leu 370 375 380	1152
cgt ggt gtt gaa gga gta gaa ttt tct aca cct aca aat agc ttt acg Arg Gly Val Glu Gly Val Glu Phe Ser Thr Pro Thr Asn Ser Phe Thr 385 390 395 400	1200
tat cga gga aga ggt acg gtt gat tct tta act gaa ttg ccg cct gag Tyr Arg Gly Arg Gly Thr Val Asp Ser Leu Thr Glu Leu Pro Pro Glu 405 410 415	1248
gat aat agt gtg cca cct cgc gaa gga tat agt cat cgt tta tgt cat Asp Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser His Arg Leu Cys His 420 425 430	1296
gca act ttt gtt caa aga tct gga aca cct ttt tta aca act ggt gta Ala Thr Phe Val Gln Arg Ser Gly Thr Pro Phe Leu Thr Thr Gly Val 435 440 445	1344
gta ttt tct tgg acg cat cgt agt gca act ctt aca aat aca att gat Val Phe Ser Trp Thr His Arg Ser Ala Thr Leu Thr Asn Thr Ile Asp 450 455 460	1392
cca gag aga att aat caa ata cct tta gtg aaa gga ttt aga gtt tgg Pro Glu Arg Ile Asn Gln Ile Pro Leu Val Lys Gly Phe Arg Val Trp 465 470 475 480	1440
ggg ggc acc tct gtc att aca gga cca gga ttt aca gga ggg gat atc Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly Gly Asp Ile 485 490 495	1488
ctt cga aga aat acc ttt ggt gat ttt gta tct cta caa gtc aat att Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln Val Asn Ile 500 505 510	1536

aat tca cca att acc caa aga tac cgt tta aga ttt cgt tac gct tcc 1584
 Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg Tyr Ala Ser
 515 520 525

agt agg gat gca cga gtt ata gta tta aca gga gcg gca tcc aca gga 1632
 Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala Ser Thr Gly
 530 535 540

gtg gga ggc caa gtt agt gta aat atg cct ctt cag aaa act atg gaa 1680
 Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys Thr Met Glu
 545 550 555 560

ata ggg gag aac tta aca tct aga aca ttt aga tat acc gat ttt agt 1728
 Ile Gly Glu Asn Leu Thr Ser Arg Thr Phe Arg Tyr Thr Asp Phe Ser
 565 570 575

aat cct ttt tca ttt aga gct aat cca gat ata att ggg ata agt gaa 1776
 Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly Ile Ser Glu
 580 585 590

caa cct cta ttt ggt gca ggt tct att agt agc ggt gaa ctt tat ata 1824
 Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu Leu Tyr Ile
 595 600 605

gat aaa att gaa att att cta gca gat gca aca ttt gaa taa 1866
 Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu
 610 615 620

<210> 26
 <211> 621
 <212> PRT
 <213> Bacillus thuringiensis

<400> 26

Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
 1 5 10 15

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

Asn Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Phe Leu Val Ser
 35 40 45

Asn Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val
 50 55 60

Trp Gly Ile Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
 65 70 75 80

Glu Gln Leu Ile Asn Glu Arg Ile Ala Glu Phe Ala Arg Asn Ala Ala
 85 90 95

Ile Ala Asn Leu Glu Gly Leu Gly Asn Asn Phe Asn Ile Tyr Val Glu
 100 105 110

Ala Phe Lys Glu Trp Glu Glu Asp Pro Lys Asn Pro Ala Thr Arg Thr
 115 120 125

Arg Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp
 130 135 140

Ile Pro Ser Phe Arg Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val
 145 150 155 160

Tyr Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val
 165 170 175

Ile Phe Gly Glu Arg Trp Gly Leu Thr Thr Ile Asn Val Asn Glu Asn
 180 185 190

Tyr Asn Arg Leu Ile Arg His Ile Asp Glu Tyr Ala Asp His Cys Ala
 195 200 205

Asn Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln
 210 215 220

Asp Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp Leu Thr Leu Thr Val
 225 230 235 240

Leu Asp Ile Ala Ala Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro
 245 250 255

Ile Gln Pro Val Gly Gln Leu Thr Arg Glu Val Tyr Thr Asp Pro Leu
 260 265 270

Ile Asn Phe Asn Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe
 275 280 285

Asn Val Met Glu Asn Ser Ala Ile Arg Asn Pro His Leu Phe Asp Ile
 290 295 300

Leu Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe Ser Val Gly Arg Asn
 305 310 315 320

Phe Tyr Trp Gly Gly His Arg Val Ile Ser Ser Leu Ile Gly Gly Gly
 325 330 335

Asn Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala Asn Gln Glu Pro Pro
 340 345 350

Arg Ser Phe Thr Phe Asn Gly Pro Val Phe Arg Thr Leu Ser Asn Pro
 355 360 365

Thr Leu Arg Leu Leu Gln Gln Pro Trp Pro Ala Pro Pro Phe Asn Leu
 370 375 380

Arg Gly Val Glu Gly Val Glu Phe Ser Thr Pro Thr Asn Ser Phe Thr
 385 390 395 400

Tyr Arg Gly Arg Gly Thr Val Asp Ser Leu Thr Glu Leu Pro Pro Glu
 405 410 415

Asp Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser His Arg Leu Cys His
 420 425 430

Ala Thr Phe Val Gln Arg Ser Gly Thr Pro Phe Leu Thr Thr Gly Val
 435 440 445

Val Phe Ser Trp Thr His Arg Ser Ala Thr Leu Thr Asn Thr Ile Asp
 450 455 460

Pro Glu Arg Ile Asn Gln Ile Pro Leu Val Lys Gly Phe Arg Val Trp
 465 470 475 480

Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly Gly Asp Ile
 485 490 495

Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln Val Asn Ile
 500 505 510

Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg Tyr Ala Ser
 515 520 525

Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala Ser Thr Gly
 530 535 540

Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys Thr Met Glu
 545 550 555 560

Ile Gly Glu Asn Leu Thr Ser Arg Thr Phe Arg Tyr Thr Asp Phe Ser
 565 570 575

Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly Ile Ser Glu
 580 585 590

Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu Leu Tyr Ile
 595 600 605

Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu
 610 615 620

<210> 27
 <211> 1866
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Синтетическая последовательность ДНК, кодирующая ядерный токсин Cry1Ca *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1

<220>
 <221> CDS
 <222> (1)..(1866)

<400> 27

atg gat aac aat ccg aac atc aat gag tgc atc ccg tac aac tgc ctg	48
Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu	
1 5 10 15	
agc aac ccg gaa gaa gtg ctg ttg gat gga gaa cgg ata tca act ggc	96
Ser Asn Pro Glu Glu Val Leu Leu Asp Gly Glu Arg Ile Ser Thr Gly	
20 25 30	
aat tca tcc att gat att tct ctg tca ctt gtt cag ttt ctg gtg tct	144
Asn Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Phe Leu Val Ser	
35 40 45	
aac ttc gtc ccc ggc gga gga ttc ctg gtt gga tta ata gat ttt gta	192
Asn Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val	
50 55 60	
tgg gga ata gtt ggc cct tct caa tgg gac gca ttt cta gta caa att	240
Trp Gly Ile Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile	
65 70 75 80	
gaa caa tta att aat gaa aga ata gct gaa ttt gct agg aac gct gct	288
Glu Gln Leu Ile Asn Glu Arg Ile Ala Glu Phe Ala Arg Asn Ala Ala	
85 90 95	
att gct aat tta gaa gga tta gga aac aat ttc aat ata tat gtg gaa	336
Ile Ala Asn Leu Glu Gly Leu Gly Asn Asn Phe Asn Ile Tyr Val Glu	
100 105 110	
gca ttt aag gaa tgg gaa gaa gat cct aag aat cca gca acg agg acc	384
Ala Phe Lys Glu Trp Glu Glu Asp Pro Lys Asn Pro Ala Thr Arg Thr	
115 120 125	
aga gta att gat cgc ttt cgt ata ctt gat ggg cta ctt gaa agg gac	432
Arg Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp	
130 135 140	
att cct tcg ttt cga att tct gga ttt gaa gta ccc ctt tta tcc gtt	480
Ile Pro Ser Phe Arg Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val	
145 150 155 160	
tat gct caa gcg gcc aat ctg cat cta gct ata tta aga gat tct gta	528
Tyr Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val	
165 170 175	
att ttt gga gaa aga tgg gga ttg aca acg ata aat gtc aat gaa aac	576
Ile Phe Gly Glu Arg Trp Gly Leu Thr Thr Ile Asn Val Asn Glu Asn	
180 185 190	

tat aat aga cta att agg cat att gat gaa tat gct gat cac tgt gca	624
Tyr Asn Arg Leu Ile Arg His Ile Asp Glu Tyr Ala Asp His Cys Ala	
195 200 205	
aat acg tat aat cgg gga tta aat aat tta ccg aaa tct acg tat caa	672
Asn Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln	
210 215 220	
gat tgg ata aca tat aat cga ttg cgg aga gac tta aca ttg act gta	720
Asp Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp Leu Thr Leu Thr Val	
225 230 235 240	
tta gat atc gcc gct ttc ttt cca aac tat gac aat agg aga tat cca	768
Leu Asp Ile Ala Ala Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro	
245 250 255	
att cag cca gtt ggt caa cta aca agg gaa gtt tat acg gac cca tta	816
Ile Gln Pro Val Gly Gln Leu Thr Arg Glu Val Tyr Thr Asp Pro Leu	
260 265 270	
att aat ttt aat cca cag tta cag tct gta gct caa tta cct act ttt	864
Ile Asn Phe Asn Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe	
275 280 285	
aac gtt atg gag aac agc gca att aga aat cct cat tta ttt gat ata	912
Asn Val Met Glu Asn Ser Ala Ile Arg Asn Pro His Leu Phe Asp Ile	
290 295 300	
ttg aat aat ctt aca atc ttt acg gat tgg ttt agt gtt gga cgc aat	960
Leu Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe Ser Val Gly Arg Asn	
305 310 315 320	
ttt tat tgg gga gga cat cga gta ata tct agc ctt ata gga ggt ggg	1008
Phe Tyr Trp Gly Gly His Arg Val Ile Ser Ser Leu Ile Gly Gly Gly	
325 330 335	
aac atc aca tcg cct ata tat gga aga gag gcg aac caa gag cct cca	1056
Asn Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala Asn Gln Glu Pro Pro	
340 345 350	
aga tcc ttt act ttt aat gga ccc gtg ttt agg act tta tca aat cct	1104
Arg Ser Phe Thr Phe Asn Gly Pro Val Phe Arg Thr Leu Ser Asn Pro	
355 360 365	
act tta cga tta tta cag caa cct tgg cca gcg cca cca ttt aat tta	1152
Thr Leu Arg Leu Leu Gln Gln Pro Trp Pro Ala Pro Pro Phe Asn Leu	
370 375 380	
cgt ggt gtt gaa gga gta gaa ttt tct aca cct aca aat agc ttt acg	1200
Arg Gly Val Glu Gly Val Glu Phe Ser Thr Pro Thr Asn Ser Phe Thr	
385 390 395 400	
tat cga gga aga ggg acg gtt gat tct tta act gaa ttg ccg cct gag	1248
Tyr Arg Gly Arg Gly Thr Val Asp Ser Leu Thr Glu Leu Pro Pro Glu	
405 410 415	
gat aat agt gtg cca cct cgc gaa gga tat agt cat cgt tta tgt cat	1296
Asp Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser His Arg Leu Cys His	
420 425 430	
gca act ttt gtt caa aga tcg gga aca cct ttt tta aca act ggt gta	1344
Ala Thr Phe Val Gln Arg Ser Gly Thr Pro Phe Leu Thr Thr Gly Val	
435 440 445	

gta ttt tct tgg acg cat cgt agt gca act ctt aca aat aca atc gac 1392
 Val Phe Ser Trp Thr His Arg Ser Ala Thr Leu Thr Asn Thr Ile Asp
 450 455 460

cca gag aga att aat caa ata cct tta gtg aag gga ttt aga gtt tgg 1440
 Pro Glu Arg Ile Asn Gln Ile Pro Leu Val Lys Gly Phe Arg Val Trp
 465 470 475 480

ggg ggc acc tct gtc att acc gga ccc gga ttt acc gga ggg gat atc 1488
 Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly Gly Asp Ile
 485 490 495

ctt cga aga aat acc ttt ggt gat ttt gta tct cta caa gtc aac att 1536
 Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln Val Asn Ile
 500 505 510

aat tca cca att acc caa aga tac cgt tta aga ttt cgt tac gct tcc 1584
 Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg Tyr Ala Ser
 515 520 525

agt agg gat gca cga gtt ata gta tta acg gga gcg gca tcc acc gga 1632
 Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala Ser Thr Gly
 530 535 540

gtg gga ggc caa gtt agt gta aat atg cct ctt cag aaa act atg gaa 1680
 Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys Thr Met Glu
 545 550 555 560

ata ggg gag aac tta aca tcc aga aca ttt aga tat acc gat ttt agt 1728
 Ile Gly Glu Asn Leu Thr Ser Arg Thr Phe Arg Tyr Thr Asp Phe Ser
 565 570 575

aat cct ttt tca ttt aga gct aat cca gat ata att ggg ata agt gaa 1776
 Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly Ile Ser Glu
 580 585 590

caa cct cta ttt ggg gcg ggt tct att agt agc ggt gaa ctt tat ata 1824
 Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu Leu Tyr Ile
 595 600 605

gat aaa att gaa att att cta gca gat gca aca ttt gaa tga 1866
 Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu
 610 615 620

<210> 28
 <211> 621
 <212> PRT
 <213> Искусственная последовательность

<220>
 <223> Синтетическая конструкция

<400> 28

Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
 1 5 10 15

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

55

Asn Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Phe Leu Val Ser
 35 40 45

Asn Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val
 50 55 60

Trp Gly Ile Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
 65 70 75 80

Glu Gln Leu Ile Asn Glu Arg Ile Ala Glu Phe Ala Arg Asn Ala Ala
 85 90 95

Ile Ala Asn Leu Glu Gly Leu Gly Asn Asn Phe Asn Ile Tyr Val Glu
 100 105 110

Ala Phe Lys Glu Trp Glu Glu Asp Pro Lys Asn Pro Ala Thr Arg Thr
 115 120 125

Arg Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp
 130 135 140

Ile Pro Ser Phe Arg Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val
 145 150 155 160

Tyr Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val
 165 170 175

Ile Phe Gly Glu Arg Trp Gly Leu Thr Thr Ile Asn Val Asn Glu Asn
 180 185 190

Tyr Asn Arg Leu Ile Arg His Ile Asp Glu Tyr Ala Asp His Cys Ala
 195 200 205

Asn Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln
 210 215 220

Asp Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp Leu Thr Leu Thr Val
 225 230 235 240

Leu Asp Ile Ala Ala Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro
 245 250 255

Ile Gln Pro Val Gly Gln Leu Thr Arg Glu Val Tyr Thr Asp Pro Leu
 260 265 270

Ile Asn Phe Asn Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe
 275 280 285

56

Asn Val Met Glu Asn Ser Ala Ile Arg Asn Pro His Leu Phe Asp Ile
 290 295 300

Leu Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe Ser Val Gly Arg Asn
 305 310 315 320

Phe Tyr Trp Gly Gly His Arg Val Ile Ser Ser Leu Ile Gly Gly Gly
 325 330 335

Asn Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala Asn Gln Glu Pro Pro
 340 345 350

Arg Ser Phe Thr Phe Asn Gly Pro Val Phe Arg Thr Leu Ser Asn Pro
 355 360 365

Thr Leu Arg Leu Leu Gln Gln Pro Trp Pro Ala Pro Pro Phe Asn Leu
 370 375 380

Arg Gly Val Glu Gly Val Glu Phe Ser Thr Pro Thr Asn Ser Phe Thr
 385 390 395 400

Tyr Arg Gly Arg Gly Thr Val Asp Ser Leu Thr Glu Leu Pro Pro Glu
 405 410 415

Asp Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser His Arg Leu Cys His
 420 425 430

Ala Thr Phe Val Gln Arg Ser Gly Thr Pro Phe Leu Thr Thr Gly Val
 435 440 445

Val Phe Ser Trp Thr His Arg Ser Ala Thr Leu Thr Asn Thr Ile Asp
 450 455 460

Pro Glu Arg Ile Asn Gln Ile Pro Leu Val Lys Gly Phe Arg Val Trp
 465 470 475 480

Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly Gly Asp Ile
 485 490 495

Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln Val Asn Ile
 500 505 510

Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg Tyr Ala Ser
 515 520 525

Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala Ser Thr Gly
 530 535 540

57

Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys Thr Met Glu
545 550 555 560

Ile Gly Glu Asn Leu Thr Ser Arg Thr Phe Arg Tyr Thr Asp Phe Ser
565 570 575

Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly Ile Ser Glu
580 585 590

Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu Leu Tyr Ile
595 600 605

Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu
610 615 620

<210> 29
<211> 1866
<212> ДНК
<213> Искусственная последовательность

<220>
<223> Синтетическая последовательность ДНК по изобретению, кодирующая ядерный токсин Cry1Ca *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой удаляются последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1

<220>
<221> CDS
<222> (1)..(1866)

<400> 29
atg gat aac aat ccg aac atc aat gag tgc atc ccg tac aac tgc ctg 48
Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
1 5 10 15
agc aac ccg gaa gaa gtg ctg ttg gat gga gaa cgg ata tca act ggc 96
Ser Asn Pro Glu Glu Val Leu Leu Asp Gly Glu Arg Ile Ser Thr Gly
20 25 30
aat tca tcc att gac att tct ctg tca ctt gtt cag ttt ctg gtg tct 144
Asn Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Phe Leu Val Ser
35 40 45
aac ttc gtc ccc ggc gga gga ttc ctg gtt gga tta ata gat ttc gta 192
Asn Phe Val Pro Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val
50 55 60
tgg gga ata gtt ggc cct tct caa tgg gac gca ttt cta gta caa att 240
Trp Gly Ile Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
65 70 75 80
gaa caa tta att aat gaa aga ata gct gaa ttt gct agg aac gct gct 288
Glu Gln Leu Ile Asn Glu Arg Ile Ala Glu Phe Ala Arg Asn Ala Ala
85 90 95
att gct aat tta gaa gga tta gga aac aat ttc aac atc tat gtg gaa 336
Ile Ala Asn Leu Glu Gly Leu Gly Asn Asn Phe Asn Ile Tyr Val Glu
100 105 110

gca ttt aag gaa tgg gaa gaa gat cct aag aat cca gca acg agg acc 384
 Ala Phe Lys Glu Trp Glu Glu Asp Pro Lys Asn Pro Ala Thr Arg Thr
 115 120 125

aga gta att gat cgc ttt cgt ata ctt gat ggg cta ctt gaa agg gac 432
 Arg Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp
 130 135 140

att cct tcg ttt cga att tct gga ttt gaa gta ccc ctt ctc tcc gtt 480
 Ile Pro Ser Phe Arg Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val
 145 150 155 160

tat gct caa gcg gcc aat ctg cat cta gct atc tta aga gat tct gtc 528
 Tyr Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val
 165 170 175

atc ttt gga gaa aga tgg gga ttg aca acg ata aat gtc aat gaa aac 576
 Ile Phe Gly Glu Arg Trp Gly Leu Thr Thr Ile Asn Val Asn Glu Asn
 180 185 190

tat aat aga cta att agg cat att gat gaa tat gct gat cac tgt gca 624
 Tyr Asn Arg Leu Ile Arg His Ile Asp Glu Tyr Ala Asp His Cys Ala
 195 200 205

aat acg tat aat cgg gga tta aat aat tta ccg aaa tct acg tat caa 672
 Asn Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln
 210 215 220

gat tgg ata aca tat aat cga ttg cgg aga gac tta aca ttg act gta 720
 Asp Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp Leu Thr Leu Thr Val
 225 230 235 240

tta gat atc gcc gct ttc ttt cca aac tat gac aat agg aga tat cca 768
 Leu Asp Ile Ala Ala Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro
 245 250 255

att cag cca gtt ggt caa cta aca agg gaa gtt tat acg gac cca tta 816
 Ile Gln Pro Val Gly Gln Leu Thr Arg Glu Val Tyr Thr Asp Pro Leu
 260 265 270

att aat ttt aat cca cag tta cag tct gta gct caa tta cct act ttt 864
 Ile Asn Phe Asn Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe
 275 280 285

aac gtt atg gag aac agc gca att aga aat cct cat ttg ttc gac ata 912
 Asn Val Met Glu Asn Ser Ala Ile Arg Asn Pro His Leu Phe Asp Ile
 290 295 300

ttg aat aat ctt aca atc ttt acg gat tgg ttt agt gtt gga cgc aac 960
 Leu Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe Ser Val Gly Arg Asn
 305 310 315 320

ttc tat tgg gga gga cat cga gta ata tct agc ctt ata gga ggt ggg 1008
 Phe Tyr Trp Gly Gly His Arg Val Ile Ser Ser Leu Ile Gly Gly Gly
 325 330 335

aac atc aca tcg cct atc tat gga aga gag gcg aac caa gag cct cca 1056
 Asn Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala Asn Gln Glu Pro Pro
 340 345 350

aga tcc ttt act ttt aat gga ccc gtg ttt agg act tta tca aat cct 1104
 Arg Ser Phe Thr Phe Asn Gly Pro Val Phe Arg Thr Leu Ser Asn Pro
 355 360 365

act tta cga tta tta cag caa cct tgg cca gcg cca cca ttt aat tta Thr Leu Arg Leu Leu Gln Gln Pro Trp Pro Ala Pro Pro Phe Asn Leu 370 375 380	1152
cgt ggt gtt gaa gga gta gaa ttt tct aca cct aca aat agc ttt acg Arg Gly Val Glu Gly Val Glu Phe Ser Thr Pro Thr Asn Ser Phe Thr 385 390 395 400	1200
tat cga gga aga ggg acg gtt gat tct tta act gaa ttg ccg cct gag Tyr Arg Gly Arg Gly Thr Val Asp Ser Leu Thr Glu Leu Pro Pro Glu 405 410 415	1248
gat aat agt gtg cca cct cgc gaa gga tat agt cat cgt tta tgt cat Asp Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser His Arg Leu Cys His 420 425 430	1296
gca acc ttt gtt caa aga tcg gga aca cct ttc tta aca act ggt gta Ala Thr Phe Val Gln Arg Ser Gly Thr Pro Phe Leu Thr Thr Gly Val 435 440 445	1344
gta ttc tct tgg acg cat cgt agt gca act ctt aca aat aca atc gac Val Phe Ser Trp Thr His Arg Ser Ala Thr Leu Thr Asn Thr Ile Asp 450 455 460	1392
cca gag aga att aat caa ata cct tta gtg aag gga ttt aga gtt tgg Pro Glu Arg Ile Asn Gln Ile Pro Leu Val Lys Gly Phe Arg Val Trp 465 470 475 480	1440
ggg ggc acc tct gtc att acc gga ccc gga ttt acc gga ggg gat atc Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly Gly Asp Ile 485 490 495	1488
ctt cga aga aat acc ttt ggt gat ttc gta tct cta caa gtc aac att Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln Val Asn Ile 500 505 510	1536
aat tca cca att acc caa aga tac cgt tta aga ttt cgt tac gct tcc Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg Tyr Ala Ser 515 520 525	1584
agt agg gat gca cga gtt ata gta tta acg gga gcg gca tcc acc gga Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala Ser Thr Gly 530 535 540	1632
gtg gga ggc caa gtt agt gta aat atg cct ctt cag aaa act atg gaa Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys Thr Met Glu 545 550 555 560	1680
ata ggg gag aac tta aca tcc aga aca ttt aga tat acc gat ttt agt Ile Gly Glu Asn Leu Thr Ser Arg Thr Phe Arg Tyr Thr Asp Phe Ser 565 570 575	1728
aat cct ttt tca ttt aga gct aat cca gat ata att ggg ata agt gaa Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly Ile Ser Glu 580 585 590	1776
caa cct cta ttt ggg gcg ggt tct att agt agc ggt gaa ctt tac ata Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu Leu Tyr Ile 595 600 605	1824
gat aaa att gaa att att cta gca gat gca aca ttt gaa tga Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu 610 615 620	1866

60

<210> 30
 <211> 621
 <212> PRT
 <213> Искусственная последовательность

 <220>
 <223> Синтетическая конструкция

 <400> 30

 Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
 1 5 10 15

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

 Asn Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Phe Leu Val Ser
 35 40 45

 Asn Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val
 50 55 60

 Trp Gly Ile Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile
 65 70 75 80

 Glu Gln Leu Ile Asn Glu Arg Ile Ala Glu Phe Ala Arg Asn Ala Ala
 85 90 95

 Ile Ala Asn Leu Glu Gly Leu Gly Asn Asn Phe Asn Ile Tyr Val Glu
 100 105 110

 Ala Phe Lys Glu Trp Glu Glu Asp Pro Lys Asn Pro Ala Thr Arg Thr
 115 120 125

 Arg Val Ile Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp
 130 135 140

 Ile Pro Ser Phe Arg Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val
 145 150 155 160

 Tyr Ala Gln Ala Ala Asn Leu His Leu Ala Ile Leu Arg Asp Ser Val
 165 170 175

 Ile Phe Gly Glu Arg Trp Gly Leu Thr Thr Ile Asn Val Asn Glu Asn
 180 185 190

 Tyr Asn Arg Leu Ile Arg His Ile Asp Glu Tyr Ala Asp His Cys Ala
 195 200 205

61

Asn Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln
 210 215 220

Asp Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp Leu Thr Leu Thr Val
 225 230 235 240

Leu Asp Ile Ala Ala Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro
 245 250 255

Ile Gln Pro Val Gly Gln Leu Thr Arg Glu Val Tyr Thr Asp Pro Leu
 260 265 270

Ile Asn Phe Asn Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe
 275 280 285

Asn Val Met Glu Asn Ser Ala Ile Arg Asn Pro His Leu Phe Asp Ile
 290 295 300

Leu Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe Ser Val Gly Arg Asn
 305 310 315 320

Phe Tyr Trp Gly Gly His Arg Val Ile Ser Ser Leu Ile Gly Gly Gly
 325 330 335

Asn Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala Asn Gln Glu Pro Pro
 340 345 350

Arg Ser Phe Thr Phe Asn Gly Pro Val Phe Arg Thr Leu Ser Asn Pro
 355 360 365

Thr Leu Arg Leu Leu Gln Gln Pro Trp Pro Ala Pro Pro Phe Asn Leu
 370 375 380

Arg Gly Val Glu Gly Val Glu Phe Ser Thr Pro Thr Asn Ser Phe Thr
 385 390 395 400

Tyr Arg Gly Arg Gly Thr Val Asp Ser Leu Thr Glu Leu Pro Pro Glu
 405 410 415

Asp Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser His Arg Leu Cys His
 420 425 430

Ala Thr Phe Val Gln Arg Ser Gly Thr Pro Phe Leu Thr Thr Gly Val
 435 440 445

Val Phe Ser Trp Thr His Arg Ser Ala Thr Leu Thr Asn Thr Ile Asp
 450 455 460

62

Pro Glu Arg Ile Asn Gln Ile Pro Leu Val Lys Gly Phe Arg Val Trp
465 470 475 480

Gly Gly Thr Ser Val Ile Thr Gly Pro Gly Phe Thr Gly Gly Asp Ile
485 490 495

Leu Arg Arg Asn Thr Phe Gly Asp Phe Val Ser Leu Gln Val Asn Ile
500 505 510

Asn Ser Pro Ile Thr Gln Arg Tyr Arg Leu Arg Phe Arg Tyr Ala Ser
515 520 525

Ser Arg Asp Ala Arg Val Ile Val Leu Thr Gly Ala Ala Ser Thr Gly
530 535 540

Val Gly Gly Gln Val Ser Val Asn Met Pro Leu Gln Lys Thr Met Glu
545 550 555 560

Ile Gly Glu Asn Leu Thr Ser Arg Thr Phe Arg Tyr Thr Asp Phe Ser
565 570 575

Asn Pro Phe Ser Phe Arg Ala Asn Pro Asp Ile Ile Gly Ile Ser Glu
580 585 590

Gln Pro Leu Phe Gly Ala Gly Ser Ile Ser Ser Gly Glu Leu Tyr Ile
595 600 605

Asp Lys Ile Glu Ile Ile Leu Ala Asp Ala Thr Phe Glu
610 615 620

<210> 31
<211> 1428
<212> ДНК
<213> *Bacillus thuringiensis*

<220>
<221> CDS
<222> (1)..(1428)
<223> Нативная последовательность ДНК, кодирующая токсин Cry6Aa *Bacillus thuringiensis*

<400> 31
atg att att gat agt aaa acg act tta cct aga cat tca ctt att cat 48
Met Ile Ile Asp Ser Lys Thr Thr Leu Pro Arg His Ser Leu Ile His
1 5 10 15
aca att aaa tta aat tct aat aag aaa tat ggt cct ggt gat atg act 96
Thr Ile Lys Leu Asn Ser Asn Lys Lys Tyr Gly Pro Gly Asp Met Thr
20 25 30
aat gga aat caa ttt att att tca aaa caa gaa tgg gct acg att gga 144
Asn Gly Asn Gln Phe Ile Ile Ser Lys Gln Glu Trp Ala Thr Ile Gly
35 40 45

gca tat att cag act gga tta ggt tta cca gta aat gaa caa caa tta Ala Tyr Ile Gln Thr Gly Leu Gly Leu Pro Val Asn Glu Gln Gln Leu 50 55 60	192
aga aca cat gtt aat tta agt cag gat ata tca ata cct agt gat ttt Arg Thr His Val Asn Leu Ser Gln Asp Ile Ser Ile Pro Ser Asp Phe 65 70 75 80	240
tct caa tta tat gat gtt tat tgt tct gat aaa act tca gca gaa tgg Ser Gln Leu Tyr Asp Val Tyr Cys Ser Asp Lys Thr Ser Ala Glu Trp 85 90 95	288
tgg aat aaa aat tta tat cct tta att att aaa tct gct aat gat att Trp Asn Lys Asn Leu Tyr Pro Leu Ile Ile Lys Ser Ala Asn Asp Ile 100 105 110	336
gct tca tat ggt ttt aaa gtt gct ggt gat cct tct att aag aaa gat Ala Ser Tyr Gly Phe Lys Val Ala Gly Asp Pro Ser Ile Lys Lys Asp 115 120 125	384
gga tat ttt aaa aaa ttg caa gat gaa tta gat aat att gtt gat aat Gly Tyr Phe Lys Lys Leu Gln Asp Glu Leu Asp Asn Ile Val Asp Asn 130 135 140	432
aat tcc gat gat gat gca ata gct aaa gct att aaa gat ttt aaa gcg Asn Ser Asp Asp Asp Ala Ile Ala Lys Ala Ile Lys Asp Phe Lys Ala 145 150 155 160	480
cga tgt ggt att tta att aaa gaa gct aaa caa tat gaa gaa gct gca Arg Cys Gly Ile Leu Ile Lys Glu Ala Lys Gln Tyr Glu Glu Ala Ala 165 170 175	528
aaa aat att gta aca tct tta gat caa ttt tta cat ggt gat cag aaa Lys Asn Ile Val Thr Ser Leu Asp Gln Phe Leu His Gly Asp Gln Lys 180 185 190	576
aaa tta gaa ggt gtt atc aat att caa aaa cgt tta aaa gaa gtt caa Lys Leu Glu Gly Val Ile Asn Ile Gln Lys Arg Leu Lys Glu Val Gln 195 200 205	624
aca gct ctt aat caa gcc cat ggg gaa agt agt cca gct cat aaa gag Thr Ala Leu Asn Gln Ala His Gly Glu Ser Ser Pro Ala His Lys Glu 210 215 220	672
tta tta gaa aaa gta aaa aat tta aaa aca aca tta gaa agg act att Leu Leu Glu Lys Val Lys Asn Leu Lys Thr Thr Leu Glu Arg Thr Ile 225 230 235 240	720
aaa gct gaa caa gat tta gag aaa aaa gta gaa tat agt ttt cta tta Lys Ala Glu Gln Asp Leu Glu Lys Lys Val Glu Tyr Ser Phe Leu Leu 245 250 255	768
gga cca ttg tta gga ttt gtt gtt tat gaa att ctt gaa aat act gct Gly Pro Leu Leu Gly Phe Val Val Tyr Glu Ile Leu Glu Asn Thr Ala 260 265 270	816
gtt cag cat ata aaa aat caa att gat gag ata aag aaa caa tta gat Val Gln His Ile Lys Asn Gln Ile Asp Glu Ile Lys Lys Gln Leu Asp 275 280 285	864
tct gct cag cat gat ttg gat aga gat gtt aaa att ata gga atg tta Ser Ala Gln His Asp Leu Asp Arg Asp Val Lys Ile Ile Gly Met Leu 290 295 300	912

aat agt att aat aca gat att gat aat tta tat agt caa gga caa gaa 960
 Asn Ser Ile Asn Thr Asp Ile Asp Asn Leu Tyr Ser Gln Gly Gln Glu
 305 310 315 320

gca att aaa gtt ttc caa aag tta caa ggt att tgg gct act att gga 1008
 Ala Ile Lys Val Phe Gln Lys Leu Gln Gly Ile Trp Ala Thr Ile Gly
 325 330 335

gct caa ata gaa aat ctt aga aca acg tcg tta caa gaa gtt caa gat 1056
 Ala Gln Ile Glu Asn Leu Arg Thr Thr Ser Leu Gln Glu Val Gln Asp
 340 345 350

tct gat gat gct gat gag ata caa att gaa ctt gag gac gct tct gat 1104
 Ser Asp Asp Ala Asp Glu Ile Gln Ile Glu Leu Glu Asp Ala Ser Asp
 355 360 365

gct tgg tta gtt gtg gct caa gaa gct cgt gat ttt aca cta aat gct 1152
 Ala Trp Leu Val Val Ala Gln Glu Ala Arg Asp Phe Thr Leu Asn Ala
 370 375 380

tat tca act aat agt aga caa aat tta ccg att aat gtt ata tca gat 1200
 Tyr Ser Thr Asn Ser Arg Gln Asn Leu Pro Ile Asn Val Ile Ser Asp
 385 390 395 400

tca tgt aat tgt tca aca aca aat atg aca tca aat caa tac agt aat 1248
 Ser Cys Asn Cys Ser Thr Thr Asn Met Thr Ser Asn Gln Tyr Ser Asn
 405 410 415

cca aca aca aat atg aca tca aat caa tat atg att tca cat gaa tat 1296
 Pro Thr Thr Asn Met Thr Ser Asn Gln Tyr Met Ile Ser His Glu Tyr
 420 425 430

aca agt tta cca aat aat ttt atg tta tca aga aat agt aat tta gaa 1344
 Thr Ser Leu Pro Asn Asn Phe Met Leu Ser Arg Asn Ser Asn Leu Glu
 435 440 445

tat aaa tgt cct gaa aat aat ttt atg ata tat tgg tat aat aat tcg 1392
 Tyr Lys Cys Pro Glu Asn Asn Phe Met Ile Tyr Trp Tyr Asn Asn Ser
 450 455 460

gat tgg tat aat aat tcg gat tgg tat aat aat tga 1428
 Asp Trp Tyr Asn Asn Ser Asp Trp Tyr Asn Asn
 465 470 475

<210> 32
 <211> 475
 <212> PRT
 <213> Bacillus thuringiensis

<400> 32

Met Ile Ile Asp Ser Lys Thr Thr Leu Pro Arg His Ser Leu Ile His
 1 5 10 15

Thr Ile Lys Leu Asn Ser Asn Lys Lys Tyr Gly Pro Gly Asp Met Thr
 20 25 30

Asn Gly Asn Gln Phe Ile Ile Ser Lys Gln Glu Trp Ala Thr Ile Gly
 35 40 45

65

Ala Tyr Ile Gln Thr Gly Leu Gly Leu Pro Val Asn Glu Gln Gln Leu
50 55 60

Arg Thr His Val Asn Leu Ser Gln Asp Ile Ser Ile Pro Ser Asp Phe
65 70 75 80

Ser Gln Leu Tyr Asp Val Tyr Cys Ser Asp Lys Thr Ser Ala Glu Trp
85 90 95

Trp Asn Lys Asn Leu Tyr Pro Leu Ile Ile Lys Ser Ala Asn Asp Ile
100 105 110

Ala Ser Tyr Gly Phe Lys Val Ala Gly Asp Pro Ser Ile Lys Lys Asp
115 120 125

Gly Tyr Phe Lys Lys Leu Gln Asp Glu Leu Asp Asn Ile Val Asp Asn
130 135 140

Asn Ser Asp Asp Asp Ala Ile Ala Lys Ala Ile Lys Asp Phe Lys Ala
145 150 155 160

Arg Cys Gly Ile Leu Ile Lys Glu Ala Lys Gln Tyr Glu Glu Ala Ala
165 170 175

Lys Asn Ile Val Thr Ser Leu Asp Gln Phe Leu His Gly Asp Gln Lys
180 185 190

Lys Leu Glu Gly Val Ile Asn Ile Gln Lys Arg Leu Lys Glu Val Gln
195 200 205

Thr Ala Leu Asn Gln Ala His Gly Glu Ser Ser Pro Ala His Lys Glu
210 215 220

Leu Leu Glu Lys Val Lys Asn Leu Lys Thr Thr Leu Glu Arg Thr Ile
225 230 235 240

Lys Ala Glu Gln Asp Leu Glu Lys Lys Val Glu Tyr Ser Phe Leu Leu
245 250 255

Gly Pro Leu Leu Gly Phe Val Val Tyr Glu Ile Leu Glu Asn Thr Ala
260 265 270

Val Gln His Ile Lys Asn Gln Ile Asp Glu Ile Lys Lys Gln Leu Asp
275 280 285

Ser Ala Gln His Asp Leu Asp Arg Asp Val Lys Ile Ile Gly Met Leu
290 295 300

Asn Ser Ile Asn Thr Asp Ile Asp Asn Leu Tyr Ser Gln Gly Gln Glu
305 310 315 320

Ala Ile Lys Val Phe Gln Lys Leu Gln Gly Ile Trp Ala Thr Ile Gly
325 330 335

Ala Gln Ile Glu Asn Leu Arg Thr Thr Ser Leu Gln Glu Val Gln Asp
340 345 350

Ser Asp Asp Ala Asp Glu Ile Gln Ile Glu Leu Glu Asp Ala Ser Asp
355 360 365

Ala Trp Leu Val Val Ala Gln Glu Ala Arg Asp Phe Thr Leu Asn Ala
370 375 380

Tyr Ser Thr Asn Ser Arg Gln Asn Leu Pro Ile Asn Val Ile Ser Asp
385 390 395 400

Ser Cys Asn Cys Ser Thr Thr Asn Met Thr Ser Asn Gln Tyr Ser Asn
405 410 415

Pro Thr Thr Asn Met Thr Ser Asn Gln Tyr Met Ile Ser His Glu Tyr
420 425 430

Thr Ser Leu Pro Asn Asn Phe Met Leu Ser Arg Asn Ser Asn Leu Glu
435 440 445

Tyr Lys Cys Pro Glu Asn Asn Phe Met Ile Tyr Trp Tyr Asn Asn Ser
450 455 460

Asp Trp Tyr Asn Asn Ser Asp Trp Tyr Asn Asn
465 470 475

<210> 33
<211> 1428
<212> ДНК
<213> Искусственная последовательность

<220>
<223> Синтетическая последовательность ДНК, кодирующая токсин Cry6Aa *Bacillus thuringiensis* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1

<220>
<221> CDS
<222> (1)..(1428)

<400> 33
atg atc atc gac tcc aag acg acc ctg cca cgg cac tcc ctt atc cac 48
Met Ile Ile Asp Ser Lys Thr Thr Leu Pro Arg His Ser Leu Ile His
1 5 10 15

aca att aaa tta aat agc aat aag aag tac ggt ccc ggt gat atg act 96
 Thr Ile Lys Leu Asn Ser Asn Lys Lys Tyr Gly Pro Gly Asp Met Thr
 20 25 30

aac gga aat caa ttc att att tca aag caa gag tgg gct acc atc gga 144
 Asn Gly Asn Gln Phe Ile Ile Ser Lys Gln Glu Trp Ala Thr Ile Gly
 35 40 45

gcg tac atc cag act ggg ctg ggc cta cca gta aat gaa caa caa tta 192
 Ala Tyr Ile Gln Thr Gly Leu Gly Leu Pro Val Asn Glu Gln Gln Leu
 50 55 60

agg acc cat gtc aac ctc agc caa gat atc agc atc cct agc gac ttt 240
 Arg Thr His Val Asn Leu Ser Gln Asp Ile Ser Ile Pro Ser Asp Phe
 65 70 75 80

tct cag ctc tac gac gtc tat tgc agc gat aaa act tcc gca gaa tgg 288
 Ser Gln Leu Tyr Asp Val Tyr Cys Ser Asp Lys Thr Ser Ala Glu Trp
 85 90 95

tgg aat aaa aac ctg tac ccc ctc atc att aaa tct gcc aac gat att 336
 Trp Asn Lys Asn Leu Tyr Pro Leu Ile Ile Lys Ser Ala Asn Asp Ile
 100 105 110

gcc agc tac ggc ttc aag gtc gcg ggt gat cct tct att aag aag gac 384
 Ala Ser Tyr Gly Phe Lys Val Ala Gly Asp Pro Ser Ile Lys Lys Asp
 115 120 125

ggc tac ttc aag aag ctg caa gat gag ctg gac aac att gtt gac aat 432
 Gly Tyr Phe Lys Lys Leu Gln Asp Glu Leu Asp Asn Ile Val Asp Asn
 130 135 140

aat tcc gat gat gat gca ata gcg aaa gcc att aaa gac ttc aag gcg 480
 Asn Ser Asp Asp Asp Ala Ile Ala Lys Ala Ile Lys Asp Phe Lys Ala
 145 150 155 160

cga tgc ggc atc cta att aaa gaa gca aag cag tat gaa gag gca gcg 528
 Arg Cys Gly Ile Leu Ile Lys Glu Ala Lys Gln Tyr Glu Glu Ala Ala
 165 170 175

aaa aat atc gta aca tcc ctc gac caa ttt ctg cat ggc gat cag aag 576
 Lys Asn Ile Val Thr Ser Leu Asp Gln Phe Leu His Gly Asp Gln Lys
 180 185 190

aaa ttg gag ggt gtg atc aac atc caa aaa cgt ctg aag gag gtg cag 624
 Lys Leu Glu Gly Val Ile Asn Ile Gln Lys Arg Leu Lys Glu Val Gln
 195 200 205

acg gct ctt aat caa gcc cac ggg gaa agt tca cca gct cat aaa gag 672
 Thr Ala Leu Asn Gln Ala His Gly Glu Ser Ser Pro Ala His Lys Glu
 210 215 220

ctg tta gag aaa gtc aag aat ctc aag acc aca ctt gag agg acc att 720
 Leu Leu Glu Lys Val Lys Asn Leu Lys Thr Thr Leu Glu Arg Thr Ile
 225 230 235 240

aaa gct gag caa gac ctg gag aag aaa gtg gag tac agt ttc ctt ctc 768
 Lys Ala Glu Gln Asp Leu Glu Lys Lys Val Glu Tyr Ser Phe Leu Leu
 245 250 255

ggc ccc ttg ctg ggc ttc gtc gtt tat gaa atc ctt gaa aat act gcc 816
 Gly Pro Leu Leu Gly Phe Val Val Tyr Glu Ile Leu Glu Asn Thr Ala
 260 265 270

gtc cag cat ata aaa aac caa att gac gag ata aag aag caa ctg gac 864
 Val Gln His Ile Lys Asn Gln Ile Asp Glu Ile Lys Lys Gln Leu Asp
 275 280 285

tct gcc cag cac gac ttg gac aga gac gtt aag atc ata ggg atg ctg 912
 Ser Ala Gln His Asp Leu Asp Arg Asp Val Lys Ile Ile Gly Met Leu
 290 295 300

aac agt att aat aca gac att gat aac ttg tat agc caa gga caa gag 960
 Asn Ser Ile Asn Thr Asp Ile Asp Asn Leu Tyr Ser Gln Gly Gln Glu
 305 310 315 320

gca att aaa gtg ttc caa aag ctc caa ggc atc tgg gca act atc gga 1008
 Ala Ile Lys Val Phe Gln Lys Leu Gln Gly Ile Trp Ala Thr Ile Gly
 325 330 335

gcg cag ata gag aac ctt agg aca acg tcg ctc caa gaa gtg caa gac 1056
 Ala Gln Ile Glu Asn Leu Arg Thr Thr Ser Leu Gln Glu Val Gln Asp
 340 345 350

tct gac gac gcc gat gag atc caa att gaa ctt gag gac gcg tct gat 1104
 Ser Asp Asp Ala Asp Glu Ile Gln Ile Glu Leu Glu Asp Ala Ser Asp
 355 360 365

gct tgg tta gtg gtg gcc caa gaa gct cgc gac ttc aca cta aat gcc 1152
 Ala Trp Leu Val Val Ala Gln Glu Ala Arg Asp Phe Thr Leu Asn Ala
 370 375 380

tac tca act aac tcg cgt cag aat cta ccg att aat gtt ata tcc gat 1200
 Tyr Ser Thr Asn Ser Arg Gln Asn Leu Pro Ile Asn Val Ile Ser Asp
 385 390 395 400

tcc tgc aac tgt tcc aca acg aac atg acc tca aat caa tac agt aat 1248
 Ser Cys Asn Cys Ser Thr Thr Asn Met Thr Ser Asn Gln Tyr Ser Asn
 405 410 415

cca acc aca aat atg acc tca aat caa tat atg atc tca cac gag tat 1296
 Pro Thr Thr Asn Met Thr Ser Asn Gln Tyr Met Ile Ser His Glu Tyr
 420 425 430

acc tcg ttg ccg aat aat ttc atg ctc tca aga aat agc aat ctg gaa 1344
 Thr Ser Leu Pro Asn Asn Phe Met Leu Ser Arg Asn Ser Asn Leu Glu
 435 440 445

tat aag tgt cct gaa aat aat ttc atg ata tac tgg tac aat aat tcg 1392
 Tyr Lys Cys Pro Glu Asn Asn Phe Met Ile Tyr Trp Tyr Asn Asn Ser
 450 455 460

gac tgg tac aat aat tcg gat tgg tac aat aat tga 1428
 Asp Trp Tyr Asn Asn Ser Asp Trp Tyr Asn Asn
 465 470 475

<210> 34
 <211> 475
 <212> PRT
 <213> Искусственная последовательность

<220>
 <223> Синтетическая конструкция

<400> 34

69

Met Ile Ile Asp Ser Lys Thr Thr Leu Pro Arg His Ser Leu Ile His
1 5 10 15

Thr Ile Lys Leu Asn Ser Asn Lys Lys Tyr Gly Pro Gly Asp Met Thr
20 25 30

Asn Gly Asn Gln Phe Ile Ile Ser Lys Gln Glu Trp Ala Thr Ile Gly
35 40 45

Ala Tyr Ile Gln Thr Gly Leu Gly Leu Pro Val Asn Glu Gln Gln Leu
50 55 60

Arg Thr His Val Asn Leu Ser Gln Asp Ile Ser Ile Pro Ser Asp Phe
65 70 75 80

Ser Gln Leu Tyr Asp Val Tyr Cys Ser Asp Lys Thr Ser Ala Glu Trp
85 90 95

Trp Asn Lys Asn Leu Tyr Pro Leu Ile Ile Lys Ser Ala Asn Asp Ile
100 105 110

Ala Ser Tyr Gly Phe Lys Val Ala Gly Asp Pro Ser Ile Lys Lys Asp
115 120 125

Gly Tyr Phe Lys Lys Leu Gln Asp Glu Leu Asp Asn Ile Val Asp Asn
130 135 140

Asn Ser Asp Asp Asp Ala Ile Ala Lys Ala Ile Lys Asp Phe Lys Ala
145 150 155 160

Arg Cys Gly Ile Leu Ile Lys Glu Ala Lys Gln Tyr Glu Glu Ala Ala
165 170 175

Lys Asn Ile Val Thr Ser Leu Asp Gln Phe Leu His Gly Asp Gln Lys
180 185 190

Lys Leu Glu Gly Val Ile Asn Ile Gln Lys Arg Leu Lys Glu Val Gln
195 200 205

Thr Ala Leu Asn Gln Ala His Gly Glu Ser Ser Pro Ala His Lys Glu
210 215 220

Leu Leu Glu Lys Val Lys Asn Leu Lys Thr Thr Leu Glu Arg Thr Ile
225 230 235 240

Lys Ala Glu Gln Asp Leu Glu Lys Lys Val Glu Tyr Ser Phe Leu Leu
245 250 255

70

Gly Pro Leu Leu Gly Phe Val Val Tyr Glu Ile Leu Glu Asn Thr Ala
 260 265 270

Val Gln His Ile Lys Asn Gln Ile Asp Glu Ile Lys Lys Gln Leu Asp
 275 280 285

Ser Ala Gln His Asp Leu Asp Arg Asp Val Lys Ile Ile Gly Met Leu
 290 295 300

Asn Ser Ile Asn Thr Asp Ile Asp Asn Leu Tyr Ser Gln Gly Gln Glu
 305 310 315 320

Ala Ile Lys Val Phe Gln Lys Leu Gln Gly Ile Trp Ala Thr Ile Gly
 325 330 335

Ala Gln Ile Glu Asn Leu Arg Thr Thr Ser Leu Gln Glu Val Gln Asp
 340 345 350

Ser Asp Asp Ala Asp Glu Ile Gln Ile Glu Leu Glu Asp Ala Ser Asp
 355 360 365

Ala Trp Leu Val Val Ala Gln Glu Ala Arg Asp Phe Thr Leu Asn Ala
 370 375 380

Tyr Ser Thr Asn Ser Arg Gln Asn Leu Pro Ile Asn Val Ile Ser Asp
 385 390 395 400

Ser Cys Asn Cys Ser Thr Thr Asn Met Thr Ser Asn Gln Tyr Ser Asn
 405 410 415

Pro Thr Thr Asn Met Thr Ser Asn Gln Tyr Met Ile Ser His Glu Tyr
 420 425 430

Thr Ser Leu Pro Asn Asn Phe Met Leu Ser Arg Asn Ser Asn Leu Glu
 435 440 445

Tyr Lys Cys Pro Glu Asn Asn Phe Met Ile Tyr Trp Tyr Asn Asn Ser
 450 455 460

Asp Trp Tyr Asn Asn Ser Asp Trp Tyr Asn Asn
 465 470 475

<210> 35

<211> 1428

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Синтетическая последовательность ДНК по изобретению, кодирующая токсин Cry6Aa *Bacillus thuringiensis* с использованием кодонов, оптимизированных для

кукурузы, и в которой удаляются последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1

```

<220>
<221> CDS
<222> (1)..(1428)

<400> 35
atg atc atc gac tcc aag acg acc ctg cca cgg cac tcc ctt atc cac      48
Met Ile Ile Asp Ser Lys Thr Thr Leu Pro Arg His Ser Leu Ile His
1           5           10           15

aca att aaa tta aat agc aat aag aag tac ggt ccc ggt gat atg act      96
Thr Ile Lys Leu Asn Ser Asn Lys Lys Tyr Gly Pro Gly Asp Met Thr
                20           25           30

aac gga aat caa ttc att atc tca aag caa gag tgg gct acc atc gga      144
Asn Gly Asn Gln Phe Ile Ile Ser Lys Gln Glu Trp Ala Thr Ile Gly
                35           40           45

gcg tac atc cag act ggg ctg ggc cta cca gta aat gaa caa caa tta      192
Ala Tyr Ile Gln Thr Gly Leu Gly Leu Pro Val Asn Glu Gln Gln Leu
                50           55           60

agg acc cat gtc aac ctc agc caa gat atc agc atc cct agc gac ttt      240
Arg Thr His Val Asn Leu Ser Gln Asp Ile Ser Ile Pro Ser Asp Phe
65           70           75           80

tct cag ctc tac gac gtc tat tgc agc gat aaa act tcc gca gaa tgg      288
Ser Gln Leu Tyr Asp Val Tyr Cys Ser Asp Lys Thr Ser Ala Glu Trp
                85           90           95

tgg aat aaa aac ctg tac ccc ctc atc att aaa tct gcc aac gat att      336
Trp Asn Lys Asn Leu Tyr Pro Leu Ile Ile Lys Ser Ala Asn Asp Ile
                100          105          110

gcc agc tac ggc ttc aag gtc gcg ggt gat cct tct att aag aag gac      384
Ala Ser Tyr Gly Phe Lys Val Ala Gly Asp Pro Ser Ile Lys Lys Asp
                115          120          125

ggc tac ttc aag aag ctg caa gat gag ctg gac aac att gtt gac aat      432
Gly Tyr Phe Lys Lys Leu Gln Asp Glu Leu Asp Asn Ile Val Asp Asn
                130          135          140

aat tcc gat gat gat gca ata gcg aaa gcc att aaa gac ttc aag gcg      480
Asn Ser Asp Asp Asp Ala Ile Ala Lys Ala Ile Lys Asp Phe Lys Ala
145           150           155           160

cga tgc ggc atc cta att aaa gaa gca aag cag tat gaa gag gca gcg      528
Arg Cys Gly Ile Leu Ile Lys Glu Ala Lys Gln Tyr Glu Glu Ala Ala
                165          170          175

aaa aat atc gta aca tcc ctc gac caa ttt ctg cat ggc gat cag aag      576
Lys Asn Ile Val Thr Ser Leu Asp Gln Phe Leu His Gly Asp Gln Lys
                180          185          190

aaa ttg gag ggt gtg atc aac atc caa aaa cgt ctg aag gag gtg cag      624
Lys Leu Glu Gly Val Ile Asn Ile Gln Lys Arg Leu Lys Glu Val Gln
                195          200          205

acg gct ctt aat caa gcc cac ggg gaa agt tca cca gct cat aaa gag      672
Thr Ala Leu Asn Gln Ala His Gly Glu Ser Ser Pro Ala His Lys Glu
                210          215          220
    
```

ctg tta gag aaa gtc aag aat ctc aag acc aca ctt gag agg acc att Leu Leu Glu Lys Val Lys Asn Leu Lys Thr Thr Leu Glu Arg Thr Ile 225 230 235 240	720
aaa gct gag caa gac ctg gag aag aaa gtg gag tac agt ttc ctt ctc Lys Ala Glu Gln Asp Leu Glu Lys Lys Val Glu Tyr Ser Phe Leu Leu 245 250 255	768
ggc ccc ttg ctg ggc ttc gtc gtt tat gaa atc ctt gaa aat act gcc Gly Pro Leu Leu Gly Phe Val Val Tyr Glu Ile Leu Glu Asn Thr Ala 260 265 270	816
gtc cag cat ata aaa aac caa att gac gag ata aag aag caa ctg gac Val Gln His Ile Lys Asn Gln Ile Asp Glu Ile Lys Lys Gln Leu Asp 275 280 285	864
tct gcc cag cac gac ttg gac aga gac gtt aag atc ata ggg atg ctg Ser Ala Gln His Asp Leu Asp Arg Asp Val Lys Ile Ile Gly Met Leu 290 295 300	912
aac agt att aat aca gac att gat aac ttg tat agc caa gga caa gag Asn Ser Ile Asn Thr Asp Ile Asp Asn Leu Tyr Ser Gln Gly Gln Glu 305 310 315 320	960
gca att aaa gtg ttc caa aag ctc caa ggc atc tgg gca act atc gga Ala Ile Lys Val Phe Gln Lys Leu Gln Gly Ile Trp Ala Thr Ile Gly 325 330 335	1008
gcg cag ata gag aac ctt agg aca acg tcg ctc caa gaa gtg caa gac Ala Gln Ile Glu Asn Leu Arg Thr Thr Ser Leu Gln Glu Val Gln Asp 340 345 350	1056
tct gac gac gcc gat gag atc caa att gaa ctt gag gac gcg tct gat Ser Asp Asp Ala Asp Glu Ile Gln Ile Glu Leu Glu Asp Ala Ser Asp 355 360 365	1104
gct tgg tta gtg gtg gcc caa gaa gct cgc gac ttc aca cta aat gcc Ala Trp Leu Val Val Ala Gln Glu Ala Arg Asp Phe Thr Leu Asn Ala 370 375 380	1152
tac tca act aac tcg cgt cag aat cta ccg att aat gtt atc tcc gat Tyr Ser Thr Asn Ser Arg Gln Asn Leu Pro Ile Asn Val Ile Ser Asp 385 390 395 400	1200
tcc tgc aac tgt tcc aca acg aac atg acc tca aat caa tac agt aat Ser Cys Asn Cys Ser Thr Thr Asn Met Thr Ser Asn Gln Tyr Ser Asn 405 410 415	1248
cca acc aca aat atg acc tca aat caa tac atg atc tca cac gag tat Pro Thr Thr Asn Met Thr Ser Asn Gln Tyr Met Ile Ser His Glu Tyr 420 425 430	1296
acc tcg ttg ccg aat aat ttc atg ctc tca aga aat agc aat ctg gaa Thr Ser Leu Pro Asn Asn Phe Met Leu Ser Arg Asn Ser Asn Leu Glu 435 440 445	1344
tat aag tgt cct gaa aat aat ttc atg ata tac tgg tac aat aat tcg Tyr Lys Cys Pro Glu Asn Asn Phe Met Ile Tyr Trp Tyr Asn Asn Ser 450 455 460	1392
gac tgg tac aat aat tcg gat tgg tac aat aat tga Asp Trp Tyr Asn Asn Ser Asp Trp Tyr Asn Asn 465 470 475	1428

<210> 36
 <211> 475
 <212> PRT
 <213> Искусственная последовательность

 <220>
 <223> Синтетическая конструкция

 <400> 36

 Met Ile Ile Asp Ser Lys Thr Thr Leu Pro Arg His Ser Leu Ile His
 1 5 10 15

 Thr Ile Lys Leu Asn Ser Asn Lys Lys Tyr Gly Pro Gly Asp Met Thr
 20 25 30

 Asn Gly Asn Gln Phe Ile Ile Ser Lys Gln Glu Trp Ala Thr Ile Gly
 35 40 45

 Ala Tyr Ile Gln Thr Gly Leu Gly Leu Pro Val Asn Glu Gln Gln Leu
 50 55 60

 Arg Thr His Val Asn Leu Ser Gln Asp Ile Ser Ile Pro Ser Asp Phe
 65 70 75 80

 Ser Gln Leu Tyr Asp Val Tyr Cys Ser Asp Lys Thr Ser Ala Glu Trp
 85 90 95

 Trp Asn Lys Asn Leu Tyr Pro Leu Ile Ile Lys Ser Ala Asn Asp Ile
 100 105 110

 Ala Ser Tyr Gly Phe Lys Val Ala Gly Asp Pro Ser Ile Lys Lys Asp
 115 120 125

 Gly Tyr Phe Lys Lys Leu Gln Asp Glu Leu Asp Asn Ile Val Asp Asn
 130 135 140

 Asn Ser Asp Asp Asp Ala Ile Ala Lys Ala Ile Lys Asp Phe Lys Ala
 145 150 155 160

 Arg Cys Gly Ile Leu Ile Lys Glu Ala Lys Gln Tyr Glu Glu Ala Ala
 165 170 175

 Lys Asn Ile Val Thr Ser Leu Asp Gln Phe Leu His Gly Asp Gln Lys
 180 185 190

 Lys Leu Glu Gly Val Ile Asn Ile Gln Lys Arg Leu Lys Glu Val Gln
 195 200 205

74

Thr Ala Leu Asn Gln Ala His Gly Glu Ser Ser Pro Ala His Lys Glu
 210 215 220

Leu Leu Glu Lys Val Lys Asn Leu Lys Thr Thr Leu Glu Arg Thr Ile
 225 230 235 240

Lys Ala Glu Gln Asp Leu Glu Lys Lys Val Glu Tyr Ser Phe Leu Leu
 245 250 255

Gly Pro Leu Leu Gly Phe Val Val Tyr Glu Ile Leu Glu Asn Thr Ala
 260 265 270

Val Gln His Ile Lys Asn Gln Ile Asp Glu Ile Lys Lys Gln Leu Asp
 275 280 285

Ser Ala Gln His Asp Leu Asp Arg Asp Val Lys Ile Ile Gly Met Leu
 290 295 300

Asn Ser Ile Asn Thr Asp Ile Asp Asn Leu Tyr Ser Gln Gly Gln Glu
 305 310 315 320

Ala Ile Lys Val Phe Gln Lys Leu Gln Gly Ile Trp Ala Thr Ile Gly
 325 330 335

Ala Gln Ile Glu Asn Leu Arg Thr Thr Ser Leu Gln Glu Val Gln Asp
 340 345 350

Ser Asp Asp Ala Asp Glu Ile Gln Ile Glu Leu Glu Asp Ala Ser Asp
 355 360 365

Ala Trp Leu Val Val Ala Gln Glu Ala Arg Asp Phe Thr Leu Asn Ala
 370 375 380

Tyr Ser Thr Asn Ser Arg Gln Asn Leu Pro Ile Asn Val Ile Ser Asp
 385 390 395 400

Ser Cys Asn Cys Ser Thr Thr Asn Met Thr Ser Asn Gln Tyr Ser Asn
 405 410 415

Pro Thr Thr Asn Met Thr Ser Asn Gln Tyr Met Ile Ser His Glu Tyr
 420 425 430

Thr Ser Leu Pro Asn Asn Phe Met Leu Ser Arg Asn Ser Asn Leu Glu
 435 440 445

Tyr Lys Cys Pro Glu Asn Asn Phe Met Ile Tyr Trp Tyr Asn Asn Ser
 450 455 460

75

Asp Trp Tyr Asn Asn Ser Asp Trp Tyr Asn Asn
465 470 475

<210> 37
<211> 888
<212> ДНК
<213> *Sphingobiurn herbicidovorans*

<220>
<221> CDS
<222> (1)..(888)
<223> Нативная последовательность ДНК, кодирующая белок AAD1 *Sphingobiurn herbicidovorans*

<400> 37
atg cat gct gca ctg tcc ccc ctc tcc cag cgc ttt gag cgc atc gcg 48
Met His Ala Ala Leu Ser Pro Leu Ser Gln Arg Phe Glu Arg Ile Ala
1 5 10 15
gtc cag ccg ctg acc gcc gtc ctg ggc gcc gag atc acc ggc gtc gac 96
Val Gln Pro Leu Thr Gly Val Leu Gly Ala Glu Ile Thr Gly Val Asp
20 25 30
ctg cgc gag ccg ctc gac gac agc acc tgg aac gaa atc ctc gac gcg 144
Leu Arg Glu Pro Leu Asp Asp Ser Thr Trp Asn Glu Ile Leu Asp Ala
35 40 45
ttc cac act tac cag gtc atc tat ttt ccc gcc cag gcg atc acc aac 192
Phe His Thr Tyr Gln Val Ile Tyr Phe Pro Gly Gln Ala Ile Thr Asn
50 55 60
gaa cag cac atc gcc ttc agc cgg cgc ttc gcc ccc gtc gat ccc gtg 240
Glu Gln His Ile Ala Phe Ser Arg Arg Phe Gly Pro Val Asp Pro Val
65 70 75 80
ccc ctg ctc aag agc atc gaa ggg tat cca gag gtg cag atg atc cgc 288
Pro Leu Leu Lys Ser Ile Glu Gly Tyr Pro Glu Val Gln Met Ile Arg
85 90 95
cgc gaa gcc aac gaa agc ggg cgt gtg atc ggt gat gac tgg cac acc 336
Arg Glu Ala Asn Glu Ser Gly Arg Val Ile Gly Asp Asp Trp His Thr
100 105 110
gac agc acc ttc ctg gac gca ccg ccg gcc gcc gtg gtg atg cgc gcg 384
Asp Ser Thr Phe Leu Asp Ala Pro Pro Ala Ala Val Val Met Arg Ala
115 120 125
atc gac gtg ccc gag cat ggc gcc gac acc ggt ttt ctg agc atg tac 432
Ile Asp Val Pro Glu His Gly Gly Asp Thr Gly Phe Leu Ser Met Tyr
130 135 140
acc gcg tgg gag acg ctg tcg ccc acc atg cag gcc acc atc gaa ggg 480
Thr Ala Trp Glu Thr Leu Ser Pro Thr Met Gln Ala Thr Ile Glu Gly
145 150 155 160
ttg aac gta gtg cac agc gcc acg cgt gtg ttc gcc tcg ctc tac cag 528
Leu Asn Val Val His Ser Ala Thr Arg Val Phe Gly Ser Leu Tyr Gln
165 170 175
gcc cag aac ccg cgc ttc agc aac acc agc gtc aag gtg atg gac gtc 576
Ala Gln Asn Arg Arg Phe Ser Asn Thr Ser Val Lys Val Met Asp Val
180 185 190

gac gcg ggc gac cgt gaa acc gtg cac ccc ctg gtg gtg acc cat ccg 624
 Asp Ala Gly Asp Arg Glu Thr Val His Pro Leu Val Val Thr His Pro
 195 200 205

ggc agc ggc cgc aag ggc ctg tac gtg aac cag gtc tat tgc cag cgc 672
 Gly Ser Gly Arg Lys Gly Leu Tyr Val Asn Gln Val Tyr Cys Gln Arg
 210 215 220

atc gag ggc atg acc gat gcc gaa agc aaa ccg ctg ctg cag ttc ctg 720
 Ile Glu Gly Met Thr Asp Ala Glu Ser Lys Pro Leu Leu Gln Phe Leu
 225 230 235 240

tac gag cat gcg aca cgg ttc gat ttc acc tgc cgc gtg cgc tgg aag 768
 Tyr Glu His Ala Thr Arg Phe Asp Phe Thr Cys Arg Val Arg Trp Lys
 245 250 255

aag gac cag gtc ctg gtc tgg gac aac ctg tgc acg atg cac cgg gcc 816
 Lys Asp Gln Val Leu Val Trp Asp Asn Leu Cys Thr Met His Arg Ala
 260 265 270

gta ccc gac tac gcg ggc aag ttc cgc tac ctg acg cgc acc acg gtc 864
 Val Pro Asp Tyr Ala Gly Lys Phe Arg Tyr Leu Thr Arg Thr Thr Val
 275 280 285

ggg ggc gtg cgc ccg gcg cgc tag 888
 Gly Gly Val Arg Pro Ala Arg
 290 295

<210> 38
 <211> 295
 <212> PRT
 <213> Sphingobiurn herbicidovorans

<400> 38

Met His Ala Ala Leu Ser Pro Leu Ser Gln Arg Phe Glu Arg Ile Ala
 1 5 10 15

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

Leu Arg Glu Pro Leu Asp Asp Ser Thr Trp Asn Glu Ile Leu Asp Ala
 35 40 45

Phe His Thr Tyr Gln Val Ile Tyr Phe Pro Gly Gln Ala Ile Thr Asn
 50 55 60

Glu Gln His Ile Ala Phe Ser Arg Arg Phe Gly Pro Val Asp Pro Val
 65 70 75 80

Pro Leu Leu Lys Ser Ile Glu Gly Tyr Pro Glu Val Gln Met Ile Arg
 85 90 95

Arg Glu Ala Asn Glu Ser Gly Arg Val Ile Gly Asp Asp Trp His Thr
 100 105 110

77

Asp Ser Thr Phe Leu Asp Ala Pro Pro Ala Ala Val Val Met Arg Ala
115 120 125

Ile Asp Val Pro Glu His Gly Gly Asp Thr Gly Phe Leu Ser Met Tyr
130 135 140

Thr Ala Trp Glu Thr Leu Ser Pro Thr Met Gln Ala Thr Ile Glu Gly
145 150 155 160

Leu Asn Val Val His Ser Ala Thr Arg Val Phe Gly Ser Leu Tyr Gln
165 170 175

Ala Gln Asn Arg Arg Phe Ser Asn Thr Ser Val Lys Val Met Asp Val
180 185 190

Asp Ala Gly Asp Arg Glu Thr Val His Pro Leu Val Val Thr His Pro
195 200 205

Gly Ser Gly Arg Lys Gly Leu Tyr Val Asn Gln Val Tyr Cys Gln Arg
210 215 220

Ile Glu Gly Met Thr Asp Ala Glu Ser Lys Pro Leu Leu Gln Phe Leu
225 230 235 240

Tyr Glu His Ala Thr Arg Phe Asp Phe Thr Cys Arg Val Arg Trp Lys
245 250 255

Lys Asp Gln Val Leu Val Trp Asp Asn Leu Cys Thr Met His Arg Ala
260 265 270

Val Pro Asp Tyr Ala Gly Lys Phe Arg Tyr Leu Thr Arg Thr Thr Val
275 280 285

Gly Gly Val Arg Pro Ala Arg
290 295

<210> 39
<211> 888
<212> ДНК
<213> Искусственная последовательность

<220>
<223> Синтетическая последовательность ДНК, кодирующая белок AAD1 с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1 и таблицы 2

<220>
<221> CDS
<222> (1)..(888)

```

<400> 39
atg cac gct gca ctg tca cca ctc tca cag cgc ttt gag aga att gcg      48
Met His Ala Ala Leu Ser Pro Leu Ser Gln Arg Phe Glu Arg Ile Ala
1          5          10          15

gtc cag ccg ctg act ggc gtc ttg ggc gct gag atc acc ggc gtc gat      96
Val Gln Pro Leu Thr Gly Val Leu Gly Ala Glu Ile Thr Gly Val Asp
          20          25          30

ctg agg gag cct ctc gac gat tca acg tgg aac gaa att ctc gac gcg      144
Leu Arg Glu Pro Leu Asp Asp Ser Thr Trp Asn Glu Ile Leu Asp Ala
          35          40          45

ttc cat act tac caa gtc atc tat ttt ccc ggg caa gct att acc aac      192
Phe His Thr Tyr Gln Val Ile Tyr Phe Pro Gly Gln Ala Ile Thr Asn
          50          55          60

gaa caa cac atc gct ttc tct cgg cga ttc ggc ccc gtc gat cca gtg      240
Glu Gln His Ile Ala Phe Ser Arg Arg Phe Gly Pro Val Asp Pro Val
65          70          75          80

ccc tta ctc aag tct atc gaa ggc tac cca gag gtg cag atg ata aga      288
Pro Leu Leu Lys Ser Ile Glu Gly Tyr Pro Glu Val Gln Met Ile Arg
          85          90          95

agg gag gcc aac gaa agc ggg cgt gtg ata ggt gat gac tgg cac act      336
Arg Glu Ala Asn Glu Ser Gly Arg Val Ile Gly Asp Asp Trp His Thr
          100          105          110

gac agc aca ttc ctg gat gca ccg ccg gcc gct gtg gtg atg agg gca      384
Asp Ser Thr Phe Leu Asp Ala Pro Pro Ala Ala Val Val Met Arg Ala
          115          120          125

atc gac gtg ccc gag cac gga ggt gac act ggt ttc ttg agt atg tac      432
Ile Asp Val Pro Glu His Gly Gly Asp Thr Gly Phe Leu Ser Met Tyr
          130          135          140

act gct tgg gag acg ctt tcg cct act atg caa gcc aca atc gag ggg      480
Thr Ala Trp Glu Thr Leu Ser Pro Thr Met Gln Ala Thr Ile Glu Gly
145          150          155          160

ttg aat gta gtt cac agc gcc acg cgt gtg ttc gga tct ctc tat caa      528
Leu Asn Val Val His Ser Ala Thr Arg Val Phe Gly Ser Leu Tyr Gln
          165          170          175

gcc caa aac cgg cgc ttt tca aat acc tcc gtc aag gtg atg gac gtt      576
Ala Gln Asn Arg Arg Phe Ser Asn Thr Ser Val Lys Val Met Asp Val
          180          185          190

gac cgc ggc gac cgt gaa acc gtg cac cct ctt gtt gta acc cat ccg      624
Asp Ala Gly Asp Arg Glu Thr Val His Pro Leu Val Val Thr His Pro
          195          200          205

ggc agt ggt cgc aag ggc cta tac gtt aac caa gtc tat tgc cag cgc      672
Gly Ser Gly Arg Lys Gly Leu Tyr Val Asn Gln Val Tyr Cys Gln Arg
          210          215          220

atc gag gga atg aca gac gca gag agt aag ccg ctc ctg caa ttc ctg      720
Ile Glu Gly Met Thr Asp Ala Glu Ser Lys Pro Leu Leu Gln Phe Leu
225          230          235          240

tac gag cac gcg aca cgg ttc gat ttc acc tgc cgc gtg cgc tgg aaa      768
Tyr Glu His Ala Thr Arg Phe Asp Phe Thr Cys Arg Val Arg Trp Lys
          245          250          255

```

79

aag gat caa gtc ctt gta tgg gac aac ctt tgt acg atg cac cgg gcc 816
 Lys Asp Gln Val Leu Val Trp Asp Asn Leu Cys Thr Met His Arg Ala
 260 265 270

ggt cct gac tac gcg ggc aag ttc aga tac ctg acg agg acc acg gtc 864
 Val Pro Asp Tyr Ala Gly Lys Phe Arg Tyr Leu Thr Arg Thr Thr Val
 275 280 285

ggt gga gtt agg cca gcg aga tga 888
 Gly Gly Val Arg Pro Ala Arg
 290 295

<210> 40
 <211> 295
 <212> PRT
 <213> Искусственная последовательность

<220>
 <223> Синтетическая конструкция

<400> 40

Met His Ala Ala Leu Ser Pro Leu Ser Gln Arg Phe Glu Arg Ile Ala
 1 5 10 15

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

Leu Arg Glu Pro Leu Asp Asp Ser Thr Trp Asn Glu Ile Leu Asp Ala
 35 40 45

Phe His Thr Tyr Gln Val Ile Tyr Phe Pro Gly Gln Ala Ile Thr Asn
 50 55 60

Glu Gln His Ile Ala Phe Ser Arg Arg Phe Gly Pro Val Asp Pro Val
 65 70 75 80

Pro Leu Leu Lys Ser Ile Glu Gly Tyr Pro Glu Val Gln Met Ile Arg
 85 90 95

Arg Glu Ala Asn Glu Ser Gly Arg Val Ile Gly Asp Asp Trp His Thr
 100 105 110

Asp Ser Thr Phe Leu Asp Ala Pro Pro Ala Ala Val Val Met Arg Ala
 115 120 125

Ile Asp Val Pro Glu His Gly Gly Asp Thr Gly Phe Leu Ser Met Tyr
 130 135 140

Thr Ala Trp Glu Thr Leu Ser Pro Thr Met Gln Ala Thr Ile Glu Gly
 145 150 155 160

80

Leu Asn Val Val His Ser Ala Thr Arg Val Phe Gly Ser Leu Tyr Gln
 165 170 175

Ala Gln Asn Arg Arg Phe Ser Asn Thr Ser Val Lys Val Met Asp Val
 180 185 190

Asp Ala Gly Asp Arg Glu Thr Val His Pro Leu Val Val Thr His Pro
 195 200 205

Gly Ser Gly Arg Lys Gly Leu Tyr Val Asn Gln Val Tyr Cys Gln Arg
 210 215 220

Ile Glu Gly Met Thr Asp Ala Glu Ser Lys Pro Leu Leu Gln Phe Leu
 225 230 235 240

Tyr Glu His Ala Thr Arg Phe Asp Phe Thr Cys Arg Val Arg Trp Lys
 245 250 255

Lys Asp Gln Val Leu Val Trp Asp Asn Leu Cys Thr Met His Arg Ala
 260 265 270

Val Pro Asp Tyr Ala Gly Lys Phe Arg Tyr Leu Thr Arg Thr Thr Val
 275 280 285

Gly Gly Val Arg Pro Ala Arg
 290 295

<210> 41
 <211> 888
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Синтетическая последовательность ДНК по изобретению, кодирующая белок AAD1 с использованием кодонов, оптимизированных для кукурузы, и в которой удаляются последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1

<220>
 <221> CDS
 <222> (1)..(888)

<400> 41
 atg cac gct gca ctg tca cca ctc tca cag cgc ttt gag aga att gcg 48
 Met His Ala Ala Leu Ser Pro Leu Ser Gln Arg Phe Glu Arg Ile Ala
 1 5 10 15
 gtc cag ccg ctg act ggc gtc ttg ggc gct gag atc acc ggc gtc gat 96
 Val Gln Pro Leu Thr Gly Val Leu Gly Ala Glu Ile Thr Gly Val Asp
 20 25 30
 ctg agg gag cct ctc gac gat tca acg tgg aac gaa att ctc gac gcg 144
 Leu Arg Glu Pro Leu Asp Asp Ser Thr Trp Asn Glu Ile Leu Asp Ala
 35 40 45

ttc cat act tac caa gtc atc tac ttt ccc ggg caa gct att acc aac	192
Phe His Thr Tyr Gln Val Ile Tyr Phe Pro Gly Gln Ala Ile Thr Asn	
50 55 60	
gaa caa cac atc gct ttc tct cgg cga ttc ggc ccc gtc gat cca gtg	240
Glu Gln His Ile Ala Phe Ser Arg Arg Phe Gly Pro Val Asp Pro Val	
65 70 75 80	
ccc tta ctc aag tct atc gaa ggc tac cca gag gtg cag atg ata aga	288
Pro Leu Leu Lys Ser Ile Glu Gly Tyr Pro Glu Val Gln Met Ile Arg	
85 90 95	
agg gag gcc aac gaa agc ggg cgt gtg ata ggt gat gac tgg cac act	336
Arg Glu Ala Asn Glu Ser Gly Arg Val Ile Gly Asp Asp Trp His Thr	
100 105 110	
gac agc aca ttc ctg gat gca ccg ccg gcc gct gtg gtg atg agg gca	384
Asp Ser Thr Phe Leu Asp Ala Pro Pro Ala Ala Val Val Met Arg Ala	
115 120 125	
atc gac gtg ccc gag cac gga ggt gac act ggt ttc ttg agt atg tac	432
Ile Asp Val Pro Glu His Gly Gly Asp Thr Gly Phe Leu Ser Met Tyr	
130 135 140	
act gct tgg gag acg ctt tcg cct act atg caa gcc aca atc gag ggg	480
Thr Ala Trp Glu Thr Leu Ser Pro Thr Met Gln Ala Thr Ile Glu Gly	
145 150 155 160	
ttg aat gta gtt cac agc gcc acg cgt gtg ttc gga tct ctc tat caa	528
Leu Asn Val Val His Ser Ala Thr Arg Val Phe Gly Ser Leu Tyr Gln	
165 170 175	
gcc caa aac cgg cgc ttt tca aat acc tcc gtc aag gtg atg gac gtt	576
Ala Gln Asn Arg Arg Phe Ser Asn Thr Ser Val Lys Val Met Asp Val	
180 185 190	
gac gcg ggc gac cgt gaa acc gtg cac cct ctt gtt gta acc cat ccg	624
Asp Ala Gly Asp Arg Glu Thr Val His Pro Leu Val Val Thr His Pro	
195 200 205	
ggc agt ggt cgc aag ggc cta tac gtt aac caa gtc tat tgc cag cgc	672
Gly Ser Gly Arg Lys Gly Leu Tyr Val Asn Gln Val Tyr Cys Gln Arg	
210 215 220	
atc gag gga atg aca gac gca gag agt aag ccg ctc ctg caa ttc ctg	720
Ile Glu Gly Met Thr Asp Ala Glu Ser Lys Pro Leu Leu Gln Phe Leu	
225 230 235 240	
tac gag cac gcg aca cgg ttc gat ttc acc tgc cgc gtg cgc tgg aaa	768
Tyr Glu His Ala Thr Arg Phe Asp Phe Thr Cys Arg Val Arg Trp Lys	
245 250 255	
aag gat caa gtc ctt gta tgg gac aac ctt tgt acg atg cac cgg gcc	816
Lys Asp Gln Val Leu Val Trp Asp Asn Leu Cys Thr Met His Arg Ala	
260 265 270	
gtt cct gac tac gcg ggc aag ttc aga tac ctg acg agg acc acg gtc	864
Val Pro Asp Tyr Ala Gly Lys Phe Arg Tyr Leu Thr Arg Thr Thr Val	
275 280 285	
ggt gga gtt agg cca gcg aga tga	888
Gly Gly Val Arg Pro Ala Arg	
290 295	

<210> 42
 <211> 295
 <212> PRT
 <213> Искусственная последовательность

 <220>
 <223> Синтетическая конструкция

 <400> 42

 Met His Ala Ala Leu Ser Pro Leu Ser Gln Arg Phe Glu Arg Ile Ala
 1 5 10 15

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

 Leu Arg Glu Pro Leu Asp Asp Ser Thr Trp Asn Glu Ile Leu Asp Ala
 35 40 45

 Phe His Thr Tyr Gln Val Ile Tyr Phe Pro Gly Gln Ala Ile Thr Asn
 50 55 60

 Glu Gln His Ile Ala Phe Ser Arg Arg Phe Gly Pro Val Asp Pro Val
 65 70 75 80

 Pro Leu Leu Lys Ser Ile Glu Gly Tyr Pro Glu Val Gln Met Ile Arg
 85 90 95

 Arg Glu Ala Asn Glu Ser Gly Arg Val Ile Gly Asp Asp Trp His Thr
 100 105 110

 Asp Ser Thr Phe Leu Asp Ala Pro Pro Ala Ala Val Val Met Arg Ala
 115 120 125

 Ile Asp Val Pro Glu His Gly Gly Asp Thr Gly Phe Leu Ser Met Tyr
 130 135 140

 Thr Ala Trp Glu Thr Leu Ser Pro Thr Met Gln Ala Thr Ile Glu Gly
 145 150 155 160

 Leu Asn Val Val His Ser Ala Thr Arg Val Phe Gly Ser Leu Tyr Gln
 165 170 175

 Ala Gln Asn Arg Arg Phe Ser Asn Thr Ser Val Lys Val Met Asp Val
 180 185 190

 Asp Ala Gly Asp Arg Glu Thr Val His Pro Leu Val Val Thr His Pro
 195 200 205

83

Gly Ser Gly Arg Lys Gly Leu Tyr Val Asn Gln Val Tyr Cys Gln Arg
 210 215 220

Ile Glu Gly Met Thr Asp Ala Glu Ser Lys Pro Leu Leu Gln Phe Leu
 225 230 235 240

Tyr Glu His Ala Thr Arg Phe Asp Phe Thr Cys Arg Val Arg Trp Lys
 245 250 255

Lys Asp Gln Val Leu Val Trp Asp Asn Leu Cys Thr Met His Arg Ala
 260 265 270

Val Pro Asp Tyr Ala Gly Lys Phe Arg Tyr Leu Thr Arg Thr Thr Val
 275 280 285

Gly Gly Val Arg Pro Ala Arg
 290 295

<210> 43
 <211> 1368
 <212> ДНК
 <213> Aspergillus nidulans

<220>
 <221> CDS
 <222> (1)..(1368)
 <223> Нативная последовательность ДНК, кодирующая белок дельта-9-десатуразы
 жирных кислот Aspergillus nidulans

<400> 43
 atg tct gca cca acg gcg gac atc agg gct cgc gcc ccg gag gcc aaa 48
 Met Ser Ala Pro Thr Ala Asp Ile Arg Ala Arg Ala Pro Glu Ala Lys
 1 5 10 15
 aag gtt cac atc gct gac act gct atc aac cgc cat aac tgg tac aag 96
 Lys Val His Ile Ala Asp Thr Ala Ile Asn Arg His Asn Trp Tyr Lys
 20 25 30
 cat gtg aac tgg ctg aac gtt ttc ctg atc atc ggt atc ccg ctt tat 144
 His Val Asn Trp Leu Asn Val Phe Leu Ile Ile Gly Ile Pro Leu Tyr
 35 40 45
 ggg tgc att cag gcg ttc tgg gtg cca ctg cag ctg aag act gcc atc 192
 Gly Cys Ile Gln Ala Phe Trp Val Pro Leu Gln Leu Lys Thr Ala Ile
 50 55 60
 tgg gcc gtc atc tac tac ttt ttc acc ggt ctc ggt atc aca gca ggt 240
 Trp Ala Val Ile Tyr Tyr Phe Phe Thr Gly Leu Gly Ile Thr Ala Gly
 65 70 75 80
 tac cat cgt cta tgg gct cac tgc tgc tac tcc gcc acc ctt cct ttg 288
 Tyr His Arg Leu Trp Ala His Cys Ser Tyr Ser Ala Thr Leu Pro Leu
 85 90 95
 cgt atc tgg ctc gct gcc gtt ggt ggt gcc gtc gaa ggt tct atc 336
 Arg Ile Trp Leu Ala Ala Val Gly Gly Ala Val Glu Gly Ser Ile
 100 105 110

cgc tgg tgg gct cgt gac cac cgc gct cac cac cgc tac acc gat acc Arg Trp Trp Ala Arg Asp His Arg Ala His His Arg Tyr Thr Asp Thr 115 120 125	384
gac aaa gac ccg tac tcc gtt cgc aag ggt ctg ctc tac tct cac ctt Asp Lys Asp Pro Tyr Ser Val Arg Lys Gly Leu Leu Tyr Ser His Leu 130 135 140	432
ggc tgg atg gtg atg aag cag aac cct aag cgt att ggc cgt acc gat Gly Trp Met Val Met Lys Gln Asn Pro Lys Arg Ile Gly Arg Thr Asp 145 150 155 160	480
att tcc gac ctg aac gag gac ccc gtc gtt gtc tgg cag cac cgc aac Ile Ser Asp Leu Asn Glu Asp Pro Val Val Val Trp Gln His Arg Asn 165 170 175	528
tac ctc aag gtc gtt ttc acg atg gga ttg gct gtg cct atg ctt gtt Tyr Leu Lys Val Val Phe Thr Met Gly Leu Ala Val Pro Met Leu Val 180 185 190	576
gct ggt ctt gga tgg ggt gac tgg ttg ggc ggc ttc gtg tat gcc ggc Ala Gly Leu Gly Trp Gly Asp Trp Leu Gly Gly Phe Val Tyr Ala Gly 195 200 205	624
att ctg cgt atc ttc ttc gtc cag cag gcg act ttc tgc gtc aac tct Ile Leu Arg Ile Phe Phe Val Gln Gln Ala Thr Phe Cys Val Asn Ser 210 215 220	672
ttg gcc cac tgg ctc ggt gac cag ccc ttc gat gac cgc aac tca cct Leu Ala His Trp Leu Gly Asp Gln Pro Phe Asp Asp Arg Asn Ser Pro 225 230 235 240	720
cgt gac cac gtt atc acc gct ctc gtc acc ctt gga gag ggc tac cac Arg Asp His Val Ile Thr Ala Leu Val Thr Leu Gly Glu Gly Tyr His 245 250 255	768
aac ttc cac cac gag ttc ccc tcg gac tac cgt aac gcc atc gaa tgg Asn Phe His His Glu Phe Pro Ser Asp Tyr Arg Asn Ala Ile Glu Trp 260 265 270	816
cac cag tat gat ccc acc aag tgg tcc atc tgg gcc tgg aag cag ctt His Gln Tyr Asp Pro Thr Lys Trp Ser Ile Trp Ala Trp Lys Gln Leu 275 280 285	864
ggt ctt gcc tac gac ctg aag aag ttc cgt gcc aac gag att gag aag Gly Leu Ala Tyr Asp Leu Lys Lys Phe Arg Ala Asn Glu Ile Glu Lys 290 295 300	912
ggt cgt gtc cag cag ctc cag aag aag ctt gac cgt aag cgt gcc act Gly Arg Val Gln Gln Leu Gln Lys Lys Leu Asp Arg Lys Arg Ala Thr 305 310 315 320	960
ctc gat tgg ggt act cct ctt gac cag ctc ccc gtc atg gag tgg gac Leu Asp Trp Gly Thr Pro Leu Asp Gln Leu Pro Val Met Glu Trp Asp 325 330 335	1008
gac tac gtc gag cag gct aag aac ggc cgc ggt ctc gtg gct att gcc Asp Tyr Val Glu Gln Ala Lys Asn Gly Arg Gly Leu Val Ala Ile Ala 340 345 350	1056
ggt gtt gtc cac gat gtc acg gac ttc atc aaa gac cac ccc ggt ggc Gly Val Val His Asp Val Thr Asp Phe Ile Lys Asp His Pro Gly Gly 355 360 365	1104

```

aag gcc atg atc agc tcc ggt att ggg aag gac gcc acc gcc atg ttc      1152
Lys Ala Met Ile Ser Ser Gly Ile Gly Lys Asp Ala Thr Ala Met Phe
    370                      375                      380

aac ggt ggt gtc tac tac cac tcc aac gcc gca cac aac ctc ctc tct      1200
Asn Gly Gly Val Tyr Tyr His Ser Asn Ala Ala His Asn Leu Leu Ser
    385                      390                      395                      400

acc atg cgt gtt ggt gtt atc cgc ggc ggc tgt gaa gtc gaa atc tgg      1248
Thr Met Arg Val Gly Val Ile Arg Gly Gly Cys Glu Val Glu Ile Trp
                405                      410                      415

aag cgt gcc cag aag gag aac gtg gag tac gtg cgt gat ggc tct ggc      1296
Lys Arg Ala Gln Lys Glu Asn Val Glu Tyr Val Arg Asp Gly Ser Gly
                420                      425                      430

cag cgc gtc atc cgt gcc ggc gag cag cca acc aag atc cca gaa ccc      1344
Gln Arg Val Ile Arg Ala Gly Glu Gln Pro Thr Lys Ile Pro Glu Pro
                435                      440                      445

att ccc aca gcg gat gcg gcg tga      1368
Ile Pro Thr Ala Asp Ala Ala
    450                      455

<210> 44
<211> 455
<212> PRT
<213> Aspergillus nidulans

<400> 44

Met Ser Ala Pro Thr Ala Asp Ile Arg Ala Arg Ala Pro Glu Ala Lys
 1                      5                      10                      15

Lys Val His Ile Ala Asp Thr Ala Ile Asn Arg His Asn Trp Tyr Lys
 20                      25                      30

His Val Asn Trp Leu Asn Val Phe Leu Ile Ile Gly Ile Pro Leu Tyr
 35                      40                      45

Gly Cys Ile Gln Ala Phe Trp Val Pro Leu Gln Leu Lys Thr Ala Ile
 50                      55                      60

Trp Ala Val Ile Tyr Tyr Phe Phe Thr Gly Leu Gly Ile Thr Ala Gly
 65                      70                      75                      80

Tyr His Arg Leu Trp Ala His Cys Ser Tyr Ser Ala Thr Leu Pro Leu
 85                      90                      95

Arg Ile Trp Leu Ala Ala Val Gly Gly Gly Ala Val Glu Gly Ser Ile
 100                      105                      110

Arg Trp Trp Ala Arg Asp His Arg Ala His His Arg Tyr Thr Asp Thr
 115                      120                      125

```

Asp Lys Asp Pro Tyr Ser Val Arg Lys Gly Leu Leu Tyr Ser His Leu
 130 135 140

Gly Trp Met Val Met Lys Gln Asn Pro Lys Arg Ile Gly Arg Thr Asp
 145 150 155 160

Ile Ser Asp Leu Asn Glu Asp Pro Val Val Val Trp Gln His Arg Asn
 165 170 175

Tyr Leu Lys Val Val Phe Thr Met Gly Leu Ala Val Pro Met Leu Val
 180 185 190

Ala Gly Leu Gly Trp Gly Asp Trp Leu Gly Gly Phe Val Tyr Ala Gly
 195 200 205

Ile Leu Arg Ile Phe Phe Val Gln Gln Ala Thr Phe Cys Val Asn Ser
 210 215 220

Leu Ala His Trp Leu Gly Asp Gln Pro Phe Asp Asp Arg Asn Ser Pro
 225 230 235 240

Arg Asp His Val Ile Thr Ala Leu Val Thr Leu Gly Glu Gly Tyr His
 245 250 255

Asn Phe His His Glu Phe Pro Ser Asp Tyr Arg Asn Ala Ile Glu Trp
 260 265 270

His Gln Tyr Asp Pro Thr Lys Trp Ser Ile Trp Ala Trp Lys Gln Leu
 275 280 285

Gly Leu Ala Tyr Asp Leu Lys Lys Phe Arg Ala Asn Glu Ile Glu Lys
 290 295 300

Gly Arg Val Gln Gln Leu Gln Lys Lys Leu Asp Arg Lys Arg Ala Thr
 305 310 315 320

Leu Asp Trp Gly Thr Pro Leu Asp Gln Leu Pro Val Met Glu Trp Asp
 325 330 335

Asp Tyr Val Glu Gln Ala Lys Asn Gly Arg Gly Leu Val Ala Ile Ala
 340 345 350

Gly Val Val His Asp Val Thr Asp Phe Ile Lys Asp His Pro Gly Gly
 355 360 365

Lys Ala Met Ile Ser Ser Gly Ile Gly Lys Asp Ala Thr Ala Met Phe
 370 375 380

Asn Gly Gly Val Tyr Tyr His Ser Asn Ala Ala His Asn Leu Leu Ser
 385 390 395 400

Thr Met Arg Val Gly Val Ile Arg Gly Gly Cys Glu Val Glu Ile Trp
 405 410 415

Lys Arg Ala Gln Lys Glu Asn Val Glu Tyr Val Arg Asp Gly Ser Gly
 420 425 430

Gln Arg Val Ile Arg Ala Gly Glu Gln Pro Thr Lys Ile Pro Glu Pro
 435 440 445

Ile Pro Thr Ala Asp Ala Ala
 450 455

<210> 45
 <211> 1368
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Синтетическая последовательность ДНК, кодирующая белок дельта-9-
 десатуразы жирных кислот *Aspergillus nidulans* с использованием кодонов,
 оптимизированных для кукурузы, и в которой поддерживаются последовательности из
 таблицы 1 и таблицы 2

<220>
 <221> CDS
 <222> (1)..(1368)

<400> 45
 atg agt gca cca acg gcg gac ata agg gcg cgc gcc ccg gag gca aaa 48
 Met Ser Ala Pro Thr Ala Asp Ile Arg Ala Arg Ala Pro Glu Ala Lys
 1 5 10 15
 aag gtt cac att gct gac act gct atc aat cgc cat aac tgg tat aag 96
 Lys Val His Ile Ala Asp Thr Ala Ile Asn Arg His Asn Trp Tyr Lys
 20 25 30
 cat gtg aat tgg ctg aac gtt ttt ctg atc atc ggc atc ccg ctt tat 144
 His Val Asn Trp Leu Asn Val Phe Leu Ile Ile Gly Ile Pro Leu Tyr
 35 40 45
 ggg tgt att caa gcg ttc tgg gtg cca ctc cag ctc aag act gcc atc 192
 Gly Cys Ile Gln Ala Phe Trp Val Pro Leu Gln Leu Lys Thr Ala Ile
 50 55 60
 tgg gcc gta atc tac tac ttc ttt acc ggt ttg gga atc aca gcg ggt 240
 Trp Ala Val Ile Tyr Tyr Phe Phe Thr Gly Leu Gly Ile Thr Ala Gly
 65 70 75 80
 tat cac aga ttg tgg gca cac tgc tcg tac tcc gcc acc ctt cct tta 288
 Tyr His Arg Leu Trp Ala His Cys Ser Tyr Ser Ala Thr Leu Pro Leu
 85 90 95
 cgt ata tgg ctc gct gcc gta gga gga ggc gcc gtc gaa ggt tca atc 336
 Arg Ile Trp Leu Ala Ala Val Gly Gly Gly Ala Val Glu Gly Ser Ile

100	105	110	
cgt tgg tgg gct aga gac cat	cgt gct cat cat aga tat	acc gat aca	384
Arg Trp Trp Ala Arg Asp His	Arg Ala His His Arg Tyr	Thr Asp Thr	
115	120	125	
gac aaa gac ccg tac tcc gtt	cgc aag ggg ctg cta tac tct cac ctt		432
Asp Lys Asp Pro Tyr Ser Val	Arg Lys Gly Leu Leu Tyr Ser His Leu		
130	135	140	
ggc tgg atg gtg atg aag cag	aac cct aag cgt att ggc aga acc gat		480
Gly Trp Met Val Met Lys Gln	Asn Pro Lys Arg Ile Gly Arg Thr Asp		
145	150	155	160
att tcc gac ctg aac gag gac	ccc gtc gtt gtc tgg cag cac cgg aac		528
Ile Ser Asp Leu Asn Glu Asp	Pro Val Val Val Trp Gln His Arg Asn		
165	170	175	
tac ctc aag gtc gtt ttc acg	atg gga ttg gct gtg cct atg ctt gtt		576
Tyr Leu Lys Val Val Phe Thr	Met Gly Leu Ala Val Pro Met Leu Val		
180	185	190	
gct ggg ctt ggc tgg gga gac	tgg ttg ggc ggc ttc gtg tat gcc ggc		624
Ala Gly Leu Gly Trp Gly Asp	Trp Leu Gly Gly Phe Val Tyr Ala Gly		
195	200	205	
ata ctg aga atc ttt ttc gtc	cag caa gcg act ttt tgc gtc aac tct		672
Ile Leu Arg Ile Phe Phe Val	Gln Gln Ala Thr Phe Cys Val Asn Ser		
210	215	220	
ttg gcc cac tgg ctc gga gat	cag ccg ttc gat gac cgg aac agt cct		720
Leu Ala His Trp Leu Gly Asp	Gln Pro Phe Asp Asp Arg Asn Ser Pro		
225	230	235	240
agg gac cac gtt atc act gct	ctc gtc acc cta gga gag ggc tac cac		768
Arg Asp His Val Ile Thr Ala	Leu Val Thr Leu Gly Glu Gly Tyr His		
245	250	255	
aac ttc cat cac gag ttc ccc	tcg gac tac cgg aac gcc atc gaa tgg		816
Asn Phe His His Glu Phe Pro	Ser Asp Tyr Arg Asn Ala Ile Glu Trp		
260	265	270	
cac cag tat gat cca acg aag	tgg agc atc tgg gcc tgg aag cag ctt		864
His Gln Tyr Asp Pro Thr Lys	Trp Ser Ile Trp Ala Trp Lys Gln Leu		
275	280	285	
ggt tta gcc tac gac ctg aag	aaa ttc aga gcc aac gag att gag aaa		912
Gly Leu Ala Tyr Asp Leu Lys	Lys Phe Arg Ala Asn Glu Ile Glu Lys		
290	295	300	
ggg cgt gtc caa cag ctg caa	aag aaa ctg gac cgt aag cgg gcg act		960
Gly Arg Val Gln Gln Leu Gln	Lys Lys Leu Asp Arg Lys Arg Ala Thr		
305	310	315	320
ctc gat tgg gga aca cct ctg	gat cag ctc ccc gtc atg gag tgg gac		1008
Leu Asp Trp Gly Thr Pro Leu	Asp Gln Leu Pro Val Met Glu Trp Asp		
325	330	335	
gac tac gtg gag caa gca aag	aac ggt cgc ggt ctc gtg gca ata gcg		1056
Asp Tyr Val Glu Gln Ala Lys	Asn Gly Arg Gly Leu Val Ala Ile Ala		
340	345	350	
ggc gtg gtg cac gat gtc acg	gat ttc atc aaa gat cac ccg ggg ggc		1104
Gly Val Val His Asp Val Thr	Asp Phe Ile Lys Asp His Pro Gly Gly		

355	360	365	
aag gcc atg atc agc tcc ggg att ggc aag gac gca acc gcc atg ttc			1152
Lys Ala Met Ile Ser Ser Gly Ile Gly Lys Asp Ala Thr Ala Met Phe			
370	375	380	
aat ggg gga gtc tac tac cac agc aac gca gca cac aat ctc ttg tca			1200
Asn Gly Gly Val Tyr Tyr His Ser Asn Ala Ala His Asn Leu Leu Ser			
385	390	395	400
aca atg agg gtg ggt gtt att agg ggc ggc tgt gaa gtc gaa atc tgg			1248
Thr Met Arg Val Gly Val Ile Arg Gly Gly Cys Glu Val Glu Ile Trp			
405	410	415	
aag agg gcg caa aag gag aat gtg gag tac gtg cga gat ggc tct ggt			1296
Lys Arg Ala Gln Lys Glu Asn Val Glu Tyr Val Arg Asp Gly Ser Gly			
420	425	430	
caa cgc gtg atc aga gcg ggc gag cag cca acc aag ata cca gaa ccg			1344
Gln Arg Val Ile Arg Ala Gly Glu Gln Pro Thr Lys Ile Pro Glu Pro			
435	440	445	
att ccc aca gcg gat gcg gcg tag			1368
Ile Pro Thr Ala Asp Ala Ala			
450	455		
<210> 46			
<211> 455			
<212> PRT			
<213> Искусственная последовательность			
<220>			
<223> Синтетическая конструкция			
<400> 46			
Met Ser Ala Pro Thr Ala Asp Ile Arg Ala Arg Ala Pro Glu Ala Lys			
1	5	10	15
Lys Val His Ile Ala Asp Thr Ala Ile Asn Arg His Asn Trp Tyr Lys			
20	25	30	
His Val Asn Trp Leu Asn Val Phe Leu Ile Ile Gly Ile Pro Leu Tyr			
35	40	45	
Gly Cys Ile Gln Ala Phe Trp Val Pro Leu Gln Leu Lys Thr Ala Ile			
50	55	60	
Trp Ala Val Ile Tyr Tyr Phe Phe Thr Gly Leu Gly Ile Thr Ala Gly			
65	70	75	80
Tyr His Arg Leu Trp Ala His Cys Ser Tyr Ser Ala Thr Leu Pro Leu			
85	90	95	
Arg Ile Trp Leu Ala Ala Val Gly Gly Gly Ala Val Glu Gly Ser Ile			
100	105	110	

Arg Trp Trp Ala Arg Asp His Arg Ala His His Arg Tyr Thr Asp Thr
 115 120 125

Asp Lys Asp Pro Tyr Ser Val Arg Lys Gly Leu Leu Tyr Ser His Leu
 130 135 140

Gly Trp Met Val Met Lys Gln Asn Pro Lys Arg Ile Gly Arg Thr Asp
 145 150 155 160

Ile Ser Asp Leu Asn Glu Asp Pro Val Val Val Trp Gln His Arg Asn
 165 170 175

Tyr Leu Lys Val Val Phe Thr Met Gly Leu Ala Val Pro Met Leu Val
 180 185 190

Ala Gly Leu Gly Trp Gly Asp Trp Leu Gly Gly Phe Val Tyr Ala Gly
 195 200 205

Ile Leu Arg Ile Phe Phe Val Gln Gln Ala Thr Phe Cys Val Asn Ser
 210 215 220

Leu Ala His Trp Leu Gly Asp Gln Pro Phe Asp Asp Arg Asn Ser Pro
 225 230 235 240

Arg Asp His Val Ile Thr Ala Leu Val Thr Leu Gly Glu Gly Tyr His
 245 250 255

Asn Phe His His Glu Phe Pro Ser Asp Tyr Arg Asn Ala Ile Glu Trp
 260 265 270

His Gln Tyr Asp Pro Thr Lys Trp Ser Ile Trp Ala Trp Lys Gln Leu
 275 280 285

Gly Leu Ala Tyr Asp Leu Lys Lys Phe Arg Ala Asn Glu Ile Glu Lys
 290 295 300

Gly Arg Val Gln Gln Leu Gln Lys Lys Leu Asp Arg Lys Arg Ala Thr
 305 310 315 320

Leu Asp Trp Gly Thr Pro Leu Asp Gln Leu Pro Val Met Glu Trp Asp
 325 330 335

Asp Tyr Val Glu Gln Ala Lys Asn Gly Arg Gly Leu Val Ala Ile Ala
 340 345 350

Gly Val Val His Asp Val Thr Asp Phe Ile Lys Asp His Pro Gly Gly
 355 360 365

Lys Ala Met Ile Ser Ser Gly Ile Gly Lys Asp Ala Thr Ala Met Phe
 370 375 380

Asn Gly Gly Val Tyr Tyr His Ser Asn Ala Ala His Asn Leu Leu Ser
 385 390 395 400

Thr Met Arg Val Gly Val Ile Arg Gly Gly Cys Glu Val Glu Ile Trp
 405 410 415

Lys Arg Ala Gln Lys Glu Asn Val Glu Tyr Val Arg Asp Gly Ser Gly
 420 425 430

Gln Arg Val Ile Arg Ala Gly Glu Gln Pro Thr Lys Ile Pro Glu Pro
 435 440 445

Ile Pro Thr Ala Asp Ala Ala
 450 455

<210> 47
 <211> 1368
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Синтетическая последовательность ДНК по изобретению, кодирующая белок дельта-9-десатуразы жирных кислот *Aspergillus nidulans* с использованием кодонов, оптимизированных для кукурузы, и в которой удаляются последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1

<220>
 <221> CDS
 <222> (1)..(1368)

<400> 47
 atg agt gca cca acg gcg gac ata agg gcg cgc gcc ccg gag gca aaa 48
 Met Ser Ala Pro Thr Ala Asp Ile Arg Ala Arg Ala Pro Glu Ala Lys
 1 5 10 15
 aag gtt cac att gct gac act gct atc aat cgc cat aac tgg tat aag 96
 Lys Val His Ile Ala Asp Thr Ala Ile Asn Arg His Asn Trp Tyr Lys
 20 25 30
 cat gtg aat tgg ctg aac gtt ttt ctg atc atc ggc atc ccg ctt tat 144
 His Val Asn Trp Leu Asn Val Phe Leu Ile Ile Gly Ile Pro Leu Tyr
 35 40 45
 ggg tgt att caa gcg ttc tgg gtg cca ctc cag ctc aag act gcc atc 192
 Gly Cys Ile Gln Ala Phe Trp Val Pro Leu Gln Leu Lys Thr Ala Ile
 50 55 60
 tgg gcc gta atc tac tac ttc ttt acc ggt ttg gga atc aca gcg ggt 240
 Trp Ala Val Ile Tyr Tyr Phe Phe Thr Gly Leu Gly Ile Thr Ala Gly
 65 70 75 80
 tat cac aga ttg tgg gca cac tgc tcg tac tcc gcc acc ctt cct tta 288
 Tyr His Arg Leu Trp Ala His Cys Ser Tyr Ser Ala Thr Leu Pro Leu

85	90	95	
cgt ata tgg ctc gct gcc gta gga gga ggc gcc gtc gaa ggt tca atc Arg Ile Trp Leu Ala Ala Val Gly Gly Gly Ala Val Glu Gly Ser Ile 100 105 110			336
cgt tgg tgg gct aga gac cat cgt gct cat cat aga tat acc gat aca Arg Trp Trp Ala Arg Asp His Arg Ala His His Arg Tyr Thr Asp Thr 115 120 125			384
gac aaa gac ccg tac tcc gtt cgc aag ggg ctg cta tac tct cac ctt Asp Lys Asp Pro Tyr Ser Val Arg Lys Gly Leu Leu Tyr Ser His Leu 130 135 140			432
ggc tgg atg gtg atg aag cag aac cct aag cgt att ggc aga acc gat Gly Trp Met Val Met Lys Gln Asn Pro Lys Arg Ile Gly Arg Thr Asp 145 150 155 160			480
att agc gac ctg aac gag gac ccc gtc gtt gtc tgg cag cac cgg aac Ile Ser Asp Leu Asn Glu Asp Pro Val Val Val Trp Gln His Arg Asn 165 170 175			528
tac ctc aag gtc gtt ttc acg atg gga ttg gct gtg cct atg ctt gtt Tyr Leu Lys Val Val Phe Thr Met Gly Leu Ala Val Pro Met Leu Val 180 185 190			576
gct ggg ctt ggc tgg gga gac tgg ttg ggc gcc ttc gtg tat gcc ggc Ala Gly Leu Gly Trp Gly Asp Trp Leu Gly Gly Phe Val Tyr Ala Gly 195 200 205			624
ata ctg aga atc ttt ttc gtc cag caa gcg act ttt tgc gtc aac tct Ile Leu Arg Ile Phe Phe Val Gln Gln Ala Thr Phe Cys Val Asn Ser 210 215 220			672
ttg gcc cac tgg ctc gga gat cag ccg ttc gat gac cgg aac agt cct Leu Ala His Trp Leu Gly Asp Gln Pro Phe Asp Asp Arg Asn Ser Pro 225 230 235 240			720
agg gac cac gtt atc act gct ctc gtc acc cta gga gag ggc tac cac Arg Asp His Val Ile Thr Ala Leu Val Thr Leu Gly Glu Gly Tyr His 245 250 255			768
aac ttc cat cac gag ttc ccc tcg gac tac ccg aac gcc atc gaa tgg Asn Phe His His Glu Phe Pro Ser Asp Tyr Arg Asn Ala Ile Glu Trp 260 265 270			816
cac cag tat gat cca acg aag tgg agc atc tgg gcc tgg aag cag ctt His Gln Tyr Asp Pro Thr Lys Trp Ser Ile Trp Ala Trp Lys Gln Leu 275 280 285			864
ggt tta gcc tac gac ctg aag aaa ttc aga gcc aac gag att gag aaa Gly Leu Ala Tyr Asp Leu Lys Lys Phe Arg Ala Asn Glu Ile Glu Lys 290 295 300			912
ggg cgt gtc caa cag ctg caa aag aaa ctg gac cgt aag ccg gcg act Gly Arg Val Gln Gln Leu Gln Lys Lys Leu Asp Arg Lys Arg Ala Thr 305 310 315 320			960
ctc gat tgg gga aca cct ctg gat cag ctc ccc gtc atg gag tgg gac Leu Asp Trp Gly Thr Pro Leu Asp Gln Leu Pro Val Met Glu Trp Asp 325 330 335			1008
gac tac gtg gag caa gca aag aac ggt cgc ggt ctc gtg gca ata gcg Asp Tyr Val Glu Gln Ala Lys Asn Gly Arg Gly Leu Val Ala Ile Ala			1056

340	345	350	
ggc gtg gtg cac gat gtc acg gat ttc atc aaa gat cac ccg ggg ggc			1104
Gly Val Val His Asp Val Thr Asp Phe Ile Lys Asp His Pro Gly Gly			
355	360	365	
aag gcc atg atc agc tcc ggg att ggc aag gac gca acc gcc atg ttc			1152
Lys Ala Met Ile Ser Ser Gly Ile Gly Lys Asp Ala Thr Ala Met Phe			
370	375	380	
aat ggg gga gtc tac tac cac agc aac gca gca cac aat ctc ttg tca			1200
Asn Gly Gly Val Tyr Tyr His Ser Asn Ala Ala His Asn Leu Leu Ser			
385	390	395	400
aca atg agg gtg ggt gtt att agg ggc ggc tgt gaa gtc gaa atc tgg			1248
Thr Met Arg Val Gly Val Ile Arg Gly Gly Cys Glu Val Glu Ile Trp			
405	410	415	
aag agg gcg caa aag gag aat gtg gag tac gtg cga gat ggc tct ggt			1296
Lys Arg Ala Gln Lys Glu Asn Val Glu Tyr Val Arg Asp Gly Ser Gly			
420	425	430	
caa cgc gtg atc aga gcg ggc gag cag cca acc aag ata cca gaa ccg			1344
Gln Arg Val Ile Arg Ala Gly Glu Gln Pro Thr Lys Ile Pro Glu Pro			
435	440	445	
att ccc aca gcg gat gcg gcg tag			1368
Ile Pro Thr Ala Asp Ala Ala			
450	455		

<210> 48
 <211> 455
 <212> PRT
 <213> Искусственная последовательность

<220>
 <223> Синтетическая конструкция

<400> 48

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

Arg Ile Trp Leu Ala Ala Val Gly Gly Gly Ala Val Glu Gly Ser Ile
 100 105 110

Arg Trp Trp Ala Arg Asp His Arg Ala His His Arg Tyr Thr Asp Thr
 115 120 125

Asp Lys Asp Pro Tyr Ser Val Arg Lys Gly Leu Leu Tyr Ser His Leu
 130 135 140

Gly Trp Met Val Met Lys Gln Asn Pro Lys Arg Ile Gly Arg Thr Asp
 145 150 155 160

Ile Ser Asp Leu Asn Glu Asp Pro Val Val Val Trp Gln His Arg Asn
 165 170 175

Tyr Leu Lys Val Val Phe Thr Met Gly Leu Ala Val Pro Met Leu Val
 180 185 190

Ala Gly Leu Gly Trp Gly Asp Trp Leu Gly Gly Phe Val Tyr Ala Gly
 195 200 205

Ile Leu Arg Ile Phe Phe Val Gln Gln Ala Thr Phe Cys Val Asn Ser
 210 215 220

Leu Ala His Trp Leu Gly Asp Gln Pro Phe Asp Asp Arg Asn Ser Pro
 225 230 235 240

Arg Asp His Val Ile Thr Ala Leu Val Thr Leu Gly Glu Gly Tyr His
 245 250 255

Asn Phe His His Glu Phe Pro Ser Asp Tyr Arg Asn Ala Ile Glu Trp
 260 265 270

His Gln Tyr Asp Pro Thr Lys Trp Ser Ile Trp Ala Trp Lys Gln Leu
 275 280 285

Gly Leu Ala Tyr Asp Leu Lys Lys Phe Arg Ala Asn Glu Ile Glu Lys
 290 295 300

Gly Arg Val Gln Gln Leu Gln Lys Lys Leu Asp Arg Lys Arg Ala Thr
 305 310 315 320

Leu Asp Trp Gly Thr Pro Leu Asp Gln Leu Pro Val Met Glu Trp Asp
 325 330 335

Asp Tyr Val Glu Gln Ala Lys Asn Gly Arg Gly Leu Val Ala Ile Ala
 340 345 350

Gly Val Val His Asp Val Thr Asp Phe Ile Lys Asp His Pro Gly Gly
 355 360 365

Lys Ala Met Ile Ser Ser Gly Ile Gly Lys Asp Ala Thr Ala Met Phe
 370 375 380

Asn Gly Gly Val Tyr Tyr His Ser Asn Ala Ala His Asn Leu Leu Ser
 385 390 395 400

Thr Met Arg Val Gly Val Ile Arg Gly Gly Cys Glu Val Glu Ile Trp
 405 410 415

Lys Arg Ala Gln Lys Glu Asn Val Glu Tyr Val Arg Asp Gly Ser Gly
 420 425 430

Gln Arg Val Ile Arg Ala Gly Glu Gln Pro Thr Lys Ile Pro Glu Pro
 435 440 445

Ile Pro Thr Ala Asp Ala Ala
 450 455

<210> 49
 <211> 798
 <212> ДНК
 <213> Xerophyta viscosa

<220>
 <221> CDS
 <222> (1)..(798)
 <223> Нативная последовательность ДНК, кодирующая белок SAP1 Xerophyta viscosa

<400> 49
 atg agg aac gag ggt ttt ctg aaa atg aag acc gac gtt gga gtc gcc 48
 Met Arg Asn Glu Gly Phe Leu Lys Met Lys Thr Asp Val Gly Val Ala
 1 5 10 15
 gac gag gtg atc tcc gga gat ctc aag cag ctt ggt gac gct gca aag 96
 Asp Glu Val Ile Ser Gly Asp Leu Lys Gln Leu Gly Asp Ala Ala Lys
 20 25 30
 cgg cta gct aaa cat gcg atc aag ctc ggc gcc agc ttc ggg gtt ggc 144
 Arg Leu Ala Lys His Ala Ile Lys Leu Gly Ala Ser Phe Gly Val Gly
 35 40 45
 tct acc ata gtc cag gct att gct tcg atc gct gct atc tat ttg ttg 192
 Ser Thr Ile Val Gln Ala Ile Ala Ser Ile Ala Ala Ile Tyr Leu Leu
 50 55 60
 ata ttg gac cgg aca aac tgg cgt aca aat atc ttg aca tca ctt cta 240
 Ile Leu Asp Arg Thr Asn Trp Arg Thr Asn Ile Leu Thr Ser Leu Leu
 65 70 75 80
 att cca tat gtt tac ttg agt ctt cct tca gtg ata ttc aac cta ttc 288
 Ile Pro Tyr Val Tyr Leu Ser Leu Pro Ser Val Ile Phe Asn Leu Phe
 85 90 95

agg ggc gac ctg ggc aga tgg ctt tca ttc att ggc gta gta atg aag 336
 Arg Gly Asp Leu Gly Arg Trp Leu Ser Phe Ile Gly Val Val Met Lys
 100 105 110

ctc ttc ttc cac cga cac ttc cca gtt acc ttg gaa ctg ctt gtg tct 384
 Leu Phe Phe His Arg His Phe Pro Val Thr Leu Glu Leu Leu Val Ser
 115 120 125

ctc att ctc ctg att gtg gtt tcc ccc act ttc att gcc cac aca atc 432
 Leu Ile Leu Leu Ile Val Val Ser Pro Thr Phe Ile Ala His Thr Ile
 130 135 140

aga ggc agt ctc att gga gtc ttc atc ttc ctt gtc atc gcc tgc tac 480
 Arg Gly Ser Leu Ile Gly Val Phe Ile Phe Leu Val Ile Ala Cys Tyr
 145 150 155 160

ctc ctc caa gag cac att aga tca gct ggt ggc ttc aaa aac gcg ttc 528
 Leu Leu Gln Glu His Ile Arg Ser Ala Gly Gly Phe Lys Asn Ala Phe
 165 170 175

aca aag agc aat ggg att tca aac agc gtc ggg atc atc att cta ctg 576
 Thr Lys Ser Asn Gly Ile Ser Asn Ser Val Gly Ile Ile Ile Leu Leu
 180 185 190

atc cac ccg atc tgg agc ttg gtg gtg tat ttc ctc tac acg tct ttg 624
 Ile His Pro Ile Trp Ser Leu Val Val Tyr Phe Leu Tyr Thr Ser Leu
 195 200 205

ctg caa ctt ctt gca tac tct cct tcc cct tgt tgt tgc ata tta tac 672
 Leu Gln Leu Leu Ala Tyr Ser Pro Ser Pro Cys Cys Cys Ile Leu Tyr
 210 215 220

aat aag tgg ttt aat ttc atg cat gtt tgt aaa tgt gta agc ctt cat 720
 Asn Lys Trp Phe Asn Phe Met His Val Cys Lys Cys Val Ser Leu His
 225 230 235 240

atg tat tct cag tca att ggg tca tgc gtg tcc ata ttt ttc gtg cag 768
 Met Tyr Ser Gln Ser Ile Gly Ser Cys Val Ser Ile Phe Phe Val Gln
 245 250 255

ttt gta ttc atc tat gaa gct gaa ttt taa 798
 Phe Val Phe Ile Tyr Glu Ala Glu Phe
 260 265

<210> 50
 <211> 265
 <212> PRT
 <213> Xerophyta viscosa

<400> 50

Met Arg Asn Glu Gly Phe Leu Lys Met Lys Thr Asp Val Gly Val Ala
 1 5 10 15

Asp Glu Val Ile Ser Gly Asp Leu Lys Gln Leu Gly Asp Ala Ala Lys
 20 25 30

Arg Leu Ala Lys His Ala Ile Lys Leu Gly Ala Ser Phe Gly Val Gly
 35 40 45

97

Ser Thr Ile Val Gln Ala Ile Ala Ser Ile Ala Ala Ile Tyr Leu Leu
50 55 60

Ile Leu Asp Arg Thr Asn Trp Arg Thr Asn Ile Leu Thr Ser Leu Leu
65 70 75 80

Ile Pro Tyr Val Tyr Leu Ser Leu Pro Ser Val Ile Phe Asn Leu Phe
85 90 95

Arg Gly Asp Leu Gly Arg Trp Leu Ser Phe Ile Gly Val Val Met Lys
100 105 110

Leu Phe Phe His Arg His Phe Pro Val Thr Leu Glu Leu Leu Val Ser
115 120 125

Leu Ile Leu Leu Ile Val Val Ser Pro Thr Phe Ile Ala His Thr Ile
130 135 140

Arg Gly Ser Leu Ile Gly Val Phe Ile Phe Leu Val Ile Ala Cys Tyr
145 150 155 160

Leu Leu Gln Glu His Ile Arg Ser Ala Gly Gly Phe Lys Asn Ala Phe
165 170 175

Thr Lys Ser Asn Gly Ile Ser Asn Ser Val Gly Ile Ile Ile Leu Leu
180 185 190

Ile His Pro Ile Trp Ser Leu Val Val Tyr Phe Leu Tyr Thr Ser Leu
195 200 205

Leu Gln Leu Leu Ala Tyr Ser Pro Ser Pro Cys Cys Cys Ile Leu Tyr
210 215 220

Asn Lys Trp Phe Asn Phe Met His Val Cys Lys Cys Val Ser Leu His
225 230 235 240

Met Tyr Ser Gln Ser Ile Gly Ser Cys Val Ser Ile Phe Phe Val Gln
245 250 255

Phe Val Phe Ile Tyr Glu Ala Glu Phe
260 265

<210> 51

<211> 798

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Синтетическая последовательность ДНК, кодирующая белок SAP1 Xerophyta

viscosa с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1 и таблицы 2

```

<220>
<221> CDS
<222> (1)..(798)

<400> 51
atg aga aac gaa ggt ttt ctg aag atg aaa acg gac gtt ggg gtt gct      48
Met Arg Asn Glu Gly Phe Leu Lys Met Lys Thr Asp Val Gly Val Ala
1          5          10          15

gac gaa gtc atc agc ggt gat ttg aag cag ttg ggt gat gct gcc aaa      96
Asp Glu Val Ile Ser Gly Asp Leu Lys Gln Leu Gly Asp Ala Ala Lys
20          25          30

cgc ctt gct aag cac gct atc aaa ctg gga gcc agc ttt ggt gtt ggt      144
Arg Leu Ala Lys His Ala Ile Lys Leu Gly Ala Ser Phe Gly Val Gly
35          40          45

tca act atc gtt caa gcc atc gca tca ata gca gcc atc tat ctt ctg      192
Ser Thr Ile Val Gln Ala Ile Ala Ser Ile Ala Ala Ile Tyr Leu Leu
50          55          60

att ctc gat agg acc aac tgg agg acc aac atc ttg acg tcc ctc ctc      240
Ile Leu Asp Arg Thr Asn Trp Arg Thr Asn Ile Leu Thr Ser Leu Leu
65          70          75          80

att ccc tac gtg tat ctg tcc ctc ccg agc gtc atc ttc aat ctc ttt      288
Ile Pro Tyr Val Tyr Leu Ser Leu Pro Ser Val Ile Phe Asn Leu Phe
85          90          95

cgt ggg gac ctc ggg aga tgg ctg tca ttc ata ggc gtt gtg atg aag      336
Arg Gly Asp Leu Gly Arg Trp Leu Ser Phe Ile Gly Val Val Met Lys
100         105         110

ctg ttc ttt cat agg cac ttt cct gtt act ttg gag ctg ctt gtg agc      384
Leu Phe Phe His Arg His Phe Pro Val Thr Leu Glu Leu Leu Val Ser
115         120         125

ctc att ctt ttg att gtc gtg tca cct acc ttc ata gct cat aca att      432
Leu Ile Leu Leu Ile Val Val Ser Pro Thr Phe Ile Ala His Thr Ile
130         135         140

cgt gga tct ttg att ggg gtg ttc atc ttc ttg gtg ata gca tgt tat      480
Arg Gly Ser Leu Ile Gly Val Phe Ile Phe Leu Val Ile Ala Cys Tyr
145         150         155         160

ctg ctt caa gag cac att aga tca gct ggt ggc ttc aag aac gcc ttt      528
Leu Leu Gln Glu His Ile Arg Ser Ala Gly Gly Phe Lys Asn Ala Phe
165         170         175

aca aag tct aat gga atc tcc aac agc gtg ggc atc atc atc ctt ctg      576
Thr Lys Ser Asn Gly Ile Ser Asn Ser Val Gly Ile Ile Ile Leu Leu
180         185         190

atc cac ccg att tgg tct ctc gtc gtc tac ttc ctc tac act tca ctt      624
Ile His Pro Ile Trp Ser Leu Val Val Tyr Phe Leu Tyr Thr Ser Leu
195         200         205

ctc cag ctt ttg gcc tac tca cca tcc ccg tgc tgc tgc ata tta tac      672
Leu Gln Leu Leu Ala Tyr Ser Pro Ser Pro Cys Cys Cys Ile Leu Tyr
210         215         220
    
```

aac aag tgg ttc aac ttc atg cat gtt tgc aag tgc gtc tct ttg cac 720
 Asn Lys Trp Phe Asn Phe Met His Val Cys Lys Cys Val Ser Leu His
 225 230 235 240

atg tac tct cag tcc ata ggc tca tgt gtt tca ata ttt ttc gtc cag 768
 Met Tyr Ser Gln Ser Ile Gly Ser Cys Val Ser Ile Phe Phe Val Gln
 245 250 255

ttc gtg ttc atc tat gag gct gag ttt taa 798
 Phe Val Phe Ile Tyr Glu Ala Glu Phe
 260 265

<210> 52
 <211> 265
 <212> PRT
 <213> Искусственная последовательность

<220>
 <223> Синтетическая конструкция

<400> 52

Met Arg Asn Glu Gly Phe Leu Lys Met Lys Thr Asp Val Gly Val Ala
 1 5 10 15

Asp Glu Val Ile Ser Gly Asp Leu Lys Gln Leu Gly Asp Ala Ala Lys
 20 25 30

Arg Leu Ala Lys His Ala Ile Lys Leu Gly Ala Ser Phe Gly Val Gly
 35 40 45

Ser Thr Ile Val Gln Ala Ile Ala Ser Ile Ala Ala Ile Tyr Leu Leu
 50 55 60

Ile Leu Asp Arg Thr Asn Trp Arg Thr Asn Ile Leu Thr Ser Leu Leu
 65 70 75 80

Ile Pro Tyr Val Tyr Leu Ser Leu Pro Ser Val Ile Phe Asn Leu Phe
 85 90 95

Arg Gly Asp Leu Gly Arg Trp Leu Ser Phe Ile Gly Val Val Met Lys
 100 105 110

Leu Phe Phe His Arg His Phe Pro Val Thr Leu Glu Leu Leu Val Ser
 115 120 125

Leu Ile Leu Leu Ile Val Val Ser Pro Thr Phe Ile Ala His Thr Ile
 130 135 140

Arg Gly Ser Leu Ile Gly Val Phe Ile Phe Leu Val Ile Ala Cys Tyr
 145 150 155 160

100

Leu Leu Gln Glu His Ile Arg Ser Ala Gly Gly Phe Lys Asn Ala Phe
 165 170 175

Thr Lys Ser Asn Gly Ile Ser Asn Ser Val Gly Ile Ile Ile Leu Leu
 180 185 190

Ile His Pro Ile Trp Ser Leu Val Val Tyr Phe Leu Tyr Thr Ser Leu
 195 200 205

Leu Gln Leu Leu Ala Tyr Ser Pro Ser Pro Cys Cys Cys Ile Leu Tyr
 210 215 220

Asn Lys Trp Phe Asn Phe Met His Val Cys Lys Cys Val Ser Leu His
 225 230 235 240

Met Tyr Ser Gln Ser Ile Gly Ser Cys Val Ser Ile Phe Phe Val Gln
 245 250 255

Phe Val Phe Ile Tyr Glu Ala Glu Phe
 260 265

<210> 53
 <211> 798
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Синтетическая последовательность ДНК по изобретению, кодирующая белок
 SAP1 Xerophyta viscosa с использованием кодонов, оптимизированных для кукурузы,
 и в которой удаляются последовательности, идентифицированные в таблице 2, и
 поддерживаются последовательности из таблицы 1

<220>
 <221> CDS
 <222> (1)..(798)

<400> 53
 atg aga aac gaa ggt ttt ctg aag atg aaa acg gac gtt ggg gtt gct 48
 Met Arg Asn Glu Gly Phe Leu Lys Met Lys Thr Asp Val Gly Val Ala
 1 5 10 15
 gac gaa gtc atc agc ggt gat ttg aag cag ttg ggt gat gct gcc aaa 96
 Asp Glu Val Ile Ser Gly Asp Leu Lys Gln Leu Gly Asp Ala Ala Lys
 20 25 30
 cgc ctt gct aag cac gct atc aaa ctg gga gcc agc ttt ggt gtt ggt 144
 Arg Leu Ala Lys His Ala Ile Lys Leu Gly Ala Ser Phe Gly Val Gly
 35 40 45
 tca act atc gtt caa gcc atc gca tca ata gca gcc atc tat ctt ctg 192
 Ser Thr Ile Val Gln Ala Ile Ala Ser Ile Ala Ala Ile Tyr Leu Leu
 50 55 60
 att ctc gat agg acc aac tgg agg acc aac atc ttg acg tcc ctc ctc 240
 Ile Leu Asp Arg Thr Asn Trp Arg Thr Asn Ile Leu Thr Ser Leu Leu
 65 70 75 80

101

```

att ccc tac gtg tat ctg tcc ctc ccg agc gtc atc ttc aat ctc ttt      288
Ile Pro Tyr Val Tyr Leu Ser Leu Pro Ser Val Ile Phe Asn Leu Phe
                85                                90                                95

cgt ggg gac ctc ggg aga tgg ctg tca ttc ata ggc gtt gtg atg aag      336
Arg Gly Asp Leu Gly Arg Trp Leu Ser Phe Ile Gly Val Val Met Lys
                100                                105                                110

ctg ttc ttt cat agg cac ttt cct gtt act ttg gag ctg ctt gtg agc      384
Leu Phe Phe His Arg His Phe Pro Val Thr Leu Glu Leu Leu Val Ser
                115                                120                                125

ctc att ctt ttg att gtc gtg tct cct acc ttc ata gct cat aca att      432
Leu Ile Leu Leu Ile Val Val Ser Pro Thr Phe Ile Ala His Thr Ile
                130                                135                                140

cgt gga tct ttg att ggg gtg ttc atc ttc ttg gtg ata gca tgt tat      480
Arg Gly Ser Leu Ile Gly Val Phe Ile Phe Leu Val Ile Ala Cys Tyr
145                                150                                155                                160

ctg ctt caa gag cac att aga tca gct ggt ggc ttc aag aac gcc ttt      528
Leu Leu Gln Glu His Ile Arg Ser Ala Gly Gly Phe Lys Asn Ala Phe
                165                                170                                175

aca aag tct aat gga atc tcc aac agc gtg ggc atc atc atc ctt ctg      576
Thr Lys Ser Asn Gly Ile Ser Asn Ser Val Gly Ile Ile Ile Leu Leu
                180                                185                                190

atc cac ccg att tgg tct ctc gtc gtc tac ttc ctc tac act tca ctt      624
Ile His Pro Ile Trp Ser Leu Val Val Tyr Phe Leu Tyr Thr Ser Leu
                195                                200                                205

ctc cag ctt ttg gcc tac tca cca tcc cca tgc tgc tgt att ctt tac      672
Leu Gln Leu Leu Ala Tyr Ser Pro Ser Pro Cys Cys Cys Ile Leu Tyr
                210                                215                                220

aac aaa tgg ttc aac ttc atg cac gtg tgc aag tgc gtc tct ttg cac      720
Asn Lys Trp Phe Asn Phe Met His Val Cys Lys Cys Val Ser Leu His
225'                                230                                235                                240

atg tac tct cag tcc att ggc tca tgt gtt tca atc ttc ttt gtc cag      768
Met Tyr Ser Gln Ser Ile Gly Ser Cys Val Ser Ile Phe Phe Val Gln
                245                                250                                255

ttc gtg ttc atc tat gag gct gag ttt taa                                798
Phe Val Phe Ile Tyr Glu Ala Glu Phe
                260                                265

```

```

<210> 54
<211> 265
<212> PRT
<213> Искусственная последовательность

```

```

<220>
<223> Синтетическая конструкция

```

```

<400> 54

```

```

Met Arg Asn Glu Gly Phe Leu Lys Met Lys Thr Asp Val Gly Val Ala
1                5                10                15

```

102

Asp Glu Val Ile Ser Gly Asp Leu Lys Gln Leu Gly Asp Ala Ala Lys
 20 25 30

Arg Leu Ala Lys His Ala Ile Lys Leu Gly Ala Ser Phe Gly Val Gly
 35 40 45

Ser Thr Ile Val Gln Ala Ile Ala Ser Ile Ala Ala Ile Tyr Leu Leu
 50 55 60

Ile Leu Asp Arg Thr Asn Trp Arg Thr Asn Ile Leu Thr Ser Leu Leu
 65 70 75 80

Ile Pro Tyr Val Tyr Leu Ser Leu Pro Ser Val Ile Phe Asn Leu Phe
 85 90 95

Arg Gly Asp Leu Gly Arg Trp Leu Ser Phe Ile Gly Val Val Met Lys
 100 105 110

Leu Phe Phe His Arg His Phe Pro Val Thr Leu Glu Leu Leu Val Ser
 115 120 125

Leu Ile Leu Leu Ile Val Val Ser Pro Thr Phe Ile Ala His Thr Ile
 130 135 140

Arg Gly Ser Leu Ile Gly Val Phe Ile Phe Leu Val Ile Ala Cys Tyr
 145 150 155 160

Leu Leu Gln Glu His Ile Arg Ser Ala Gly Gly Phe Lys Asn Ala Phe
 165 170 175

Thr Lys Ser Asn Gly Ile Ser Asn Ser Val Gly Ile Ile Ile Leu Leu
 180 185 190

Ile His Pro Ile Trp Ser Leu Val Val Tyr Phe Leu Tyr Thr Ser Leu
 195 200 205

Leu Gln Leu Leu Ala Tyr Ser Pro Ser Pro Cys Cys Cys Ile Leu Tyr
 210 215 220

Asn Lys Trp Phe Asn Phe Met His Val Cys Lys Cys Val Ser Leu His
 225 230 235 240

Met Tyr Ser Gln Ser Ile Gly Ser Cys Val Ser Ile Phe Phe Val Gln
 245 250 255

Phe Val Phe Ile Tyr Glu Ala Glu Phe
 260 265

```

<210> 55
<211> 717
<212> ДНК
<213> Aequorea victoria

<220>
<221> CDS
<222> (1)..(717)
<223> Нативная последовательность ДНК, кодирующая белок GFP1 Aequorea victoria

<400> 55
atg agt aaa gga gaa gaa ctt ttc act gga gtg gtc cca gtt ctt gtt      48
Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Val Leu Val
1                               5                               10                               15

gaa tta gat ggc gat gtt aat ggg caa aaa ttc tct gtc agt gga gag      96
Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser Gly Glu
                20                               25                               30

ggt gaa ggt gat gca aca tac gga aaa ctt acc ctt aat ttt att tgc     144
Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Asn Phe Ile Cys
                35                               40                               45

act act ggg aag cta cct gtt cca tgg cca aca ctt gtc act act ttc     192
Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
                50                               55                               60

tct tat ggt gtt caa tgc ttc tca aga tac cca gat cat atg aaa cag     240
Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
65                               70                               75                               80

cat gac ttt ttc aag agt gcc atg ccc gaa ggt tat gta cag gaa aga     288
His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
                85                               90                               95

act ata ttt tac aaa gat gac ggg aac tac aag aca cgt gct gaa gtc     336
Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
                100                              105                              110

aag ttt gaa ggt gat acc ctt gtt aat aga atc gag tta aaa ggt att     384
Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
                115                              120                              125

gat ttt aaa gaa gat gga aac att ctt gga cac aaa atg gaa tac aac     432
Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu Tyr Asn
130                              135                              140

tat aac tca cat aat gta tac atc atg gga gac aaa cca aag aat ggc     480
Tyr Asn Ser His Asn Val Tyr Ile Met Gly Asp Lys Pro Lys Asn Gly
145                              150                              155                              160

atc aaa gtt aac ttc aaa att aga cac aac att aaa gat gga agc gtt     528
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly Ser Val
                165                              170                              175

caa tta gca gac cat tat caa caa aat act cca att ggc gat ggc cct     576
Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
                180                              185                              190

gtc ctt tta cca gac aac cat tac ctg tcc aca caa tct gcc ctt tcc     624
Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
195                              200                              205

```

104

aaa gat ccc aac gaa aag aga gat cac atg atc ctt ctt gag ttt gta 672
 Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu Phe Val
 210 215 220

aca gct gct agg att aca cat ggc atg gat gaa cta tac aaa taa 717
 Thr Ala Ala Arg Ile Thr His Gly Met Asp Glu Leu Tyr Lys
 225 230 235

<210> 56
 <211> 238
 <212> PRT
 <213> Aequorea victoria

<400> 56

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Val Leu Val
 1 5 10 15

Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser Gly Glu
 20 25 30

Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Asn Phe Ile Cys
 35 40 45

Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
 50 55 60

Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
 65 70 75 80

His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
 85 90 95

Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
 100 105 110

Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
 115 120 125

Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu Tyr Asn
 130 135 140

Tyr Asn Ser His Asn Val Tyr Ile Met Gly Asp Lys Pro Lys Asn Gly
 145 150 155 160

Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly Ser Val
 165 170 175

Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
 180 185 190

105

Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
 195 200 205

Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu Phe Val
 210 215 220

Thr Ala Ala Arg Ile Thr His Gly Met Asp Glu Leu Tyr Lys
 225 230 235

<210> 57
 <211> 717
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Синтетическая последовательность ДНК, кодирующая белок GFP1 *Aequorea victoria* с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1 и таблицы 2

<220>
 <221> CDS
 <222> (1)..(717)

<400> 57
 atg agt aaa ggg gaa gaa ctt ttc acc ggc gtg gtc cca gtc ctc gtt 48
 Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Val Leu Val
 1 5 10 15
 gag ttg gat ggc gat gtg aat ggg caa aaa ttc tct gtc tcc ggg gag 96
 Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser Gly Glu
 20 25 30
 ggt gag ggt gat gca acc tac gga aag ctg acc cta aat ttt att tgc 144
 Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Asn Phe Ile Cys
 35 40 45
 acg act ggg aag ttg cct gtg cct tgg ccg aca ctg gtg acg acg ttc 192
 Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
 50 55 60
 tct tat ggt gtg cag tgt ttc tca cgc tac ccg gat cat atg aaa cag 240
 Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
 65 70 75 80
 cat gac ttt ttc aag tcg gcc atg cca gaa ggc tat gta caa gag aga 288
 His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
 85 90 95
 act ata ttt tac aag gac gac ggg aac tac aag aca cgt gct gag gtg 336
 Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
 100 105 110
 aag ttc gag ggt gat acc ctt gtt aat cgg atc gag cta aag ggc att 384
 Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
 115 120 125
 gac ttt aag gag gac gga aac att ctg gga cac aaa atg gaa tac aac 432
 Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu Tyr Asn
 130 135 140

106

tat aac tcg cac aac gta tac atc atg gga gac aaa cca aag aat ggc 480
 Tyr Asn Ser His Asn Val Tyr Ile Met Gly Asp Lys Pro Lys Asn Gly
 145 150 155 160

ata aag gtt aac ttc aag att cga cac aac att aaa gac ggc agc gtt 528
 Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly Ser Val
 165 170 175

cag ttg gcc gac cac tat caa caa aat act cca att ggc gat ggc cct 576
 Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
 180 185 190

gtc ctc tta ccc gac aac cat tac ctg tcc acg caa tca gcg ctc agc 624
 Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
 195 200 205

aag gac ccc aac gag aag agg gat cac atg atc ctc ctt gag ttt gtc 672
 Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu Phe Val
 210 215 220

acc gca gct agg ata acc cac ggc atg gat gaa ctg tac aag taa 717
 Thr Ala Ala Arg Ile Thr His Gly Met Asp Glu Leu Tyr Lys
 225 230 235

<210> 58
 <211> 238
 <212> PRT
 <213> Искусственная последовательность

<220>
 <223> Синтетическая конструкция

<400> 58

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Val Leu Val
 1 5 10 15

Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser Gly Glu
 20 25 30

Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Asn Phe Ile Cys
 35 40 45

Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
 50 55 60

Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
 65 70 75 80

His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
 85 90 95

Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
 100 105 110

Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile

107

115

120

125

Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu Tyr Asn
 130 135 140

Tyr Asn Ser His Asn Val Tyr Ile Met Gly Asp Lys Pro Lys Asn Gly
 145 150 155 160

Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly Ser Val
 165 170 175

Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
 180 185 190

Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
 195 200 205

Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu Phe Val
 210 215 220

Thr Ala Ala Arg Ile Thr His Gly Met Asp Glu Leu Tyr Lys
 225 230 235

<210> 59
 <211> 717
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Синтетическая последовательность ДНК по изобретению, кодирующая белок
 GFP1 *Aequorea victoria* с использованием кодонов, оптимизированных для кукурузы,
 и в которой удаляются последовательности, идентифицированные в таблице 2, и
 поддерживаются последовательности из таблицы 1

<220>
 <221> CDS
 <222> (1)..(717)

<400> 59
 atg agt aaa ggg gaa gaa ctt ttc acc ggc gtg gtc cca gtc ctc gtt 48
 Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Val Leu Val
 1 5 10 15

gag ttg gat ggc gat gtg aat ggg caa aaa ttc tct gtc tcc ggg gag 96
 Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser Gly Glu
 20 25 30

ggt gag ggt gat gca acc tac gga aag ctg acc cta aat ttc atc tgc 144
 Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Asn Phe Ile Cys
 35 40 45

acg act ggg aag ttg cct gtg cct tgg ccg aca ctg gtg acg acg ttc 192
 Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
 50 55 60

108

tct tat ggt gtg cag tgt ttc tca cgc tac ccg gat cat atg aaa cag 240
 Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
 65 70 75 80

cat gac ttt ttc aag tcg gcc atg cca gaa ggc tat gta caa gag aga 288
 His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
 85 90 95

act atc ttt tac aag gac gac ggg aac tac aag aca cgt gct gag gtg 336
 Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
 100 105 110

aag ttc gag ggt gat acc ctt gtt aat cgg atc gag cta aag ggc att 384
 Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
 115 120 125

gac ttt aag gag gac gga aac att ctg gga cac aaa atg gaa tac aac 432
 Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu Tyr Asn
 130 135 140

tat aac tcg cac aac gta tac atc atg gga gac aaa cca aag aat ggc 480
 Tyr Asn Ser His Asn Val Tyr Ile Met Gly Asp Lys Pro Lys Asn Gly
 145 150 155 160

ata aag gtt aac ttc aag att cga cac aac att aaa gac ggc agc gtt 528
 Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly Ser Val
 165 170 175

cag ttg gcc gac cac tat caa caa aat act cca att ggc gat ggc cct 576
 Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
 180 185 190

gtc ctc tta ccc gac aac cat tac ctg tcc acg caa tca gcg ctc agc 624
 Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
 195 200 205

aag gac ccc aac gag aag agg gat cac atg atc ctc ctt gag ttt gtc 672
 Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu Phe Val
 210 215 220

acc gca gct agg ata acc cac ggc atg gat gaa ctg tac aag taa 717
 Thr Ala Ala Arg Ile Thr His Gly Met Asp Glu Leu Tyr Lys
 225 230 235

<210> 60
 <211> 238
 <212> PRT
 <213> Искусственная последовательность

<220>
 <223> Синтетическая конструкция

<400> 60

Met Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Val Leu Val
 1 5 10 15

Glu Leu Asp Gly Asp Val Asn Gly Gln Lys Phe Ser Val Ser Gly Glu
 20 25 30

Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Asn Phe Ile Cys

109

35

40

45

Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
50 55 60

Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Gln
65 70 75 80

His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg
85 90 95

Thr Ile Phe Tyr Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
100 105 110

Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
115 120 125

Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Met Glu Tyr Asn
130 135 140

Tyr Asn Ser His Asn Val Tyr Ile Met Gly Asp Lys Pro Lys Asn Gly
145 150 155 160

Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Lys Asp Gly Ser Val
165 170 175

Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly Pro
180 185 190

Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
195 200 205

Lys Asp Pro Asn Glu Lys Arg Asp His Met Ile Leu Leu Glu Phe Val
210 215 220

Thr Ala Ala Arg Ile Thr His Gly Met Asp Glu Leu Tyr Lys
225 230 235

<210> 61

<211> 1350

<212> ДНК

<213> *Leptosphaeria nodorum*

<220>

<221> CDS

<222> (1)..(1350)

<223> Нативная последовательность ДНК, кодирующая белок дельта-9-десатуразы жирных кислот *Leptosphaeria nodorum*

<400> 61

110

atg gcg gcc ttg gac agc att cca gag gat aag gct acc tcg tcg aaa Met Ala Ala Leu Asp Ser Ile Pro Glu Asp Lys Ala Thr Ser Ser Lys 1 5 10 15	48
tcg act cat att caa tat caa gaa gta act ttt cgg aac tgg tat aag Ser Thr His Ile Gln Tyr Gln Glu Val Thr Phe Arg Asn Trp Tyr Lys 20 25 30	96
aag ata aat tgg ctc aac acg acg ctg gtg gtg ctc ata ccc gct ctt Lys Ile Asn Trp Leu Asn Thr Thr Leu Val Val Leu Ile Pro Ala Leu 35 40 45	144
gga ctc tac cta aca cgc acc acg cca ctt aca cga cct acg ctc atc Gly Leu Tyr Leu Thr Arg Thr Thr Pro Leu Thr Arg Pro Thr Leu Ile 50 55 60	192
tgg tcc gtc ctg tac tac ttc tgc aca gct ttc ggc atc aca ggc gga Trp Ser Val Leu Tyr Tyr Phe Cys Thr Ala Phe Gly Ile Thr Gly Gly 65 70 75 80	240
tat cat cga cta tgg agt cat cgc agc tac tcc gct cgt cta ccg cta Tyr His Arg Leu Trp Ser His Arg Ser Tyr Ser Ala Arg Leu Pro Leu 85 90 95	288
cgc tta ttc cta gcc ttc aca ggc gcc gga gcc atc caa ggt agt gct Arg Leu Phe Leu Ala Phe Thr Gly Ala Gly Ala Ile Gln Gly Ser Ala 100 105 110	336
cga tgg tgg agc gca aat cac cgc gcc cac cac cga tgg acc gac aca Arg Trp Trp Ser Ala Asn His Arg Ala His His Arg Trp Thr Asp Thr 115 120 125	384
atg aag gac ccc tac tcc gtt atg cgc ggc cta tta ttc tcg cac atc Met Lys Asp Pro Tyr Ser Val Met Arg Gly Leu Leu Phe Ser His Ile 130 135 140	432
gga tgg atg gta ttg aac agc gac ccc aaa gtc aaa ggc cga aca gac Gly Trp Met Val Leu Asn Ser Asp Pro Lys Val Lys Gly Arg Thr Asp 145 150 155 160	480
gtc agt gat ctc gac agc gac ccc gtc gta gtc tgg cag cac aag cac Val Ser Asp Leu Asp Ser Asp Pro Val Val Val Trp Gln His Lys His 165 170 175	528
tac ggc aag tgc ctg ctg ttc gcc gcg tgg ata ttc ccc atg atc gta Tyr Gly Lys Cys Leu Leu Phe Ala Ala Trp Ile Phe Pro Met Ile Val 180 185 190	576
gcc ggc ctc gga tgg gga gat tgg tgg gga ggc ctt gtc tac gcc ggc Ala Gly Leu Gly Trp Gly Asp Trp Trp Gly Gly Leu Val Tyr Ala Gly 195 200 205	624
atc att cga gcg tgt ttc gtc cag cag gcg aca ttt tgc gtg aac tct Ile Ile Arg Ala Cys Phe Val Gln Gln Ala Thr Phe Cys Val Asn Ser 210 215 220	672
ctc gcg cat tgg atc ggc gag cag ccg ttc gac gac aga cgc acg cct Leu Ala His Trp Ile Gly Glu Gln Pro Phe Asp Asp Arg Arg Thr Pro 225 230 235 240	720
cga gac cac gtt ttg aca gcg ttg gta acg atg gga gaa gga tat cat Arg Asp His Val Leu Thr Ala Leu Val Thr Met Gly Glu Gly Tyr His 245 250 255	768

111

```

aac ttc cac cac gaa ttc cca agc gat tat cgc aac gcg atc atc tgg      816
Asn Phe His His Glu Phe Pro Ser Asp Tyr Arg Asn Ala Ile Ile Trp
      260                      265                      270

tac caa tac gac cct acc aaa tgg ctc att tac ctc ttc tcc ctc ggc      864
Tyr Gln Tyr Asp Pro Thr Lys Trp Leu Ile Tyr Leu Phe Ser Leu Gly
      275                      280                      285

ccc ttc ccc ctc gca tac tcg ctc aaa acc ttc cgg tcc aat gag att      912
Pro Phe Pro Leu Ala Tyr Ser Leu Lys Thr Phe Arg Ser Asn Glu Ile
      290                      295                      300

gaa aaa ggg cgg ttg caa caa caa caa aaa gcc ctg gac aag aag cgc      960
Glu Lys Gly Arg Leu Gln Gln Gln Lys Ala Leu Asp Lys Lys Arg
      305                      310                      315                      320

tca gga ctt gat tgg ggc cta ccc ctc ttc caa ctc cct gtc ata tcg      1008
Ser Gly Leu Asp Trp Gln Leu Pro Leu Phe Gln Leu Pro Val Ile Ser
      325                      330                      335

tgg gac gac ttc caa gcg cgt tgc aaa gag tcc ggc gag atg ctg gtt      1056
Trp Asp Asp Phe Gln Ala Arg Cys Lys Glu Ser Gly Glu Met Leu Val
      340                      345                      350

gct gtc gca ggt gtg att cac gac gtc agc cag ttt att gaa gat cac      1104
Ala Val Ala Gly Val Ile His Asp Val Ser Gln Phe Ile Glu Asp His
      355                      360                      365

cct gga ggc agg agt ttg att cgg agt gcg gtg ggc aaa gat ggg aca      1152
Pro Gly Gly Arg Ser Leu Ile Arg Ser Ala Val Gly Lys Asp Gly Thr
      370                      375                      380

ggg atg ttt aat gga ggc gta tat gag cac agt aat gcg gcg cat aat      1200
Gly Met Phe Asn Gly Gly Val Tyr Glu His Ser Asn Ala Ala His Asn
      385                      390                      395                      400

ctg ttg tcg aca atg agg gtg gga gtg ctt aga ggt ggg cag gag gtg      1248
Leu Leu Ser Thr Met Arg Val Gly Val Leu Arg Gly Gly Gln Glu Val
      405                      410                      415

gag gtg tgg aag aag cag aga gtg gat gtt tta ggg aag agc gac att      1296
Glu Val Trp Lys Lys Gln Arg Val Asp Val Leu Gly Lys Ser Asp Ile
      420                      425                      430

ttg aga cag gtt acg cgg gtg gag agg ttg gtt gag ggg gct gtg gct      1344
Leu Arg Gln Val Thr Arg Val Glu Arg Leu Val Glu Gly Ala Val Ala
      435                      440                      445

gcg tag      1350
Ala

<210> 62
<211> 449
<212> PRT
<213> Leptosphaeria nodorum

<400> 62

Met Ala Ala Leu Asp Ser Ile Pro Glu Asp Lys Ala Thr Ser Ser Lys
1                      5                      10                      15

```

112

Ser Thr His Ile Gln Tyr Gln Glu Val Thr Phe Arg Asn Trp Tyr Lys
 20 25 30

Lys Ile Asn Trp Leu Asn Thr Thr Leu Val Val Leu Ile Pro Ala Leu
 35 40 45

Gly Leu Tyr Leu Thr Arg Thr Thr Pro Leu Thr Arg Pro Thr Leu Ile
 50 55 60

Trp Ser Val Leu Tyr Tyr Phe Cys Thr Ala Phe Gly Ile Thr Gly Gly
 65 70 75 80

Tyr His Arg Leu Trp Ser His Arg Ser Tyr Ser Ala Arg Leu Pro Leu
 85 90 95

Arg Leu Phe Leu Ala Phe Thr Gly Ala Gly Ala Ile Gln Gly Ser Ala
 100 105 110

Arg Trp Trp Ser Ala Asn His Arg Ala His His Arg Trp Thr Asp Thr
 115 120 125

Met Lys Asp Pro Tyr Ser Val Met Arg Gly Leu Leu Phe Ser His Ile
 130 135 140

Gly Trp Met Val Leu Asn Ser Asp Pro Lys Val Lys Gly Arg Thr Asp
 145 150 155 160

Val Ser Asp Leu Asp Ser Asp Pro Val Val Val Trp Gln His Lys His
 165 170 175

Tyr Gly Lys Cys Leu Leu Phe Ala Ala Trp Ile Phe Pro Met Ile Val
 180 185 190

Ala Gly Leu Gly Trp Gly Asp Trp Trp Gly Gly Leu Val Tyr Ala Gly
 195 200 205

Ile Ile Arg Ala Cys Phe Val Gln Gln Ala Thr Phe Cys Val Asn Ser
 210 215 220

Leu Ala His Trp Ile Gly Glu Gln Pro Phe Asp Asp Arg Arg Thr Pro
 225 230 235 240

Arg Asp His Val Leu Thr Ala Leu Val Thr Met Gly Glu Gly Tyr His
 245 250 255

Asn Phe His His Glu Phe Pro Ser Asp Tyr Arg Asn Ala Ile Ile Trp
 260 265 270

113

Tyr Gln Tyr Asp Pro Thr Lys Trp Leu Ile Tyr Leu Phe Ser Leu Gly
 275 280 285
 Pro Phe Pro Leu Ala Tyr Ser Leu Lys Thr Phe Arg Ser Asn Glu Ile
 290 295 300
 Glu Lys Gly Arg Leu Gln Gln Gln Lys Ala Leu Asp Lys Lys Arg
 305 310 315 320
 Ser Gly Leu Asp Trp Gly Leu Pro Leu Phe Gln Leu Pro Val Ile Ser
 325 330 335
 Trp Asp Asp Phe Gln Ala Arg Cys Lys Glu Ser Gly Glu Met Leu Val
 340 345 350
 Ala Val Ala Gly Val Ile His Asp Val Ser Gln Phe Ile Glu Asp His
 355 360 365
 Pro Gly Gly Arg Ser Leu Ile Arg Ser Ala Val Gly Lys Asp Gly Thr
 370 375 380
 Gly Met Phe Asn Gly Gly Val Tyr Glu His Ser Asn Ala Ala His Asn
 385 390 395 400
 Leu Leu Ser Thr Met Arg Val Gly Val Leu Arg Gly Gly Gln Glu Val
 405 410 415
 Glu Val Trp Lys Lys Gln Arg Val Asp Val Leu Gly Lys Ser Asp Ile
 420 425 430
 Leu Arg Gln Val Thr Arg Val Glu Arg Leu Val Glu Gly Ala Val Ala
 435 440 445

Ala

<210> 63
 <211> 1350
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Синтетическая последовательность ДНК, кодирующая белок дельта-9-
 десатуразы жирных кислот *Leptosphaeria podorum* с использованием кодонов,
 оптимизированных для кукурузы, и в которой поддерживаются последовательности из
 таблицы 1 и таблицы 2

<220>
 <221> CDS
 <222> (1)..(1350)

<400> 63
atg gca gcc ctt gac agc atc cca gag gat aag gct acc tcg tct aaa 48
Met Ala Ala Leu Asp Ser Ile Pro Glu Asp Lys Ala Thr Ser Ser Lys
1 5 10 15

tcg act cat att cag tac caa gaa gtg act ttt cgg aac tgg tac aaa 96
Ser Thr His Ile Gln Tyr Gln Glu Val Thr Phe Arg Asn Trp Tyr Lys
20 25 30

aag ata aac tgg ctc aac acg acg ctg gtg gtg ctc ata cca gct ctt 144
Lys Ile Asn Trp Leu Asn Thr Thr Leu Val Val Leu Ile Pro Ala Leu
35 40 45

ggc ctt tac cta aca agg acc acg cca ctt act agg cca acg ctc atc 192
Gly Leu Tyr Leu Thr Arg Thr Thr Pro Leu Thr Arg Pro Thr Leu Ile
50 55 60

tgg tcc gtc ctg tac tac ttt tgc acc gct ttc ggc att acc ggc gga 240
Trp Ser Val Leu Tyr Tyr Phe Cys Thr Ala Phe Gly Ile Thr Gly Gly
65 70 75 80

tat cat aga cta tgg agt cat cgc agc tac tcc gct cgt cta ccg ctt 288
Tyr His Arg Leu Trp Ser His Arg Ser Tyr Ser Ala Arg Leu Pro Leu
85 90 95

cgc ttg ttc ctg gcc ttc act ggc gcc ggg gcc atc caa ggt tca gct 336
Arg Leu Phe Leu Ala Phe Thr Gly Ala Gly Ala Ile Gln Gly Ser Ala
100 105 110

agg tgg tgg agc gca aat cac cgc gcc cat cat agg tgg acc gac aca 384
Arg Trp Trp Ser Ala Asn His Arg Ala His His Arg Trp Thr Asp Thr
115 120 125

atg aag gac ccc tac tcc gtt atg cgc ggt cta tta ttc tcg cac atc 432
Met Lys Asp Pro Tyr Ser Val Met Arg Gly Leu Leu Phe Ser His Ile
130 135 140

ggc tgg atg gtt cta aac agc gac ccc aaa gtc aaa ggc cgc act gac 480
Gly Trp Met Val Leu Asn Ser Asp Pro Lys Val Lys Gly Arg Thr Asp
145 150 155 160

gtc tca gac cta gat agc gac ccc gtc gtt gtc tgg cag cac aag cac 528
Val Ser Asp Leu Asp Ser Asp Pro Val Val Val Trp Gln His Lys His
165 170 175

tac ggc aag tgc ctg cta ttt gcc gca tgg ata ttc ccg atg atc gta 576
Tyr Gly Lys Cys Leu Leu Phe Ala Ala Trp Ile Phe Pro Met Ile Val
180 185 190

gcc ggc ctc gga tgg gga gat tgg tgg gga ggc ctt gtc tac gcc ggc 624
Ala Gly Leu Gly Trp Gly Asp Trp Trp Gly Gly Leu Val Tyr Ala Gly
195 200 205

atc att agg gcg tgt ttc gtc cag caa gca acc ttt tgc gtg aac tct 672
Ile Ile Arg Ala Cys Phe Val Gln Gln Ala Thr Phe Cys Val Asn Ser
210 215 220

ctc gcg cac tgg atc ggc gag cag ccg ttc gac gac aga cgc acc cct 720
Leu Ala His Trp Ile Gly Glu Gln Pro Phe Asp Asp Arg Arg Thr Pro
225 230 235 240

aga gac cac gtt ttg acc gcg ttg gtc act atg gga gaa ggt tat cac 768
Arg Asp His Val Leu Thr Ala Leu Val Thr Met Gly Glu Gly Tyr His
245 250 255

115

```

aac ttc cac cac gag ttc ccg tct gat tat agg aac gcg atc atc tgg      816
Asn Phe His His Glu Phe Pro Ser Asp Tyr Arg Asn Ala Ile Ile Trp
                260                265                270

tat cag tac gac cct acc aaa tgg ctc ata tac ctc ttc tcc ctc ggc      864
Tyr Gln Tyr Asp Pro Thr Lys Trp Leu Ile Tyr Leu Phe Ser Leu Gly
                275                280                285

ccg ttc cca ctg gca tac tcg ctc aaa acc ttc cgg tct aac gag atc      912
Pro Phe Pro Leu Ala Tyr Ser Leu Lys Thr Phe Arg Ser Asn Glu Ile
                290                295                300

gaa aag ggg cgg ttg caa caa caa caa aag gcc ctg gat aag aag cgc      960
Glu Lys Gly Arg Leu Gln Gln Gln Gln Lys Ala Leu Asp Lys Lys Arg
305                310                315                320

tct ggc ctt gat tgg ggc ctg ccc ctc ttc cag ctc cct gtg ata tct      1008
Ser Gly Leu Asp Trp Gly Leu Pro Leu Phe Gln Leu Pro Val Ile Ser
                325                330                335

tgg gac gac ttc caa gcg cgt tgt aag gag tcc ggc gag atg ctg gtt      1056
Trp Asp Asp Phe Gln Ala Arg Cys Lys Glu Ser Gly Glu Met Leu Val
                340                345                350

gct gtc gcc ggt gtg att cac gac gtc tca cag ttt att gaa gat cac      1104
Ala Val Ala Gly Val Ile His Asp Val Ser Gln Phe Ile Glu Asp His
                355                360                365

cct gga ggg agg agt ctg att cgg tct gcg gtg ggc aag gat ggg act      1152
Pro Gly Gly Arg Ser Leu Ile Arg Ser Ala Val Gly Lys Asp Gly Thr
370                375                380

ggg atg ttt aat gga ggc gtt tat gag cac agt aat gcg gcg cac aat      1200
Gly Met Phe Asn Gly Gly Val Tyr Glu His Ser Asn Ala Ala His Asn
385                390                395                400

ctg ttg tca aca atg agg gtg ggt gtg ctt aga ggt ggg caa gag gtg      1248
Leu Leu Ser Thr Met Arg Val Gly Val Leu Arg Gly Gly Gln Glu Val
                405                410                415

gag gtg tgg aag aag cag cgt gtg gat gtt tta ggg aag agc gat atc      1296
Glu Val Trp Lys Lys Gln Arg Val Asp Val Leu Gly Lys Ser Asp Ile
                420                425                430

ttg cgt caa gtt acg cgg gtg gag agg ctg gtt gag ggg gct gtg gct      1344
Leu Arg Gln Val Thr Arg Val Glu Arg Leu Val Glu Gly Ala Val Ala
                435                440                445

gcc tag      1350
Ala

```

```

<210> 64
<211> 449
<212> PRT
<213> Искусственная последовательность

<220>
<223> Синтетическая конструкция

<400> 64

```

116

Met Ala Ala Leu Asp Ser Ile Pro Glu Asp Lys Ala Thr Ser Ser Lys
 1 5 10 15

Ser Thr His Ile Gln Tyr Gln Glu Val Thr Phe Arg Asn Trp Tyr Lys
 20 25 30

Lys Ile Asn Trp Leu Asn Thr Thr Leu Val Val Leu Ile Pro Ala Leu
 35 40 45

Gly Leu Tyr Leu Thr Arg Thr Thr Pro Leu Thr Arg Pro Thr Leu Ile
 50 55 60

Trp Ser Val Leu Tyr Tyr Phe Cys Thr Ala Phe Gly Ile Thr Gly Gly
 65 70 75 80

Tyr His Arg Leu Trp Ser His Arg Ser Tyr Ser Ala Arg Leu Pro Leu
 85 90 95

Arg Leu Phe Leu Ala Phe Thr Gly Ala Gly Ala Ile Gln Gly Ser Ala
 100 105 110

Arg Trp Trp Ser Ala Asn His Arg Ala His His Arg Trp Thr Asp Thr
 115 120 125

Met Lys Asp Pro Tyr Ser Val Met Arg Gly Leu Leu Phe Ser His Ile
 130 135 140

Gly Trp Met Val Leu Asn Ser Asp Pro Lys Val Lys Gly Arg Thr Asp
 145 150 155 160

Val Ser Asp Leu Asp Ser Asp Pro Val Val Val Trp Gln His Lys His
 165 170 175

Tyr Gly Lys Cys Leu Leu Phe Ala Ala Trp Ile Phe Pro Met Ile Val
 180 185 190

Ala Gly Leu Gly Trp Gly Asp Trp Trp Gly Gly Leu Val Tyr Ala Gly
 195 200 205

Ile Ile Arg Ala Cys Phe Val Gln Gln Ala Thr Phe Cys Val Asn Ser
 210 215 220

Leu Ala His Trp Ile Gly Glu Gln Pro Phe Asp Asp Arg Arg Thr Pro
 225 230 235 240

Arg Asp His Val Leu Thr Ala Leu Val Thr Met Gly Glu Gly Tyr His
 245 250 255

117

Asn Phe His His Glu Phe Pro Ser Asp Tyr Arg Asn Ala Ile Ile Trp
 260 265 270

Tyr Gln Tyr Asp Pro Thr Lys Trp Leu Ile Tyr Leu Phe Ser Leu Gly
 275 280 285

Pro Phe Pro Leu Ala Tyr Ser Leu Lys Thr Phe Arg Ser Asn Glu Ile
 290 295 300

Glu Lys Gly Arg Leu Gln Gln Gln Lys Ala Leu Asp Lys Lys Arg
 305 310 315 320

Ser Gly Leu Asp Trp Gly Leu Pro Leu Phe Gln Leu Pro Val Ile Ser
 325 330 335

Trp Asp Asp Phe Gln Ala Arg Cys Lys Glu Ser Gly Glu Met Leu Val
 340 345 350

Ala Val Ala Gly Val Ile His Asp Val Ser Gln Phe Ile Glu Asp His
 355 360 365

Pro Gly Gly Arg Ser Leu Ile Arg Ser Ala Val Gly Lys Asp Gly Thr
 370 375 380

Gly Met Phe Asn Gly Gly Val Tyr Glu His Ser Asn Ala Ala His Asn
 385 390 395 400

Leu Leu Ser Thr Met Arg Val Gly Val Leu Arg Gly Gly Gln Glu Val
 405 410 415

Glu Val Trp Lys Lys Gln Arg Val Asp Val Leu Gly Lys Ser Asp Ile
 420 425 430

Leu Arg Gln Val Thr Arg Val Glu Arg Leu Val Glu Gly Ala Val Ala
 435 440 445

Ala

<210> 65

<211> 1350

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Синтетическая последовательность ДНК по изобретению, кодирующая белок дельта-9-десатуразы жирных кислот *Leptosphaeria podorum* с использованием кодонов, оптимизированных для кукурузы, и в которой удаляются последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1

```

<220>
<221> CDS
<222> (1)..(1350)

<400> 65
atg gca gcc ctt gac agc atc cca gag gat aag gct acc tcg tct aaa      48
Met Ala Ala Leu Asp Ser Ile Pro Glu Asp Lys Ala Thr Ser Ser Lys
1          5          10          15

tcg act cat att cag tac caa gaa gtg act ttt cgg aac tgg tac aaa      96
Ser Thr His Ile Gln Tyr Gln Glu Val Thr Phe Arg Asn Trp Tyr Lys
20          25          30

aag ata aac tgg ctc aac acg acg ctg gtg gtg ctc ata cca gct ctt     144
Lys Ile Asn Trp Leu Asn Thr Thr Leu Val Val Leu Ile Pro Ala Leu
35          40          45

ggt ctt tac cta aca agg acc acg cca ctt act agg cca acg ctc atc     192
Gly Leu Tyr Leu Thr Arg Thr Thr Pro Leu Thr Arg Pro Thr Leu Ile
50          55          60

tgg tcc gtc ctg tac tac ttt tgc acc gct ttc ggc att acc ggc gga     240
Trp Ser Val Leu Tyr Tyr Phe Cys Thr Ala Phe Gly Ile Thr Gly Gly
65          70          75          80

tat cat aga cta tgg agt cat cgc agc tac tcc gct cgt cta ccg ctt     288
Tyr His Arg Leu Trp Ser His Arg Ser Tyr Ser Ala Arg Leu Pro Leu
85          90          95

cgc ttg ttc ctg gcc ttc act ggc gcc ggg gcc atc caa ggt tca gct     336
Arg Leu Phe Leu Ala Phe Thr Gly Ala Gly Ala Ile Gln Gly Ser Ala
100         105         110

agg tgg tgg agc gca aat cac cgc gcc cat cat agg tgg acc gac aca     384
Arg Trp Trp Ser Ala Asn His Arg Ala His His Arg Trp Thr Asp Thr
115         120         125

atg aag gac ccc tac tcc gtt atg cgc ggt ctg tta ttc tcg cac atc     432
Met Lys Asp Pro Tyr Ser Val Met Arg Gly Leu Leu Phe Ser His Ile
130         135         140

ggt tgg atg gtt cta aac agc gac ccc aaa gtc aaa ggc cgc act gac     480
Gly Trp Met Val Leu Asn Ser Asp Pro Lys Val Lys Gly Arg Thr Asp
145         150         155         160

gtc tca gac cta gat agc gac ccc gtc gtt gtc tgg cag cac aag cac     528
Val Ser Asp Leu Asp Ser Asp Pro Val Val Val Trp Gln His Lys His
165         170         175

tac ggc aag tgc ctg cta ttt gcc gca tgg ata ttc ccg atg atc gta     576
Tyr Gly Lys Cys Leu Leu Phe Ala Ala Trp Ile Phe Pro Met Ile Val
180         185         190

gcc ggc ctc gga tgg gga gat tgg tgg gga ggc ctt gtc tac gcc ggc     624
Ala Gly Leu Gly Trp Gly Asp Trp Trp Gly Gly Leu Val Tyr Ala Gly
195         200         205

atc att agg gcg tgt ttc gtc cag caa gca acc ttt tgc gtg aac tct     672
Ile Ile Arg Ala Cys Phe Val Gln Gln Ala Thr Phe Cys Val Asn Ser
210         215         220

ctc gcg cac tgg atc ggc gag cag ccg ttc gac gac aga cgc acc cct     720
Leu Ala His Trp Ile Gly Glu Gln Pro Phe Asp Asp Arg Arg Thr Pro

```

119

225	230	235	240	
aga gac cac gtt ttg acc gcg ttg gtc act atg gga gaa ggt tat cac				768
Arg Asp His Val Leu Thr Ala Leu Val Thr Met Gly Glu Gly Tyr His	245	250	255	
aac ttc cac cac gag ttc ccg tct gat tat agg aac gcg atc atc tgg				816
Asn Phe His His Glu Phe Pro Ser Asp Tyr Arg Asn Ala Ile Ile Trp	260	265	270	
tat cag tac gac cct acc aaa tgg ctc ata tac ctc ttc tcc ctc ggc				864
Tyr Gln Tyr Asp Pro Thr Lys Trp Leu Ile Tyr Leu Phe Ser Leu Gly	275	280	285	
ccg ttc cca ctg gca tac tcg ctc aaa acc ttc cgg tct aac gag atc				912
Pro Phe Pro Leu Ala Tyr Ser Leu Lys Thr Phe Arg Ser Asn Glu Ile	290	295	300	
gaa aag ggg cgg ttg caa caa caa caa aag gcc ctg gat aag aag cgc				960
Glu Lys Gly Arg Leu Gln Gln Gln Gln Lys Ala Leu Asp Lys Lys Arg	305	310	315	320
tct ggc ctt gat tgg ggc ctg ccc ctc ttc cag ctc cct gtg ata tct				1008
Ser Gly Leu Asp Trp Gly Leu Pro Leu Phe Gln Leu Pro Val Ile Ser	325	330	335	
tgg gac gac ttc caa gcg cgt tgt aag gag tcc ggc gag atg ctg gtt				1056
Trp Asp Asp Phe Gln Ala Arg Cys Lys Glu Ser Gly Glu Met Leu Val	340	345	350	
gct gtc gcc ggt gtg att cac gac gtc tca cag ttc att gaa gat cac				1104
Ala Val Ala Gly Val Ile His Asp Val Ser Gln Phe Ile Glu Asp His	355	360	365	
cct gga ggg agg agt ctg att cgg tct gcg gtg ggc aag gat ggg act				1152
Pro Gly Gly Arg Ser Leu Ile Arg Ser Ala Val Gly Lys Asp Gly Thr	370	375	380	
ggg atg ttt aat gga ggc gtt tat gag cac agt aat gcg gcg cac aat				1200
Gly Met Phe Asn Gly Gly Val Tyr Glu His Ser Asn Ala Ala His Asn	385	390	395	400
ctg ttg tca aca atg agg gtg ggt gtg ctt aga ggt ggg caa gag gtg				1248
Leu Leu Ser Thr Met Arg Val Gly Val Leu Arg Gly Gly Gln Glu Val	405	410	415	
gag gtg tgg aag aag cag cgt gtg gat gta tta ggg aag agc gat atc				1296
Glu Val Trp Lys Lys Gln Arg Val Asp Val Leu Gly Lys Ser Asp Ile	420	425	430	
ttg cgt caa gtt acg cgg gtg gag agg ctg gtt gag ggg gct gtg gct				1344
Leu Arg Gln Val Thr Arg Val Glu Arg Leu Val Glu Gly Ala Val Ala	435	440	445	
gcc tag				1350
Ala				

<210> 66

<211> 449

<212> PRT

<213> Искусственная последовательность

120

<220>

<223> Синтетическая конструкция

<400> 66

Met Ala Ala Leu Asp Ser Ile Pro Glu Asp Lys Ala Thr Ser Ser Lys
 1 5 10 15

Ser Thr His Ile Gln Tyr Gln Glu Val Thr Phe Arg Asn Trp Tyr Lys
 20 25 30

Lys Ile Asn Trp Leu Asn Thr Thr Leu Val Val Leu Ile Pro Ala Leu
 35 40 45

Gly Leu Tyr Leu Thr Arg Thr Thr Pro Leu Thr Arg Pro Thr Leu Ile
 50 55 60

Trp Ser Val Leu Tyr Tyr Phe Cys Thr Ala Phe Gly Ile Thr Gly Gly
 65 70 75 80

Tyr His Arg Leu Trp Ser His Arg Ser Tyr Ser Ala Arg Leu Pro Leu
 85 90 95

Arg Leu Phe Leu Ala Phe Thr Gly Ala Gly Ala Ile Gln Gly Ser Ala
 100 105 110

Arg Trp Trp Ser Ala Asn His Arg Ala His His Arg Trp Thr Asp Thr
 115 120 125

Met Lys Asp Pro Tyr Ser Val Met Arg Gly Leu Leu Phe Ser His Ile
 130 135 140

Gly Trp Met Val Leu Asn Ser Asp Pro Lys Val Lys Gly Arg Thr Asp
 145 150 155 160

Val Ser Asp Leu Asp Ser Asp Pro Val Val Val Trp Gln His Lys His
 165 170 175

Tyr Gly Lys Cys Leu Leu Phe Ala Ala Trp Ile Phe Pro Met Ile Val
 180 185 190

Ala Gly Leu Gly Trp Gly Asp Trp Trp Gly Gly Leu Val Tyr Ala Gly
 195 200 205

Ile Ile Arg Ala Cys Phe Val Gln Gln Ala Thr Phe Cys Val Asn Ser
 210 215 220

Leu Ala His Trp Ile Gly Glu Gln Pro Phe Asp Asp Arg Arg Thr Pro
 225 230 235 240

121

Arg Asp His Val Leu Thr Ala Leu Val Thr Met Gly Glu Gly Tyr His
245 250 255

Asn Phe His His Glu Phe Pro Ser Asp Tyr Arg Asn Ala Ile Ile Trp
260 265 270

Tyr Gln Tyr Asp Pro Thr Lys Trp Leu Ile Tyr Leu Phe Ser Leu Gly
275 280 285

Pro Phe Pro Leu Ala Tyr Ser Leu Lys Thr Phe Arg Ser Asn Glu Ile
290 295 300

Glu Lys Gly Arg Leu Gln Gln Gln Gln Lys Ala Leu Asp Lys Lys Arg
305 310 315 320

Ser Gly Leu Asp Trp Gly Leu Pro Leu Phe Gln Leu Pro Val Ile Ser
325 330 335

Trp Asp Asp Phe Gln Ala Arg Cys Lys Glu Ser Gly Glu Met Leu Val
340 345 350

Ala Val Ala Gly Val Ile His Asp Val Ser Gln Phe Ile Glu Asp His
355 360 365

Pro Gly Gly Arg Ser Leu Ile Arg Ser Ala Val Gly Lys Asp Gly Thr
370 375 380

Gly Met Phe Asn Gly Gly Val Tyr Glu His Ser Asn Ala Ala His Asn
385 390 395 400

Leu Leu Ser Thr Met Arg Val Gly Val Leu Arg Gly Gly Gln Glu Val
405 410 415

Glu Val Trp Lys Lys Gln Arg Val Asp Val Leu Gly Lys Ser Asp Ile
420 425 430

Leu Arg Gln Val Thr Arg Val Glu Arg Leu Val Glu Gly Ala Val Ala
435 440 445

Ala

<210> 67
 <211> 660
 <212> ДНК
 <213> Xerophyta viscosa

<220>

<221> CDS
 <222> (1)..(660)
 <223> Нативная последовательность ДНК, кодирующая белок PER1 Xerophyta viscosa

```

<400> 67
atg ccg ggg ctc acc att ggc gac acg atc ccc aac ctg gag ctt gac      48
Met Pro Gly Leu Thr Ile Gly Asp Thr Ile Pro Asn Leu Glu Leu Asp
1           5           10           15

acc acc cag ggt agg atc aaa atc cac gat tac gtc ggc aac ggc tac      96
Thr Thr Gln Gly Arg Ile Lys Ile His Asp Tyr Val Gly Asn Gly Tyr
           20           25           30

gtc atc ttg ttc tca cac cct gga gac ttc act cct gtc tgc acc acc     144
Val Ile Leu Phe Ser His Pro Gly Asp Phe Thr Pro Val Cys Thr Thr
           35           40           45

gaa ctt gga aag atg gct gct tac gcc gac gag ttc agc aag cgc ggg     192
Glu Leu Gly Lys Met Ala Ala Tyr Ala Asp Glu Phe Ser Lys Arg Gly
           50           55           60

gtt aag ctt ctt ggt ctt tcc tgc gac gat gta cag agc cac aag gag     240
Val Lys Leu Leu Gly Leu Ser Cys Asp Asp Val Gln Ser His Lys Glu
65           70           75           80

tgg atc aag gat atc gaa gcc tat acg ccg gga tgt cac gta aaa tat     288
Trp Ile Lys Asp Ile Glu Ala Tyr Thr Pro Gly Cys His Val Lys Tyr
           85           90           95

cct atc gcg gcg gac cca acc cgc gag att atc cag cag cta aac atg     336
Pro Ile Ala Ala Asp Pro Thr Arg Glu Ile Ile Gln Gln Leu Asn Met
           100          105          110

gta gac cca gac gag aca gag tcc agc aaa tgc gcc gtg cct tcg cga     384
Val Asp Pro Asp Glu Thr Glu Ser Ser Lys Cys Ala Val Pro Ser Arg
           115          120          125

gct ctg cac atc att ggg ccc gac aag agg atc aag ctg agt ttc ctg     432
Ala Leu His Ile Ile Gly Pro Asp Lys Arg Ile Lys Leu Ser Phe Leu
           130          135          140

tac ccc gcg tcg acg ggg cga aac atg gat gag gtg ctg agg gca gtg     480
Tyr Pro Ala Ser Thr Gly Arg Asn Met Asp Glu Val Leu Arg Ala Val
145          150          155          160

gag tcg ctc cag cag gcg gca aag cac aag gtg gca acg ccg gcg aac     528
Glu Ser Leu Gln Gln Ala Ala Lys His Lys Val Ala Thr Pro Ala Asn
           165          170          175

tgg aag cct ggt gaa cct gtt gtg atc aag cct gat gtg tcc agc gag     576
Trp Lys Pro Gly Glu Pro Val Val Ile Lys Pro Asp Val Ser Ser Glu
           180          185          190

gag gcc aag aag ctt ttc ccg cag ggt tat aaa agt gtt gat ctt cca     624
Glu Ala Lys Lys Leu Phe Pro Gln Gly Tyr Lys Ser Val Asp Leu Pro
           195          200          205

tcc aag aag gat tac ctt cgt ttt acg aac gtc tga                       660
Ser Lys Lys Asp Tyr Leu Arg Phe Thr Asn Val
           210          215

<210> 68
<211> 219
    
```

123

<212> PRT

<213> Xerophyta viscosa

<400> 68

Met Pro Gly Leu Thr Ile Gly Asp Thr Ile Pro Asn Leu Glu Leu Asp
 1 5 10 15

Thr Thr Gln Gly Arg Ile Lys Ile His Asp Tyr Val Gly Asn Gly Tyr
 20 25 30

Val Ile Leu Phe Ser His Pro Gly Asp Phe Thr Pro Val Cys Thr Thr
 35 40 45

Glu Leu Gly Lys Met Ala Ala Tyr Ala Asp Glu Phe Ser Lys Arg Gly
 50 55 60

Val Lys Leu Leu Gly Leu Ser Cys Asp Asp Val Gln Ser His Lys Glu
 65 70 75 80

Trp Ile Lys Asp Ile Glu Ala Tyr Thr Pro Gly Cys His Val Lys Tyr
 85 90 95

Pro Ile Ala Ala Asp Pro Thr Arg Glu Ile Ile Gln Gln Leu Asn Met
 100 105 110

Val Asp Pro Asp Glu Thr Glu Ser Ser Lys Cys Ala Val Pro Ser Arg
 115 120 125

Ala Leu His Ile Ile Gly Pro Asp Lys Arg Ile Lys Leu Ser Phe Leu
 130 135 140

Tyr Pro Ala Ser Thr Gly Arg Asn Met Asp Glu Val Leu Arg Ala Val
 145 150 155 160

Glu Ser Leu Gln Gln Ala Ala Lys His Lys Val Ala Thr Pro Ala Asn
 165 170 175

Trp Lys Pro Gly Glu Pro Val Val Ile Lys Pro Asp Val Ser Ser Glu
 180 185 190

Glu Ala Lys Lys Leu Phe Pro Gln Gly Tyr Lys Ser Val Asp Leu Pro
 195 200 205

Ser Lys Lys Asp Tyr Leu Arg Phe Thr Asn Val
 210 215

<210> 69

<211> 660

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Синтетическая последовательность ДНК, кодирующая белок PER1 Xerophyta viscosa с использованием кодонов, оптимизированных для кукурузы, и в которой поддерживаются последовательности из таблицы 1 и таблицы 2

<220>

<221> CDS

<222> (1)..(660)

<400> 69

atg cct gga ttg act att ggt gac aca att ccc aac ttg gag ctg gat	48
Met Pro Gly Leu Thr Ile Gly Asp Thr Ile Pro Asn Leu Glu Leu Asp	
1 5 10 15	
acg aca caa ggt cgc atc aag atc cac gac tat gtc ggg aat gga tac	96
Thr Thr Gln Gly Arg Ile Lys Ile His Asp Tyr Val Gly Asn Gly Tyr	
20 25 30	
gtg att ctc ttc tca cat cct ggt gat ttc act ccg gtg tgt acc acc	144
Val Ile Leu Phe Ser His Pro Gly Asp Phe Thr Pro Val Cys Thr Thr	
35 40 45	
gaa ttg ggc aag atg gct gct tat gcc gac gag ttc tct aag cgt ggt	192
Glu Leu Gly Lys Met Ala Ala Tyr Ala Asp Glu Phe Ser Lys Arg Gly	
50 55 60	
gtg aag ctg ctt ggg ttg tcc tgt gat gat gtc caa tca cat aag gag	240
Val Lys Leu Leu Gly Leu Ser Cys Asp Asp Val Gln Ser His Lys Glu	
65 70 75 80	
tgg atc aaa gac ata gag gct tac aca cct ggc tgt cac gta aaa tat	288
Trp Ile Lys Asp Ile Glu Ala Tyr Thr Pro Gly Cys His Val Lys Tyr	
85 90 95	
ccg att gct gct gat cca acc aga gaa atc ata cag cag ctg aac atg	336
Pro Ile Ala Ala Asp Pro Thr Arg Glu Ile Ile Gln Gln Leu Asn Met	
100 105 110	
gtg gac cct gat gag acg gaa agc tct aag tgc gct gtg cct tct agg	384
Val Asp Pro Asp Glu Thr Glu Ser Ser Lys Cys Ala Val Pro Ser Arg	
115 120 125	
gca ctt cac atc ata gga cca gat aag agg atc aag ctg tcc ttc ctc	432
Ala Leu His Ile Ile Gly Pro Asp Lys Arg Ile Lys Leu Ser Phe Leu	
130 135 140	
tac cct gcc tct act ggt cgc aac atg gac gaa gtt ctt aga gcc gtt	480
Tyr Pro Ala Ser Thr Gly Arg Asn Met Asp Glu Val Leu Arg Ala Val	
145 150 155 160	
gag tct ctt cag caa gca gct aaa cac aaa gtt gca act cct gct aac	528
Glu Ser Leu Gln Gln Ala Ala Lys His Lys Val Ala Thr Pro Ala Asn	
165 170 175	
tgg aaa cct ggc gaa cca gtc gtc atc aaa cca gac gtc agc tcc gag	576
Trp Lys Pro Gly Glu Pro Val Val Ile Lys Pro Asp Val Ser Ser Glu	
180 185 190	
gag gcc aag aag ctc ttt cct caa ggt tat aaa agc gtt gat ttg cct	624
Glu Ala Lys Lys Leu Phe Pro Gln Gly Tyr Lys Ser Val Asp Leu Pro	
195 200 205	

125

tca aag aag gac tac ttg agg ttc acc aat gtt tga
 Ser Lys Lys Asp Tyr Leu Arg Phe Thr Asn Val
 210 215

660

<210> 70
 <211> 219
 <212> PRT
 <213> Искусственная последовательность

<220>
 <223> Синтетическая конструкция

<400> 70

Met Pro Gly Leu Thr Ile Gly Asp Thr Ile Pro Asn Leu Glu Leu Asp
 1 5 10 15

Thr Thr Gln Gly Arg Ile Lys Ile His Asp Tyr Val Gly Asn Gly Tyr
 20 25 30

Val Ile Leu Phe Ser His Pro Gly Asp Phe Thr Pro Val Cys Thr Thr
 35 40 45

Glu Leu Gly Lys Met Ala Ala Tyr Ala Asp Glu Phe Ser Lys Arg Gly
 50 55 60

Val Lys Leu Leu Gly Leu Ser Cys Asp Asp Val Gln Ser His Lys Glu
 65 70 75 80

Trp Ile Lys Asp Ile Glu Ala Tyr Thr Pro Gly Cys His Val Lys Tyr
 85 90 95

Pro Ile Ala Ala Asp Pro Thr Arg Glu Ile Ile Gln Gln Leu Asn Met
 100 105 110

Val Asp Pro Asp Glu Thr Glu Ser Ser Lys Cys Ala Val Pro Ser Arg
 115 120 125

Ala Leu His Ile Ile Gly Pro Asp Lys Arg Ile Lys Leu Ser Phe Leu
 130 135 140

Tyr Pro Ala Ser Thr Gly Arg Asn Met Asp Glu Val Leu Arg Ala Val
 145 150 155 160

Glu Ser Leu Gln Gln Ala Ala Lys His Lys Val Ala Thr Pro Ala Asn
 165 170 175

Trp Lys Pro Gly Glu Pro Val Val Ile Lys Pro Asp Val Ser Ser Glu
 180 185 190

126

Glu Ala Lys Lys Leu Phe Pro Gln Gly Tyr Lys Ser Val Asp Leu Pro
 195 200 205

Ser Lys Lys Asp Tyr Leu Arg Phe Thr Asn Val
 210 215

<210> 71
 <211> 660
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Синтетическая последовательность ДНК по изобретению, кодирующая белок PER1 Xerophyta viscosa с использованием кодонов, оптимизированных для кукурузы, и в которой удаляются последовательности, идентифицированные в таблице 2, и поддерживаются последовательности из таблицы 1

<220>
 <221> CDS
 <222> (1)..(660)

<400> 71
 atg cct gga ttg act att ggt gac aca att ccc aac ttg gag ctg gat 48
 Met Pro Gly Leu Thr Ile Gly Asp Thr Ile Pro Asn Leu Glu Leu Asp
 1 5 10 15
 acg aca caa ggt cgc atc aag atc cac gac tat gtc ggg aat gga tac 96
 Thr Thr Gln Gly Arg Ile Lys Ile His Asp Tyr Val Gly Asn Gly Tyr
 20 25 30
 gtg att ctc ttc tca cat cct ggt gat ttc act ccg gtg tgt acc acc 144
 Val Ile Leu Phe Ser His Pro Gly Asp Phe Thr Pro Val Cys Thr Thr
 35 40 45
 gaa ttg ggc aag atg gct gct tat gcc gac gag ttc tct aag cgt ggt 192
 Glu Leu Gly Lys Met Ala Ala Tyr Ala Asp Glu Phe Ser Lys Arg Gly
 50 55 60
 gtg aag ctg ctt ggg ttg tcc tgt gat gat gtc caa tca cat aag gag 240
 Val Lys Leu Leu Gly Leu Ser Cys Asp Asp Val Gln Ser His Lys Glu
 65 70 75 80
 tgg atc aaa gac ata gag gct tac aca cct ggc tgt cac gta aaa tat 288
 Trp Ile Lys Asp Ile Glu Ala Tyr Thr Pro Gly Cys His Val Lys Tyr
 85 90 95
 ccg att gct gct gat cca acc aga gaa atc ata cag cag ctg aac atg 336
 Pro Ile Ala Ala Asp Pro Thr Arg Glu Ile Ile Gln Gln Leu Asn Met
 100 105 110
 gtg gac cct gat gag acg gaa agc tct aag tgc gct gtg cct tct agg 384
 Val Asp Pro Asp Glu Thr Glu Ser Ser Lys Cys Ala Val Pro Ser Arg
 115 120 125
 gca ctt cac atc ata gga cca gat aag agg atc aag ctg tcc ttc ctc 432
 Ala Leu His Ile Ile Gly Pro Asp Lys Arg Ile Lys Leu Ser Phe Leu
 130 135 140
 tac cct gcc tct act ggt cgc aac atg gac gaa gtt ctt aga gcc gtt 480
 Tyr Pro Ala Ser Thr Gly Arg Asn Met Asp Glu Val Leu Arg Ala Val
 145 150 155 160

127

gag tct ctt cag caa gca gct aaa cac aaa gtt gca act cct gct aac 528
 Glu Ser Leu Gln Gln Ala Ala Lys His Lys Val Ala Thr Pro Ala Asn
 165 170 175

tgg aaa cct ggc gaa cca gtc gtc atc aaa cca gac gtc agc tcc gag 576
 Trp Lys Pro Gly Glu Pro Val Val Ile Lys Pro Asp Val Ser Ser Glu
 180 185 190

gag gcc aag aag ctc ttt cct caa ggt tat aaa agc gtt gat ttg cct 624
 Glu Ala Lys Lys Leu Phe Pro Gln Gly Tyr Lys Ser Val Asp Leu Pro
 195 200 205

tca aag aag gac tac ttg agg ttc acc aat gtt tga 660
 Ser Lys Lys Asp Tyr Leu Arg Phe Thr Asn Val
 210 215

<210> 72
 <211> 219
 <212> PRT
 <213> Искусственная последовательность

<220>
 <223> Синтетическая конструкция

<400> 72

Met Pro Gly Leu Thr Ile Gly Asp Thr Ile Pro Asn Leu Glu Leu Asp
 1 5 10 15

Thr Thr Gln Gly Arg Ile Lys Ile His Asp Tyr Val Gly Asn Gly Tyr
 20 25 30

Val Ile Leu Phe Ser His Pro Gly Asp Phe Thr Pro Val Cys Thr Thr
 35 40 45

Glu Leu Gly Lys Met Ala Ala Tyr Ala Asp Glu Phe Ser Lys Arg Gly
 50 55 60

Val Lys Leu Leu Gly Leu Ser Cys Asp Asp Val Gln Ser His Lys Glu
 65 70 75 80

Trp Ile Lys Asp Ile Glu Ala Tyr Thr Pro Gly Cys His Val Lys Tyr
 85 90 95

Pro Ile Ala Ala Asp Pro Thr Arg Glu Ile Ile Gln Gln Leu Asn Met
 100 105 110

Val Asp Pro Asp Glu Thr Glu Ser Ser Lys Cys Ala Val Pro Ser Arg
 115 120 125

Ala Leu His Ile Ile Gly Pro Asp Lys Arg Ile Lys Leu Ser Phe Leu
 130 135 140

128

Tyr Pro Ala Ser Thr Gly Arg Asn Met Asp Glu Val Leu Arg Ala Val
 145 150 155 160

Glu Ser Leu Gln Gln Ala Ala Lys His Lys Val Ala Thr Pro Ala Asn
 165 170 175

Trp Lys Pro Gly Glu Pro Val Val Ile Lys Pro Asp Val Ser Ser Glu
 180 185 190

Glu Ala Lys Lys Leu Phe Pro Gln Gly Tyr Lys Ser Val Asp Leu Pro
 195 200 205

Ser Lys Lys Asp Tyr Leu Arg Phe Thr Asn Val
 210 215