

(19)



**Евразийское
патентное
ведомство**

(11) **015908**

(13) **B1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2011.12.30

(21) Номер заявки
200800448

(22) Дата подачи заявки
2006.08.30

(51) Int. Cl. **C07K 14/325** (2006.01)
C12N 15/82 (2006.01)
A01H 5/00 (2006.01)
A01H 5/10 (2006.01)
A01N 63/00 (2006.01)

(54) ИНСЕКТИЦИДНЫЙ БЕЛОК В.THURINGIENSIS CRY1A.105, КОДИРУЮЩИЙ ЕГО ПОЛИНУКЛЕОТИД И ИХ ПРИМЕНЕНИЕ

(31) 60/713,144

(32) 2005.08.31

(33) US

(43) 2008.08.29

(86) PCT/US2006/033868

(87) WO 2007/027777 2007.03.08

(71)(73) Заявитель и патентовладелец:
МОНСАНТО ТЕКНОЛОДЖИ, ЛЛС
(US)

(72) Изобретатель:
Богданова Наталья Н., Корбин
Дэвид Р., Малвар Томас М., Перлак
Фредерик Дж., Робертс Джеймс К.,
Романо Чарльз П. (US)

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

(56) WO-A-02/14517
US-A1-5689052
EP-A-1103616
US-A1-2004/093637
US-B1-6180774
WO-A-03/093484
WO-A-00/26371

(57) Изобретение относится к нуклеотидным последовательностям, кодирующим инсектицидный белок, демонстрирующий инсектицидные свойства по отношению к чешуекрылым насекомым, а также к новому инсектицидному белку, называемому инсектицид Cry1A.105, трансгенным растениям, экспрессирующим этот инсектицид, а также к способам обнаружения данного инсектицида или кодирующих его нуклеотидных последовательностей в биологической пробе.

B1

015908

015908
B1

Предшествующий уровень техники

Изобретение относится к новым кодирующим последовательностям для использования в растениях. Данные последовательности кодируют химерный белок-инсектицид, токсичный для широкого видового спектра чешуекрылых паразитов зерновых культур.

На протяжении долгого времени для биологического контроля за численностью сельскохозяйственных насекомых-вредителей использовались коммерчески доступные препараты природных штаммов *B. thuringiensis*. Споры *Bacillus thuringiensis*, а также кристаллические продукты ферментации этих бактерий, в виде концентрированных препаратов используются для обработки листьев растений в соответствии с принятой сельскохозяйственной практикой.

Кристаллические белки семейства *CryI* известны своей биоактивностью против личинок чешуекрылых насекомых и применяются в качестве агентов для контроля за численностью чешуекрылых насекомых-вредителей. Прекурсорная форма *CryI* δ -эндотоксина состоит из двух приблизительно равных по размеру сегментов. С-концевая часть прекурсорной формы, или про-токсичный сегмент, стабилизирует кристаллическую структуру и не имеет инсектицидной активности. N-концевая часть прекурсора содержит токсичный сегмент белка *CryI*, который на основании множественного выравнивания консервативных или практически консервативных последовательностей данных сегментов у различных членов семейства *CryI*, может быть разделен на три структурных домена. Основой для разделения этих трех субдоменов служит трехмерная кристаллографическая структурная модель *CryIA* δ -эндотоксина, три субдомена которого обозначаются как домен I, домен II и домен III (нумерация идет с N-конца токсичного сегмента белка *CryI*). Первую треть активного токсичного сегмента занимает домен I, и его наличие является необходимым для образования каналов (Thompson с соавт., 1995). Домены II и III занимают центральную и С-концевую части активного токсичного сегмента соответственно. Оба эти домена участвуют в связывании с рецепторами и в специфическом видовом распознавании насекомых в зависимости от конкретных изучаемых видов насекомых и δ -эндотоксинов (Thompson с соавт., 1995).

В том случае, если химерный белок планируется создавать при помощи произвольной комбинации доменных структур различных кристаллических инсектицидных белков, известных в данной области, вероятность того, что полученный химерный белок будет обладать улучшенными (по сравнению с парентальными белками) свойствами, низка. Причиной этого является сложность различных аспектов природы белков, таких как структура, фолдинг, олигомеризация и активация, в том числе правильный протеолитический процессинг химерного прекурсора (в том случае, если белок экспрессируется в подобной форме), в результате которого должен выделяться инсектицидный токсин. Таким образом, функциональные химерные инсектицидные токсины, демонстрирующие повышенную по сравнению с каждым из парентальных белков инсектицидную активность, могут быть получены только путем тщательного отбора специфических целевых регионов каждого из парентальных белков, которые в дальнейшем предстоит включить в химерную конструкцию. На практике, "перетасовка" токсичных доменов, т.е. сборка химерного токсина, в котором домены I, II и III различных токсинов, приводит к получению некристаллизуемого белка и/или к полной потере инсектицидной активности по отношению к предпочтительным видам насекомых-вредителей. В некоторых случаях химерный токсин демонстрирует хорошие способности к кристаллообразованию, однако лишен всякой инсектицидной активности. Таким образом, эффективные инсектицидные химерные белки получают только методом проб и ошибок, и даже в этом случае вполне квалифицированный специалист не может быть уверен, что в результате его исследований он получит химерный белок, демонстрирующий инсектицидную активность, эквивалентную или превышающую таковую для каждого из парентальных токсинов, фрагменты которых были взяты для получения химерного белка.

В литературе описана конструкция или сборка химерных белков из двух или более кристаллических инсектицидных белков-прекурсоров *B. thuringiensis*, при этом не все из полученных химер обладают способностью к кристаллообразованию, эквивалентными или превышающими таковые для парентальных прекурсорных белков, из которых был получен химерный конструктор. (Bosch с соавт. (WO 95/06730); Thompson с соавт. (WO 95/30753); Thompson с соавт. (WO 95/30752); Malvar с соавт. (WO 98/22595); Gilroy с соавт. (патент США № 5128130); Gilroy (патент США № 5055294); Lee с соавт. (1992) Gene 267:3115-3121; Honee с соавт. (1991) Mol. Microbiol. 5:2799-2806; Schnepf с соавт. (1990) J. Biol. Chem. 265:20923-20930; Perlak с соавт. (1990) Bio/Technol. 8:939-9943; Perlak с соавт. (1993) Plant Mol. Biol. 22:313-321).

Было показано, что экспрессия δ -эндотоксинов *B. thuringiensis* в трансгенной кукурузе представляет собой эффективный метод контроля за численностью сельскохозяйственно значимых насекомых-вредителей (Perlak с соавт. 1990; 1993). Использование трансгенных культур, экспрессирующих δ -эндотоксины *B. thuringiensis*, позволяют значительно снизить временные и денежные затраты на борьбу с насекомыми-вредителями, связанные с топическим использованием химических инсектицидов. Использование трансгенов, кодирующих δ -эндотоксины *B. thuringiensis*, является в особенности предпочтительным. Сельскохозяйственные культуры, экспрессирующие δ -эндотоксины *B. thuringiensis*, засеянные в районах с повышенной концентрацией насекомых-вредителей, демонстрируют лучшие урожаи, нежели

аналогичные нетрансгенные коммерческие сорта. В то же время предполагается, что насекомые могут вырабатывать резистентность к экспрессируемым в трансгенных растениях δ -эндотоксинам *B. thuringiensis*. Подобная резистентность, при условии широкого распространения, может значительно снизить коммерческую ценность содержащихся в зародышевой плазме генов, таких как гены, кодирующие δ -эндотоксины *B. thuringiensis*. Одним из способов повышения эффективности трансгенных инсектицидов против целевых насекомых-вредителей, и одновременно - борьбы с появлением резистентных к инсектициду насекомых, является получение трансгенных растений, экспрессирующих большие количества δ -эндотоксинов *B. thuringiensis* (McGaughey и Whalon 1993; Roush 1994). Помимо этого, создание единой базы данных, в которую занесены гены, кодирующие инсектицидные белковые продукты, эффективные против различных насекомых-вредителей и осуществляющие свои инсектицидные эффекты различными путями, может гарантированно предотвратить появление любой резистентности. Другим способом контроля за развитием резистентности является экспрессия в одном растении двух или более различных инсектицидных композиций, токсичных для одного и того же вида насекомых, причем каждая композиция экспрессируется на уровне, достаточно высоком для успешной задержки развития резистентности. Примеры инсектицидов, которые могут быть использованы в подобных комбинациях, включают в себя без ограничения токсина *B. thuringiensis*, инсектицидные белки представителей родов *Xenorhabdus* и *Photorhabdus*, деаллергенизированные и дегликозилированные белки пататины и/или пермутеины, растительные лектины и т.д. Однако на сегодняшний момент коэкспрессия нескольких инсектицидноактивных белков в одном растении и/или достижение высокого уровня экспрессии этих инсектицидных белков, не сопровождающиеся нежелательными морфологическими изменениями трансгенного растения, остаются труднодостижимыми.

Среди более двухсот пятидесяти идентифицированных инсектицидных белков *Bacillus thuringiensis*, всего лишь для нескольких были предприняты попытки экспрессии в растениях. Так, в растениях были успешно экспрессированы некоторые белки семейств Cry1, Cry3, Cry2Aa, Cry2Ab, бинарные токсины Cry33/34 и Cry23/37, а также Cry9. Белки семейства Cry1 представляют собой наибольший класс белков, для которого были предприняты попытки экспрессии в растениях, однако ни для одного из них не удалось добиться экспрессии на высоком уровне. Для того чтобы избежать нежелательных фитотоксических эффектов, необходимо направить Cry2Ab в хлоропласты. Большая часть подобных рекомбинантных растений экспрессирует белки Cry1A. Вероятность развития резистентности к белкам Cry1A может быть в значительной степени снижена в том случае, если вместе с cry1-аллелем будет экспрессироваться аллель контроля резистентности, либо если экспрессия аллеля cry1 будет на высоком уровне. Таким образом, актуальной задачей является разработка альтернативных токсинных генов для экспрессии в растениях, предназначенных для дополнения или замещения токсинов, ранее использовавшихся в первом и втором поколениях резистентных к насекомым трансгенных растений.

Краткое описание сущности изобретения

Изобретение относится к выделенным нуклеотидным последовательностям, предназначенным для экспрессии в растениях и кодирующим белки, демонстрирующие инсектицидные свойства по отношению к чешуекрылым насекомым. SEQ ID NO:1 представляет собой пример подобной нуклеотидной последовательности; она состоит из гена cry1A.105 и кодирует инсектицидный белок Cry1A.105. Последовательность SEQ ID NO:1 подобна последовательности SEQ ID NO:3; обе из них кодируют белок Cry1A.105. Последовательность SEQ ID NO:1 предпочтительна для использования в клетках двудольных растений, в то время как SEQ ID NO:3 предпочтительна для использования в клетках однодольных растений. Аминокислотная последовательность SEQ ID NO:4 кодируется нуклеотидной последовательностью SEQ ID NO:3 и идентична аминокислотной последовательности SEQ ID NO:2. Предполагается, что выделенные нуклеотидные последовательности включают в себя последовательности, по меньшей мере приблизительно на 88-90% идентичные последовательности SEQ ID NO:1, либо гибридизирующиеся с SEQ ID NO:1 в жестких условиях. Также предполагается, что выделенные нуклеотидные последовательности включают в себя последовательности, по меньшей мере приблизительно на 90% идентичные последовательности SEQ ID NO:3, либо гибридизирующиеся с SEQ ID NO:3 в жестких условиях.

Изобретение также относится к выделенным и очищенным инсектицидным белкам, обладающим инсектицидными свойствами по отношению к чешуекрылым насекомым. Используемый в настоящей заявке термин "инсектицидный белок" относится к белку, содержащему, по меньшей мере, токсинный фрагмент Cry1A.105, аминокислотная последовательность которого соответствует последовательности SEQ ID NO:2. Также, термин "инсектицидный белок Cry1A.105" относится к полноразмерному белку-прекурзору, состоящему из приблизительно 1177 аминокислот согласно последовательности SEQ ID NO:2; в то же время, термин "инсектицидный белок Cry1A.105" также относится к любому фрагменту белка-прекурсора, демонстрирующему инсектицидную биоактивность; в том числе данный термин относится к инсектицидному белку Cry1A.105, аминокислотная последовательность которого соответствует сегменту, состоящему из от приблизительно аминокислоты 1 до приблизительно аминокислоты 612 согласно последовательности SEQ ID NO:2, а также к сегменту, состоящему из от приблизительно аминокислоты 2 до приблизительно аминокислоты 610 согласно последовательности SEQ ID NO:2. Любая ком-

позиция, содержащая инсектицидно-эффективные количества инсектицидного белка, включена в объем настоящего изобретения.

Настоящее изобретение также относится к экспрессионной кассете для экспрессии инсектицидного белка, соответствующего аминокислотной последовательности SEQ ID NO:2, в клетке-хозяине. Экспрессионная кассета предпочтительно содержит функциональный (в выбранной для экспрессии клетке-хозяине) промотор, напрямую связанный с нуклеотидной последовательностью, кодирующей инсектицидный сегмент белка Cry1A.105, и регулирующий экспрессию этого сегмента. Примерами экспрессионных кассет могут служить нуклеотидные последовательности SEQ ID NO:5 и SEQ ID NO:7, предназначенные для использования в клетках двудольных и однодольных растений соответственно. Промотор и кодирующая последовательность оперативно связаны и функционируют вместе в клетке-хозяине. Экспрессионная кассета предназначена для использования в любых клетках-хозяевах, но предпочтительно - в клетках бактерий, грибов, млекопитающих или растений. Бактериальные клетки предпочтительно выбирают из группы, включающей в себя бактерии родов *Bacillus*, *Enterobacteriaceae*, *Pseudomonas*, *Clostridium*, *Rhizobium* и *Agrobacterium*. В том случае, если клетки-хозяева представляют собой растительные клетки, их предпочтительно выбирают из группы, включающей в себя клетки сельскохозяйственных культур, предпочтительно двудольных или однодольных растений. Примерами клеток двудольных растений могут служить клетки люцерны, яблока, абрикоса, спаржи, фасоли, кофе, ежевики, черники, канолы, моркови, цветной капусты, сельдерея, вишни, нута, цитрусовых, хлопчатника, вигны китайской, клюквы, огурца, тыквы, баклажана, плодовых деревьев, винограда, лимона, салата-латука, льна, дыни, горчицы, орехоплодных деревьев, бамии, окры, апельсина, гороха, персика, арахиса, груши, сливы, картофеля, сои, крупноплодной тыквы, земляники, сахарной свеклы, подсолнуха, батата, табака, томата, репы и других овощей. Примерами клеток однодольных растений могут служить клетки кукурузы, пшеницы, овса, риса, сорго, гречихи, ржи, злаков (овсяница, тимофеевка луговая, костер, ежа сборная, августина трава, свиной пальчатый, полевица) и ячменя. Экспрессионные кассеты, предназначенные для использования в растительных клетках, содержат оперативно связанные последовательности, регулирующие уровень и эффективность экспрессии целевого белка, такого как инсектицидный белок Cry1A.105. Подобные последовательности могут представлять собой экспрессионный энхансер, нетранслируемую лидерную последовательность, интрон, последовательность, кодирующую пептид, направляющий экспрессируемый белок в хлоропласты, а также сигналы терминации транскрипции и полиаденилирования.

Экспрессионная кассета предпочтительно включается в состав вектора для стабилизации персистенции последовательности, кодирующей Cry1A.105, в клетке-хозяине. Вектор может представлять собой любую из известных в данной области структур, но в типичном случае это плазида или репликон, в котором конструируется, или в который вставляется экспрессионная кассета перед трансфекцией вектора в клетку-хозяина. Вектор может представлять собой без ограничения плазмиду, космиду, ВАС-миду, фагмиду, YAC, BAC, «суицидный» вектор, инсерционную последовательность, транспозон, или даже линейную нуклеотидную последовательность, с которой связана или в которую заключена экспрессионная кассета.

В одном из воплощений настоящее изобретение относится к трансгенным растениям, устойчивым к заражению чешуекрылыми насекомыми. Подобные растения содержат нуклеотидную последовательность, кодирующую инсектицидный белок Cry1A.105, аминокислотная последовательность которого соответствует по меньшей мере фрагменту последовательности SEQ ID NO:2 (от приблизительно аминокислоты 2 до приблизительно аминокислоты 612). Использование подобных трансгенных растений является эффективным способом борьбы с заражением чешуекрылыми насекомыми, такими как листовертки, гусеницы озимой совки, «походные черви», точильщики, мешочница поденкоподобная и кормовые гусеницы. Предпочтительными паразитами являются совка травяная, мотылек стеблевой кукурузный, гусеница совки хлопковой американской, огневка кукурузная юго-западная, а также совка-ипсилон.

Объем настоящего изобретения также включает в себя потомство, а также семена, фруктовые плоды и продукты, получаемые из трансгенных растений согласно настоящему изобретению, в том случае, если кодирующая инсектицидный белок Cry1A.105 нуклеотидная последовательность согласно настоящему изобретению содержится в составе геномной или плазмидной ДНК растения, его потомства, семян и т.д.

Настоящее изобретение также относится к способам борьбы с заражением растений чешуекрылыми насекомыми-вредителями, подразумевающим скормливание насекомым композиции, содержащей инсектицидно-эффективное количество инсектицидного белка Cry1A.105. Одним из вариантов подобной композиции могут быть растительные клетки, полученные из (или являющиеся потомством) растительной клетки, трансформированной нуклеотидной последовательностью, кодирующей инсектицидный сегмент аминокислотной последовательности Cry1A.105 согласно SEQ ID NO:2. Одним из источников подобной инсектицидной композиции может быть трансгенное растение, полученное из трансформированной растительной клетки, содержащей экспрессионную кассету, подобную последовательностям SEQ ID NO:5 и SEQ DD NO:7. Другим способом доставки подобной инсектицидной композиции может быть экспрессия инсектицидно-эффективного количества Cry1A.105 в клетках бактерий или грибов с последующим

скармливанием этих клеток, их гомогенатов или очищенного белка Cry1A.105, целевым насекомым-вредителям, восприимчивым к инсектицидному действию Cry1A.105.

Настоящее изобретение также относится к способу идентификации нуклеотидной последовательности, кодирующей аминокислотную последовательность Cry1A.105, в биологической пробе. Данный способ подразумевает приведение тестируемой на наличие кодирующей последовательности Cry1A.105 пробы в контакт с полинуклеотидным зондом, специфично связывающимся с кодирующей последовательностью Cry1A.105. В частности, последовательность зонд связывается, или гибридизуется, с кодирующей последовательностью Cry1A.105 в жестких гибридизационных условиях. Детекция связывания в реакционной смеси является диагностичной в отношении присутствия кодирующей последовательности Cry1A.105.

Настоящее изобретение также относится к способу идентификации инсектицидного белкового фрагмента Cry1A.105. Данный способ подразумевает приведение тестируемой на наличие инсектицидного белкового фрагмента Cry1A.105 пробы в контакт с антителом, специфично связывающимся с инсектицидным белковым фрагментом Cry1A.105. Детекция связывания в реакционной смеси является диагностичной в отношении присутствия инсектицидного белкового фрагмента Cry1A.105.

Настоящее изобретение также относится к химерным или гибридным инсектицидным белкам. Подобные гибриды состоят из двух или более различных инсектицидных белков, каждый из которых обладает инсектицидной активностью по отношению по меньшей мере к одному виду насекомых, принадлежащих к одной и той же группе. Гибридные инсектицидные белки конструируются из частей различных инсектицидных белков. Сегменты инсектицидных белков, используемые для конструирования гибридов, состоят по меньшей мере из приблизительно от 50 до по меньшей мере приблизительно 200 аминокислот, составляющих непрерывную последовательность в составе непрерывной аминокислотной последовательности любого из инсектицидных белков. Инсектицидный белок Cry1A.105, аминокислотная последовательность которого соответствует сегменту с приблизительно 2-й по приблизительно 612-ю аминокислоту в последовательности SEQ ID NO:2 входит в состав группы различных инсектицидных белков, каждый из которых может быть выбран в качестве сегмента для конструирования гибридного инсектицидного белка.

Различные преимущества, особенности и сущность настоящего изобретения станут более очевидны из нижеследующего подробного описания изобретения, примеров и формулы изобретения.

Краткое описание последовательностей

SEQ ID NO:1 представляет собой синтетическую нуклеотидную последовательность для экспрессии инсектицидного белка Cry1A.105, предпочтительно в двудольных растениях.

SEQ ID NO:2 представляет собой аминокислотную последовательность белка Cry1A.105, кодируемую нуклеотидной последовательностью SEQ ID NO:1.

SEQ ID NO:3 представляет собой синтетическую нуклеотидную последовательность для экспрессии инсектицидного белка Cry1A.105, предпочтительно в однодольных растениях.

SEQ ID NO:4 представляет собой аминокислотную последовательность белка Cry1A.105, кодируемую нуклеотидной последовательностью SEQ ID NO:3.

SEQ ID NO:5 представляет собой нуклеотидную последовательность, состоящую из экспрессионной кассеты, функционирующей в растительных клетках, предпочтительно в клетках двудольных растений, предназначенную для экспрессии инсектицидного белка Cry1A.105.

SEQ ID NO:7 представляет собой нуклеотидную последовательность, состоящую из экспрессионной кассеты, функционирующей в растительных клетках, предпочтительно в клетках однодольных растений, предназначенную для экспрессии инсектицидного белка Cry1A.105.

SEQ ID NO:8 представляет собой аминокислотную последовательность инсектицидного белка Cry1A.105, кодируемую кодирующим сегментом экспрессионной кассеты SEQ ID NO:7.

Подробное описание изобретения

Согласно настоящему изобретению авторы сконструировали нуклеотидную последовательность, кодирующую новый инсектицидный белок Cry1A.105. Было показано, что аминокислотная последовательность белка Cry1A.105 согласно SEQ ID NO:2 обладает улучшенными свойствами по сравнению с естественными инсектицидными белками *B. thuringiensis*, токсичными по отношению к чешуекрылым насекомым. В частности, белок Cry1A.105 может быть экспрессирован на высоком уровне как в однодольных, так и в двудольных растениях без характерных для трансгеноза фитотоксических эффектов, возникающих как результат повышенного по сравнению с естественным уровнем экспрессии белков Cry1. Помимо этого, при экспрессии в *Bacillus thuringiensis*, белок Cry1A.105 формирует стабильные кристаллы, по-видимому, из-за стабилизирующего эффекта протоксинового сегмента Cry1Ac, связанного с токсинным фрагментом в составе химерного белка Cry1A.105. Помимо этого, инсектицидный белок Cry1A.105 обладает диапазоном различных биологических активностей по отношению к чешуекрылым насекомым, которые не были найдены у других естественных белков Cry1, идентифицированных к настоящему времени. Таким образом, экспрессия белка Cry1A.105 в трансгенных растениях приводит к увеличению числа морфологически нормальных трансгенных растений, экспрессирующих повышенные уровни аналога токсина Cry1, демонстрирующего широкий диапазон средств контроля за численностью

чешуекрылых насекомых-вредителей в любом трансгенном растении, выбранном для дальнейшего коммерческого внедрения. Использование подобных трансгенных растений может привести к задержке развития резистентности к аналогу токсина Cry1A, а в случае комбинированного использования со вторым токсином, который является токсичным по отношению к одному или более видам насекомых-вредителей, по отношению к которым аналог токсина Cry1A также является токсичным, и в том случае, если эти два токсина осуществляют свое инсектицидное действие различными путями, вероятность развития резистентности к любому из двух токсинов крайне мала.

Авторы настоящего изобретения сконструировали по меньшей мере две различных нуклеотидных последовательности для использования в растениях, при этом обе последовательности кодируют один и тот же белок Cry1A.105. Первые (или аминоконцевые) приблизительно две трети инсектицидного фрагмента белка Cry1A.105 состоят из аминокислотной последовательности Cry1Ab. Эта последовательность связана с С-концом токсинной части и фрагментом протоксинного домена инсектицидного белка Cry1, полученного из штамма aizawai EG6346 *B. thuringiensis* (Ecogen) (Chambers с соавт., 1991, J. Bacteriol. 173:3966-3976). Токсинный сегмент Cry1A.105 далее связан с сегментом, значительную часть которого занимает пептидная последовательность протоксина Cry1Ac. Авторы настоящего изобретения продемонстрировали, что данная конструкция кодирует уникальную аминокислотную последовательность, которая демонстрирует неожиданно улучшенные инсектицидные свойства по сравнению с таковыми для парентальных белков, из которых была получена химера. Помимо этого, прекурсорный белок Cry1A.105 обладает прекрасными кристаллообразующими свойствами и эффективно растворяется и процессируется в активную токсинную форму в кишечнике целевых чешуекрылых насекомых-вредителей.

Нуклеотидные последовательности, составляющие различные воплощения настоящего изобретения, были сконструированы при помощи способов, описанных в патентах США № 5500365 и 5689052, в частности - при их конструкции избегали использования некоторых подпоследовательностей в кодирующем регионе, которые интерферируют с экспрессией гетерологичных генов последовательностей в клетках растений. Сегмент, кодирующий токсинный фрагмент белка Cry1A.105, состоит из нуклеотидов в положении от приблизительно 1 до приблизительно 1830 (либо может иметь несколько большую или меньшую длину) последовательностей SEQ ID NO:1 и SEQ ID NO:3. Последовательность SEQ ID NO:1 была сконструирована для использования в двудольных растениях, в частности - в хлопчатнике. Последовательность SEQ ID NO:3 была сконструирована для использования в однодольных растениях, в частности - в кукурузе.

Нуклеотидные последовательности, составляющие предмет настоящего изобретения, гомологичны друг по отношению к другу на 94,3% и идентичны друг другу в области положений нуклеотидов от приблизительно 1330 до приблизительно 3534. Сегменты в области положений нуклеотидов от приблизительно 1 до приблизительно 1830 идентичны друг другу приблизительно на 88,9%. Сегменты, кодирующие первые два домена белка Cry1A.105, характеризуются значительно меньшей взаимной гомологией идентичность составляет около 84,7% друг по отношению к другу.

Авторы настоящего изобретения конструировали трансгенные растения при помощи вышеописанных последовательностей.

SEQ ID NO:1 вводили в плазмидный вектор, содержащий экспрессионную кассету, состоящую из усиленного промотора вируса норичниковой мозаики (eFMV), оперативно связанного с нетранслируемой лидерной последовательностью гена Hsp70 *Petunia hybrida* (Ph.Hsp70, DnaK), кодирующей последовательностью хлоропласт-таргеттирующего пептида малой субъединицы рибулозо-бис-фосфат-карбоксилазы *Arabidopsis thaliana*, а также сигналами терминации транскрипции и полиаденилирования гена рибулозо-бис-фосфат-карбоксилазы E9 *Pisum sativum*. Кодирующую последовательность Cry1A.105, согласно SEQ ID NO:1, вставляли в экспрессионную кассету в открытой рамке считывания непосредственно вплотную к 3'-концу кодирующей последовательности хлоропласт-таргеттирующего пептида, с 5'-конца по отношению к сигналу терминации трансляции E9. Нуклеотидная последовательность полученной экспрессионной кассеты приведена в SEQ ID NO:5. Сегмент вектора, содержащий экспрессионную кассету Cry1A.105 и связанный со второй экспрессионной кассетой, содержащей растительно-экспрессируемый GUS-маркер, вырезали и использовали для трансгеноза, который осуществляли при помощи биолистических методов. Полученные продукты трансгеноза тестировали на наличие инсектицидной активности по отношению к нескольким видам чешуекрылых насекомых-вредителей; было показано, что продукты трансгеноза обладают значительно лучшими характеристиками контроля за численностью насекомых по сравнению с ранее полученным резистентным к заражению насекомыми хлопчатником, содержащим только Cry1Ac или комбинацию белков Cry1Ac и Cry2Ab. Помимо этого, некоторые продукты трансгеноза демонстрировали накопление белка Cry1A.105, составляющее более 10-20 частей на миллион за один вегетационный период; подобное накопление также имело место в семенных коробочках хлопчатника, при этом не было обнаружено никаких фитотоксических эффектов на растение или репродуктивные ткани. Полученные продукты трансгеноза значительно отличались от протестированных ранее продуктов, содержащих другие белки Cry1, аккумуляция которых составляла менее приблизительно 10 частей на миллион, вне зависимости от того, таргеттировался ли продукт экспрессии в хлоропласты или нет. Помимо этого, при тестировании других белков Cry1 в хлопчатнике, в особенности если

уровень накопления трансгенного белка составлял более приблизительно 10 частей на миллион, были зарегистрированы различные фитотоксические эффекты.

SEQ ID NO:3 вводили в плазмидный вектор, содержащий экспрессионную кассету, состоящую из усиленного промотора вируса мозаики цветной капусты (eCaMV), оперативно связанного с нетранслируемой лидерной последовательностью гена хлорофилл-*a/b*-связывающего белка *Triticum aestivum*, интронной последовательностью гена актина *Oryza sativa*, а также сигналами терминации трансляции и полиаденилирования гена *hsp17*. Кодировующую последовательность *Cry1A.105* согласно SEQ ID NO:3 вставляли в экспрессионную кассету непосредственно вплотную к 3'-концу интрона, с 5'-конца по отношению к сигналу терминации трансляции. Нуклеотидная последовательность полученной экспрессионной кассеты приведена в SEQ ID NO:7. Вектор также содержит ген устойчивости к гербициду глифосату, который служит в качестве селекционного маркера для отбора тех клеток или растений, в которых произошло трансгенное событие. Продукты трансгеноза кукурузы, отобранные после трансформации экспрессионной кассетой *Cry1A.105*, тестировали на наличие инсектицидной активности по отношению к нескольким видам чешуекрылых насекомых-вредителей; было показано, что продукты трансгеноза обладают широким спектром инсектицидных свойств, нехарактерных для трансформантов, полученных с использованием других инсектицидных белков *B. thuringiensis*, таких как. Демонстрируемая экспрессирующими инсектицидными количества белка *Cry1A.105* продуктами трансгеноза кукурузы инсектицидная активность по отношению к совке травяной и совке-ипсилон суммарно с инсектицидной активностью *Cry1A.105* по отношению к совке хлопковой и точильщику зерновому расширяет спектр инсектицидной активности *Cry1Ab*, демонстрируемой продуктами трансгеноза *Cry1A.105*.

Нуклеотидные последовательности согласно настоящему изобретению являются лишь примерами. Другие нуклеотидные последовательности также могут быть использованы для экспрессии инсектицидного белка *Cry1A.105* в растительных клетках, либо их дизайн может быть проведен таким образом, чтобы обеспечить экспрессию *Cry1A.105* в других типах клеток-хозяев. Без ограничения объема раскрытого изобретения, нуклеотидная последовательность для экспрессии инсектицидного фрагмента *Cry1A.105* по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95% или по меньшей мере приблизительно на 99% идентична иллюстративным последовательностям согласно настоящему изобретению. Другие нуклеотидные последовательности, предназначенные для экспрессии инсектицидного фрагмента *Cry1A.105* в нерастительных клетках-хозяевах, могут демонстрировать любой процент гомологии или идентичности с модельными последовательностями согласно настоящему изобретению. Нуклеотидные последовательности могут варьироваться в силу вырожденности генетического кода; таким образом, можно синтезировать любое число нуклеотидных последовательностей, кодирующих любой фрагмент аминокислотной последовательности SEQ ID NO:2, и все подобные последовательности включены в объем настоящего изобретения. В объем настоящего изобретения также включены любые выделенные и очищенные нуклеотидные последовательности, кодирующие по меньшей мере инсектицидный фрагмент белка *Cry1.105*, а также любые композиции, при помощи антителиа, ДНК/РНК-зонда или двух или более пар праймеров, дизайн которых проведен таким образом, что при проведении ПЦР с этими праймерами синтезируется ампликон, содержащий одну из обсуждаемых нуклеотидных последовательностей.

Кодирующий регион экспрессируемой в кукурузе модельной последовательности содержит только последовательность, кодирующую белок-прекурсор *Cry1A.105*, в то время как модельная последовательность, экспрессируемая в хлопчатнике, содержит также кодирующую последовательность хлоропласт-таргеттирующего пептида в виде химерной конструкции с *Cry1A.105*. Было показано, что экспрессия белков *Cry1* в растениях весьма проблематична. Априори неизвестно, будет ли тот или иной белок *Cry1* хорошо экспрессироваться в растениях, поэтому неизбежно использование метода "проб и ошибок". Некоторые белки *Cry1*, будучи экспрессированными в кукурузе, приводят к фитотоксическим эффектам; подобные эффекты иногда могут быть ослаблены путем направления синтезированного белка в хлоропласты. Вышесказанное также справедливо и для хлопчатника, экспрессирующего белки *Cry1*. Приведенные в настоящей заявке примеры не имеют целью показать, что экспрессия *Cry1A.105* в кукурузе возможна только в том случае, если белковый продукт трансгена локализуется в цитоплазматическом пространстве, и аналогично - что экспрессия *Cry1A.105* в хлопчатнике возможна только в том случае, если белковый продукт трансгена локализуется в хлоропластах. Приведенные примеры имеют целью продемонстрировать, что любая локализация является функциональной для белка, составляющего предмет настоящего изобретения, причем полученные морфологически нормальные трансгенные растения демонстрируют высокие уровни экспрессии и накопления белка *Cry1A.105*, а также коммерчески приемлемые уровни резистентности к широкому спектру чешуекрылых насекомых-вредителей, выбранных из группы, состоящей из родов *Anticarsia*, *Pseudoplusia*, *Rachiplusia*, *Helicoverpa*, *Heliothis*, *Spodoptera*, *Epinotia* и *Armigera*. Авторы изобретения считают, что последовательность, кодирующая любой из хлоропласт-таргеттирующих пептидов, будет эффективна для направления белка-прекурсора *Cry1A.105* в хлоропласты/пластиды.

Нетранслируемые лидерные последовательности, интроны и 3'-сигналы терминации трансляции и полиаденилирования хорошо известны в данной области, и соответствующему специалисту будет понят-

но, что в некоторых случаях экспрессия может быть улучшена или стабилизирована путем включения подобных последовательностей в экспрессионную кассету. В генной инженерии известно большое число подобных последовательностей, и использование их всех в контексте экспрессии белка Cry1A.105 включено в объем настоящего изобретения. Аналогично, в области генной инженерии хорошо известны различные промоторы, регулирующие экспрессию контролируемой и связанной с ними последовательности; использование их всех в контексте экспрессии белка Cry1A.105 также включено в объем настоящего изобретения. Промоторы могут быть выбраны таким образом, чтобы регулировать экспрессию контролируемой последовательности в соответствии с любым числом комбинаций различных параметров, к примеру, чтобы обеспечивать временной, пространственный или тканеспецифичный контроль экспрессии, а также контроль накопления белкового продукта трансгена в той или иной растительной ткани или клетке.

Выделенный и очищенный белок, содержащий инсектицидный фрагмент аминокислотной последовательности Cry1A.105, также включен в объем настоящего изобретения, равно как и различные варианты подобного белка, в том случае, если аминокислотная(ые) замена(ы) в подобных вариантах являются консервативными и не приводят к снижению инсектицидной биоактивности или сужению спектра видовой специфичности последней. Подразумевается, что инсектицидный фрагмент белка Cry1A.105 представляет собой сегмент, включающий в себя аминокислоты в положении от приблизительно 1 до приблизительно 650, от приблизительно 2 до приблизительно 612, от приблизительно 5 до приблизительно 610 или от приблизительно 10 до приблизительно 600 аминокислотной последовательности SEQ ID NO:2. В качестве альтернативы, инсектицидный фрагмент белка Cry1A.105 может представлять собой аминокислотную последовательность длиной от приблизительно 550 до приблизительно 650 аминокислот, составляющих непрерывную аминокислотную последовательность в составе последовательности SEQ ID NO:2. Полноразмерный белок-прекурсор (аминокислоты от положения 1 до приблизительно 3534) обладает прекрасными кристаллообразующими характеристиками и хорошо переносится как односторонними, так и двусторонними видами. Белок-прекурсор также демонстрирует превосходную стабильность в кристаллической форме, а также прекрасную растворимость при щелочных значениях pH, в частности при pH от приблизительно 8,0 до приблизительно 12,0, от приблизительно 8,5 до приблизительно 11,5 либо от приблизительно 9,0 до приблизительно 11,0.

Белок, составляющий предмет настоящего изобретения, может быть очищен и использован в инсектицидно-эффективном количестве в составе любого числа композиций, предназначенных для контроля численности чешуекрылых насекомых-вредителей; данный белок может быть также скомбинирован в инсектицидно-эффективном количестве с любым числом других пестицидных агентов, отличных от белка Cry1A.105. Подобные пестицидные агенты включают в себя без ограничения другие белки *Cry V. thuringiensis*, другие инсектицидные композиции, в том числе химические инсектициды (вне зависимости от их токсичности по отношению к чешуекрылым насекомым), фунгицидные и фунгистатические агенты, антибиотики, антибактериальные агенты, бактериостатические агенты, а также нематоцидные и нематостатические агенты. Подобные пестицидные комбинации, содержащие в своем составе белок Cry1A.105 и любое число других пестицидных агентов, могут продуцироваться трансгенной клеткой, либо же они могут быть получены смешиванием очищенных или в значительной степени очищенных индивидуальных пестицидных агентов с получением единой пестицидной композиции в виде порошка, гранулированного материала масляной суспензии, водной суспензии, водно-масляной эмульсии или смачиваемого порошка, которая далее в виде смеси с сельскохозяйственно приемлемым носителем может быть использована для обработки листьев.

Композиция также может быть получена в виде, пригодном для обработки семян - либо в виде композиции, содержащей Cry1A.105 и предназначенной для обработки семян, либо в виде композиции, предназначенной для обработки семян, полученных из трансгенного растения, экспрессирующего инсектицидно-эффективные количества Cry1A.105 - в этом случае целевые чешуекрылые насекомые-вредители получают как композицию, содержащую пестицидные агенты, так и клетки растений, выросших из семян, продуцирующих пестицидно-эффективные количества белка Cry1A.105. Комбинация инсектицидных белков, каждый из которых токсичен по отношению к одним и тем же видам насекомых, и которые в то же время реализуют свой токсический потенциал по разным механизмам, может быть в особенности полезной для контроля за численностью чешуекрылых насекомых-вредителей, либо для задержки развития резистентности к каждому из входящих в комбинацию пестицидных агентов, которые в отсутствие резистентности эффективны против данных конкретных видов чешуекрылых насекомых-вредителей. Модельной комбинацией подобных белков может служить комбинация белка Cry1A.105 (т.е. первого инсектицидного белка), составляющего предмет настоящего изобретения, по меньшей мере с одним другим инсектицидным белком, отличным от первого. Подобные инсектицидные белки включают в себя без ограничения другие кристаллические белки *V. thuringiensis* (другие белки Cry1, белки Cry2, белки Cry5 и белки Cry9), белки VIP, токсичные по отношению к чешуекрылым насекомым белки T1C, а также инсектицидные белки, продуцируемые бактериями родов *Xenorhabdus* и *Photorhabdus*. Комбинация одного или более инсектицидных белков с агентом, обеспечивающим dsРНК-опосредованную экспрессию одного или более генов, критичных для выживания насекомого, является в особенности полезной

для контроля за численностью чешуекрылых насекомых-вредителей, либо для задержки развития резистентности к каждому из входящих в комбинацию пестицидных агентов, которые в отсутствие резистентности эффективны против данных конкретных видов чешуекрылых насекомых-вредителей.

В другом воплощении настоящее изобретение относится к растениям, трансформированным нуклеотидными последовательностями, составляющими предмет настоящего изобретения. Способы стабильной трансформации растительных клеток хорошо известны в области генной инженерии растений и включают в себя без ограничения вакуумную инфльтрацию, *Agrobacterium*- или *Rhizobium*-опосредованную трансформацию, электропорацию и различные баллистические методы. ДНК, вводимая в растения, обычно таргеттируется для интеграции в ядерную хромосому, хотя интеграция в ДНК плазмидов/хлоропластов также может быть осуществлена. ДНК, вводимая в растения, обычно связана с последовательностью, кодирующей маркерные белки, позволяющие идентифицировать или отобрать клетку(и), в которой(ых) произошла стабильная трансформация целевой ДНК. Подобные маркерные гены включают в себя без ограничения гены, кодирующие флуоресцентные или светорассеивающие белки, пигменты или ферменты, наделяющие экспрессирующие их клетки колориметрическими свойствами. В качестве альтернативы, подобные маркерные гены могут представлять собой селекционные маркеры, при помощи которых можно провести позитивную селекцию трансформированных клеток или тканей и которые предоставляют трансформированным клеткам преимущество в росте, причем таким образом, что нетрансформированные клетки при этом становятся статичными или погибают. Примерами подобных селекционных маркеров могут без ограничения служить гены *basta* и *bar*, ген устойчивости к метотрексату, ген неомизинфосфотрансферазы, гены нечувствительных к глифосату ферментов EPSPS, гены глифосат-оксидоредуктаз (GOX), ген *phnO* *E. coli*, эквиваленты всего вышеперечисленного и т.д.

Также, в объем настоящего изобретения включены векторы и другие последовательности, предназначенные для поддержания и манипуляций с модельными нуклеотидными последовательностями в лаборатории, а также для введения их в клетки-хозяева; примеры подобных последовательностей включают в себя без ограничения фаги, плазмиды, ВАС-миды, YAC-миды, космиды и т.д.

Трансформированные растения также включены в объем настоящего изобретения. В особенности, настоящее изобретение относится к трансформированным растениям, содержащим по меньшей мере инсектицидный фрагмент белка Cry1A.105. В объем настоящего изобретения включены как однодольные, так и двудольные трансформированные растения. Однодольные растения включают в себя без ограничения кукурузу, пшеницу, овес, рис, сорго, гречиху, рожь, злаки (овсяницу, тимopheевку луговую, костер, ежу сборную, августинову траву, свиной пальчатый, полевицу) и ячмень; двудольные растения включают в себя без ограничения люцерну, яблоко, абрикос, спаржу, фасоль, кофе, ежевику, чернику, канолу, морковь, цветную капусту, сельдерей, вишню, нут, цитрусовые, хлопчатник, вигну китайскую, клюкву, огурец, тыкву, баклажан, плодовые деревья, виноград, лимон, салат-латук, лен, дыню, горчицу, орехоплодные деревья, бамию, округу, апельсин, горох, персик, арахис, грушу, сливу, картофель, сою, крупноплодную тыкву, землянику, сахарную свеклу, подсолнук, батат, табак, томат, репу и другие овощи. Полученная из этих растений продукция, равно как семена и ткани последних, также включены в объем настоящего изобретения в том случае, если эти семена, ткани или продукты содержат трансген, кодирующий инсектицидный фрагмент белка Cry1A.105.

Настоящее изобретение также относится к способам детекции белка Cry1A.105 или нуклеотидной последовательности, кодирующей инсектицидный фрагмент белка Cry1A.105, в биологической пробе. Cry1A.105 может быть использован для иммунизации животных с целью получения антител, специфичных к эпитопам Cry1A.105. Cry1A.105-специфичные антитела могут быть использованы для обнаружения белка Cry1A.105 в биологической пробе. Способы детектирования связывания антиген-антитело хорошо известны в данной области. Обнаружение связывания антитела с эпитопами Cry1A.105 является диагностичным в отношении присутствия в биологической пробе белка Cry1A.105.

Нуклеотидные последовательности, кодирующие инсектицидный фрагмент белка Cry1A.105, также могут быть детектированы в биологической пробе. Так, для связывания с детектируемой нуклеотидной последовательностью, кодирующей инсектицидный фрагмент белка Cry1A.105, может быть использован синтетический олигонуклеотидный зонд. Способы детектирования связывания зонд-мишень хорошо известны в данной области. Обнаружение связывания зонда с мишенью, т.е. с кодирующей последовательностью Cry1A.105, является диагностичным в отношении присутствия в биологической пробе кодирующей последовательности Cry1A.105.

Синтетические олигонуклеотидные праймеры могут быть использованы в реакции амплификации для получения ампликона из биологической пробы, предположительно содержащей нуклеотидную последовательность, кодирующую инсектицидный фрагмент белка Cry1A.105. Присутствие в реакционной смеси после проведения амплификации подобного ампликона является диагностичным в отношении присутствия в биологической пробе искомой нуклеотидной последовательности. В особенности полезными в качестве зондов, диагностических в отношении присутствия в биологической пробе кодирующих последовательностей белка Cry1A.105 согласно настоящему изобретению, являются последовательности, соответствующие или полностью комплементарные (1) нуклеотидам 1401-1420 последовательностей SEQ ID NO:1 или SEQ ID NO:3, либо (2) нуклеотидам 1821-1840 последовательностей SEQ ID NO:1 или SEQ

ID NO:3. Эти последовательности соответствуют (1) 20-ти нуклеотидам, перекрывающим область стыка между доменами II и III, принадлежащих сегментам различных инсектицидных белков, использованных для конструирования инсектицидного фрагмента белков, составляющих предмет настоящего изобретения, и (2) 20-ти нуклеотидам, перекрывающим область стыка домена III и протоксинового кодирующего сегмента, принадлежащих кодирующим сегментам различных инсектицидных белков, использованных для конструирования кодирующей последовательности пре-про-токсинового белка Cry1Ab.105. Нуклеотидные последовательности, соответствующие или комплементарные любому из этих сегментов ДНК (1401-1420 или 1821-1840) могут быть использованы в качестве зондов для детектирования этих кодирующих последовательностей в биологических пробах. Обнаружение подобного связывания является диагностичным в отношении присутствия в биологической пробе обсуждаемых кодирующих последовательностей. Специалисту в данной области будет понятно, что любые последовательности, фланкирующие эти сегменты ДНК, могут быть использованы в качестве праймеров для амплификации сегментов различной длины в различных биологических пробах, и что обнаружение подобных ампликонов после окончания реакции амплификации является диагностичным в отношении присутствия подобных кодирующих последовательностей в биологической пробе. К примеру, первый праймер, нуклеотидная последовательность которого соответствует участку 1201-1220 последовательности SEQ ID NO:1, может быть использован в качестве прямого праймера, а второй праймер, нуклеотидная последовательность которого соответствует перевернутому на 180° комплементу участка 1581-1600 последовательности SEQ ID NO:1 - в качестве обратного праймера. При совместном использовании этих праймеров и биологической пробы, содержащей SEQ ID NO:1 в качестве матрицы, в реакции амплификации, получаемый ампликон имеет длину 400 нуклеотидов и представляет собой отрезок 1201-1600 последовательности SEQ ID NO:1, содержащий 20-нуклеотидный сегмент 1401-1420 последовательности SEQ ID NO:1 и, таким образом, являющийся диагностичным в отношении присутствия кодирующей последовательности Cry1A.105 в биологической пробе.

Настоящее изобретение также относится к набору для детектирования присутствия Cry1A.105 или нуклеотидной последовательности, кодирующей Cry1A.105, в биологической пробе. Набор содержит все необходимые реагенты и контрольные пробы, необходимые для выполнения определения целевых веществ, а также инструкции по применению.

Нижеследующие примеры иллюстрируют предпочтительные воплощения настоящего изобретения. Исходя из соображений спецификаций и практики настоящего изобретения и согласно раскрытому в настоящей заявке, специалисту в данной области также будут очевидны другие воплощения, входящие в объем настоящего изобретения. Описание и примеры должны рассматриваться только как пояснительные и неограничивающие сущность и объем настоящего изобретения, указанной в приведенной вслед за разделом "Примеры" формуле изобретения.

Примеры

Пример 1. В данном примере демонстрируются синтетические нуклеотидные последовательности, кодирующие инсектицидный белок Cry1A.105.

Для использования в двудольных растениях была сконструирована нуклеотидная последовательность SEQ ID NO:1, кодирующая инсектицидный белок Cry1A.105. Соответствующая транслированная аминокислотная последовательность приведена в SEQ ID NO:2. Токсин-кодирующий сегмент соответствует нуклеотидам в положении от приблизительно 1 до приблизительно 1830 (либо может иметь несколько большую или меньшую длину).

Для использования в однодольных растениях была сконструирована нуклеотидная последовательность SEQ ID NO:3, кодирующая инсектицидный белок Cry1A.105. Соответствующая транслированная аминокислотная последовательность приведена в SEQ ID NO:4. Токсин-кодирующий сегмент соответствует нуклеотидам в положении от приблизительно 1 до приблизительно 1830 (либо может иметь несколько большую или меньшую длину).

Нуклеотидные последовательности SEQ ID NO:1 и SEQ ID NO:3 в значительной степени эквивалентны друг другу. Так, SEQ ID NO:1 и SEQ ID NO:3 гомологичны друг по отношению к другу на 94,3%. Кодирующие последовательности идентичны друг другу в области положений нуклеотидов от приблизительно 1330 до приблизительно 3534. Токсин-кодирующая часть каждой из последовательностей соответствует нуклеотидам в области от приблизительно 1 до приблизительно 1830; эти сегменты последовательностей SEQ ID NO:1 и SEQ ID NO:3 идентичны друг другу приблизительно на 88,9%. Существенными отличиями между двумя последовательностями характеризуются регионы, соответствующие нуклеотидам от приблизительно 1 до приблизительно 1329, т.е. первые две трети токсин-кодирующего сегмента Cry1A.105. В этом регионе идентичность SEQ ID NO:1 и SEQ ID NO:3 составляет около 84,7% друг по отношению к другу.

Штамм *E. coli* (TOP10, Invitrogen, Inc.), трансформированный плазмидой pMON70522, содержащей бета-лактамазный селекционный маркер и кодирующую Cry1A.105 последовательность SEQ ID NO:3, был депонирован в Коллекцию Исследовательских Сельскохозяйственных Культур при руководстве Международного Депозитария (Agriculture Research Culture Collection (NRRL) International Depository Authority) по адресу 1815, North University Street, Peoria, Illinois 61604, США, в соответствии с Будапешт-

ской Конвенцией о международном признании депонирования микроорганизмов для патентных целей, под номером NRRL B-30873.

Пример 2. В данном примере демонстрируется трансгенный хлопчатник, экспрессирующий белок Cry1A.105.

Семена хлопчатника Delta и Pineland DP50 подвергали поверхностной стерилизации и проращивали в течение ночи. Экспланты меристемы выделяли, и первичные листья удаляли при помощи микродиссекции. Препарированные экспланты помещали в среду для трансфекции таким образом, чтобы меристемы были ориентированы перпендикулярно потоку частиц. Трансформирующий вектор pMON47740 содержал экспрессионную кассету SEQ ID NO:9. KpnI - фрагмент, содержащий маркерный ген GUS под контролем промотора e35S кодирующую последовательность направляемого в хлоропласты Cry1A.105 под контролем промотора eFMV, вырезали из плазмиды, очищали при помощи ВЭЖХ и использовали для "gus"-трансформации эксплантов меристемы. Очищенная ДНК, содержащая экспрессионную кассету Cry1A.105 и маркерный ген GUS осаждали на микроскопические золотые шарики и тонким слоем нанесли на листья майлара. Частицы с ДНК доставляли в ткань меристемы при помощи электрических разрядов в низком вакууме. После бомбардировки экспланты помещали в среду без гормонов и селекционных агентов. Листья регенерированных ростков собирали и ткань листьев анализировали на присутствие маркера GUS. Трансгенные растения с высоким уровнем экспрессии GUS далее помещали в теплицы для дальнейших тестов. Отобранные растения снова тестировали на экспрессию GUS и негативные части растений обрезали. Данный цикл пробоотборов и обрезаний продолжали до тех пор, пока все секторы всех растений не оказались GUS-позитивными. После этого растения выращивали в стандартных тепличных условиях до момента сбора семян.

Ткани GUS-позитивных трансгенных растений F1 далее тестировали на наличие инсектицидной активности против совки хлопковой (CBW) и совки травяной (FAW). В качестве позитивного контроля использовали ранее полученные изогенные растения, экспрессирующие инсектицидные количества Cry1Ac или комбинации Cry1Ac и Cry2Ab, в качестве негативного контроля использовали изогенные нетрансгенные растения.

В качестве одного из способов определения инсектицидной активности трансгенного хлопчатника был выбран метод листовых квадратов (т.н. "CBW square assay") (Adamczyk с соавт., (2001) J. Econ. Entomol. 94:284-290; Kranthi с соавт., (2005) Current Science 89:291-298). Из листьев вырезали квадраты (величиной со спичечную головку и более), которые помещали в лунки аналитического планшета по одному квадрату на лунку. На каждый квадрат помещали одну гусеницу CBW третьей возрастной стадии. Через пять дней подсчитывали количество выживших насекомых.

Для исследования инсектицидной активности ткани выделенных из трансгенных растений семенных коробочек применяли т.н. "CBW boll assay". С каждого трансгенного растения после цветения собирали по 8 ярко-зеленых семенных коробочек, которые помещали в индивидуальные емкости, куда далее сажали одну гусеницу CBW третьей возрастной стадии. Через пять дней подсчитывали количество выживших насекомых.

Листовые анализы применяли для изучения инсектицидной активности трансгенной листовой ткани против FAW. С верхушек растений хлопчатника собирали новые листья. В каждую из 16 лунок аналитического планшета помещали по 2 круглых кусочка листа, полученных при помощи дыропробивного пресса (диаметр приблизительно 3/4"). В каждую лунку помещали по одной гусенице FAW второй или третьей возрастной стадии. Через пять дней подсчитывали количество выживших насекомых.

Результаты проведенных исследований приведены в таблице 1. Полученные данные указывают на то, что трансгенный хлопчатник, экспрессирующий Cry1A.105, демонстрирует большую инсектицидную активность по отношению к FAW и CBW, чем трансгенные растения, экспрессирующие Cry1Ac или комбинацию Cry1Ac и Cry2Ab.

Таблица 1 Результаты биоанализа FAW и CBW с использованием тканей трансгенного хлопчатника

Растение	FAW (% выживших особей) (ткань листа)	CBW (% выживших особей) (листовые квадраты)	CBW (% выживших особей) (семенные коробочки)
Cry1Ac/Cry2Ab	74,5	32,0	35,8
Cry1Ac	92,7	35,5	35,8
Изогенный негативный контроль	99,6	96,8	54
17238	10,9	9,4	25
17567	0	12,5	12,5
17774	1,6	1,2	0
17875	3,1	4,2	0
18026	1,6	18,8	12,5
18122	7,8	22,9	0

Помимо экспериментов, результаты которых приведены в табл. 1, аналогичные исследования были проведены на гусеницах табачной листовертки-почкоеда и совки хлопковой американской. Во всех экспериментах, растения, экспрессирующие Cry1A.105, демонстрировали инсектицидную активность также и против этих видов.

Пример 3. Данный пример демонстрирует трансгенную кукурузу, экспрессирующую белок Cry1A.105.

Растения трансгенной кукурузы регенерировали из клеток, трансформированных вектором pMON40232. pMON40232 содержит экспрессионную кассету, имеющую нуклеотидную последовательность SEQ ID NO:7 и содержащую оперативно связанные друг с другом усиленный промотор C7AMV 35S, лидерную последовательность гена САВ пшеницы, интрон гена актина 1 риса, кодирующую последовательность Cry1A.105 и 3'-сигналы терминации трансляции и полиаденилирования гена hsp17 пшеницы. Нуклеотидная последовательность хлоропласт-таргеттирующего пептида гена EPSPS *Arabidopsis thaliana* (At.EPSES-CTP2) помещена с 5'-конца относительно кодирующей последовательности Cry1A.105 в одной рамке считывания с последней. pMON40232 содержит предназначенный для селекции трансгенных растений рекомбинантный ген, кодирующий фермент EPSPS, нечувствительный к гербициду глифосату. Трансгенные растения, полученные из тканей, трансформированных pMON4 0232, кодировали как LAJ 105. Трансгенные растения скринировали на отсутствие каких-либо остатков векторных последовательностей за пределами экспрессионной кассеты, на присутствие одной простой интегрированной последовательности, а также на интактность экспрессионной кассеты, содержащей нуклеотидную последовательность, кодирующую белок Cry1A.105.

Трансгенные растения, удовлетворяющие всем вышеуказанным скрининговым условиям, далее подвергали различным биологическим тестам. Трансгенные растения кукурузы LAJ105 сравнивали с изогенной кукурузой LH198 в качестве негативного контроля, а также с экспрессирующим инсектицидный фрагмент белка Cry1Ab сортом MON810. С каждого из Cry1A.105-трансгенных и контрольных растений получали по пять дисков листовой ткани диаметром приблизительно 10 мм каждый. Диски помещали в заполненные агаром лунки для поддержания тканей в набухшем состоянии, после чего скармливали новорожденным гусеницам FAW, совки-ипсилон (BCW), мотылька кукурузного (ECB), совки хлопковой американской (CEW) и огневки кукурузной юго-западной (SWCB). В каждую лунку сажали по одной новорожденной гусенице FAW, по одной новорожденной гусенице CEW, по две новорожденных гусеницы BCW, по две новорожденных гусеницы SWCB, либо по четыре новорожденных гусеницы ECB. Повреждения листа в результате поедания гусеницами оценивали через четыре дня при помощи шкалы повреждения листа (LDR), в которой балл 0 соответствует отсутствию видимых повреждений, балл 11 соответствует уничтожению половины (50%) каждого из дисков, при этом каждый балл в промежутке между этими значениями соответствует увеличению видимых повреждений листовых дисков на 5%.

Результаты биологических тестов указывают на то, что растения, экспрессирующие белок Cry1A.105, демонстрируют большую инсектицидную активность (LDR) по отношению к FAW, ECB и CEW, нежели экспрессирующие Cry1Ab контрольные растения. Так, значения LDR для трех вышеуказанных вредителей составляли менее 1 при тестировании Cry1A.105-трансгенов по сравнению с приблизительно 8 - приблизительно 10 при тестировании Cry1Ab-контролей. В то же время, при тестировании этих же растений на SWCB, показатели LDR как для Cry1A. 105-трансгенов, так и для Cry1Ab-контролей, составляли 1-2, что указывает на то, что белок Cry1A.105 не является более токсичным по отношению к SWCB, нежели Cry1Ab. Результаты этих тестов подтверждают предыдущие данные, указывающие на то, что Cry1Ab не эффективен для борьбы с BCW. Cry1A.105-трансгены не являются более эффективными против BCW, нежели Cry1Ab-контроли. Таким образом, при уровнях экспрессии белка Cry1A.105, достижимых *in planta*, полученные трансгены могут быть эффективны для борьбы с чешуекрылыми насекомыми-вредителями, являющимися представителями других родов, включающих в себя без ограничения *Anticarsia*, *Pseudoplusia*, *Rachiplusia*, *Heliothis*, *Helicoverpa*, *Spodoptera*, *Epinotia* и *Argemera*.

СПИСОК ПОСЛЕДОВАТЕЛЬНОСТЕЙ

<110> Monsanto Technology LLC
 Bogdanova, Natalia N.
 Corbin, David R.
 Malvar, Thomas M.
 Perlak, Frederick J.
 Roberts, James K.
 Romano, Charles P.

<120> НУКЛЕОТИДНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ, КОДИРУЮЩИЕ ИНСЕКТИЦИДНЫЕ БЕЛКИ

<130> 38-21 (52905) A

<150> 60/713144

<151> 2005-08-31

<160> 8

<170> PatentIn version 3.3

<210> 1

<211> 3534

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Синтетическая нуклеотидная последовательность, кодирующая аминокислотную последовательность Cry1A.105

<220>

<221> Кодированная последовательность

<222> (1)..(3534)

<400> 1

atg gac aac aac cca aac atc aac gaa tgc att cca tac aac tgc ttg	48
Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu	
1 5 10 15	
agt aac cca gaa gtt gaa gta ctt ggt gga gaa cgc att gaa acc ggt	96
Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly	
20 25 30	
tac act ccc atc gac atc tcc ttg tcc ttg aca cag ttt ctg ctc agc	144
Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser	
35 40 45	
gag ttc gtg cca ggt gct ggg ttc gtt ctc gga cta gtt gac atc atc	192
Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile	
50 55 60	
tgg ggt atc ttt ggt cca tct caa tgg gat gca ttc ctg 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 ttg atc aac cag agg atc gaa gag ttc gcc agg aac cag gcc	288
Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala	
85 90 95	

015908

atc tct agg ttg gaa gga ttg agc aat ctc tac caa atc tat gca gag Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu 100 105 110	336
agc ttc aga gag tgg gaa gcc gat cct act aac cca gct ctc cgc gag Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu 115 120 125	384
gaa atg cgt att caa ttc aac gac atg aac agc gcc ttg acc aca gct Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala 130 135 140	432
atc cca ttg ttc gca gtc cag aac tac caa gtt cct ctc ttg tcc 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 gca gct aat ctt cac ctc agc gtg ctt cga gac gtt agc Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser 165 170 175	528
gtg ltt ggg caa agg lgg gga ltc gal gcl gca acc atc aat agc cgt Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg 180 185 190	576
tac aac gac ctt act agg ctg att gga aac tac acc gac cac gct gtt Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp His Ala Val 195 200 205	624
cgt tgg tac aac act ggc ttg gag cgt gtc tgg ggt cct 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 att aga tac aac cag ttc agg aga gaa ttg acc ctc aca gtt Asp Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val 225 230 235 240	720
ttg gac att gtg tct ctc ttc ccg aac tat gac tcc aga acc tac cct Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr Tyr Pro 245 250 255	768
atc cgt aca gtg tcc caa ctt acc aga gaa atc tat act aac cca gtt Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val 260 265 270	816
ctt gag aac ttc gac ggt agc ttc cgt ggt tct gcc caa ggt atc gaa Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu 275 280 285	864
ggc tcc atc agg agc cca cac ttg atg gac atc ttg aac agc ata act Gly Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr 290 295 300	912
atc tac acc gat gct cac aga gga gag tat tac tgg tct gga cac cag Ile Tyr Thr Asp Ala His Arg Gly Glu Tyr Tyr Trp Ser Gly His Gln 305 310 315 320	960
atc atg gcc tct cca gtt gga ttc agc ggg ccc gag ttt acc ttt cct Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro 325 330 335	1008
ctc tat gga act atg gga aac gcc gct cca caa caa cgt atc gtt gct	1056

015908

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

015908

580					585					590						
tca	ggg	aat	gaa	ggt	tac	atc	gac	agg	ttt	gaa	ttg	att	cca	gtt	act	1824
Ser	Gly	Asn	Glu	Val	Tyr	Ile	Asp	Arg	Phe	Glu	Leu	Ile	Pro	Val	Thr	
		595					600					605				
gca	acc	ctc	gag	gct	gag	tac	aac	ctt	gag	aga	gcc	cag	aag	gct	gtg	1872
Ala	Thr	Leu	Glu	Ala	Glu	Tyr	Asn	Leu	Glu	Arg	Ala	Gln	Lys	Ala	Val	
	610					615					620					
aac	gcc	ctc	ttt	acc	tcc	acc	aat	cag	ctt	ggc	ttg	aaa	act	aac	gtt	1920
Asn	Ala	Leu	Phe	Thr	Ser	Thr	Asn	Gln	Leu	Gly	Leu	Lys	Thr	Asn	Val	
	625				630					635					640	
act	gac	tat	cac	att	gac	caa	gtg	tcc	aac	ttg	gtc	acc	tac	ctt	agc	1968
Thr	Asp	Tyr	His	Ile	Asp	Gln	Val	Ser	Asn	Leu	Val	Thr	Tyr	Leu	Ser	
				645					650					655		
gat	gag	ttc	tgc	ctc	gac	gag	aag	cgt	gaa	ctc	tcc	gag	aaa	ggt	aaa	2016
Asp	Glu	Phe	Cys	Leu	Asp	Glu	Lys	Arg	Glu	Leu	Ser	Glu	Lys	Val	Lys	
		660						665						670		
cac	gcc	aag	cgt	ctc	agc	gac	gag	agg	aat	ctc	ttg	caa	gac	tcc	aac	2064
His	Ala	Lys	Arg	Leu	Ser	Asp	Glu	Arg	Asn	Leu	Leu	Gln	Asp	Ser	Asn	
		675					680						685			
ttc	aaa	gac	atc	aac	agg	cag	cca	gaa	cgt	ggt	tgg	ggt	gga	agc	acc	2112
Phe	Lys	Asp	Ile	Asn	Arg	Gln	Pro	Glu	Arg	Gly	Trp	Gly	Gly	Ser	Thr	
		690				695					700					
ggg	atc	acc	atc	caa	gga	ggc	gac	gat	gtg	ttc	aag	gag	aac	tac	gtc	2160
Gly	Ile	Thr	Ile	Gln	Gly	Gly	Asp	Asp	Val	Phe	Lys	Glu	Asn	Tyr	Val	
	705				710					715					720	
acc	ctc	tcc	gga	act	ttc	gac	gag	tgc	tac	cct	acc	tac	ttg	tac	cag	2208
Thr	Leu	Ser	Gly	Thr	Phe	Asp	Glu	Cys	Tyr	Pro	Thr	Tyr	Leu	Tyr	Gln	
				725					730					735		
aag	atc	gat	gag	tcc	aaa	ctc	aaa	gcc	ttc	acc	agg	tat	caa	ctt	aga	2256
Lys	Ile	Asp	Glu	Ser	Lys	Leu	Lys	Ala	Phe	Thr	Arg	Tyr	Gln	Leu	Arg	
		740						745					750			
ggc	tac	atc	gaa	gac	agc	caa	gac	ctt	gaa	atc	tac	tcg	atc	agg	tac	2304
Gly	Tyr	Ile	Glu	Asp	Ser	Gln	Asp	Leu	Glu	Ile	Tyr	Ser	Ile	Arg	Tyr	
		755					760						765			
aat	gcc	aag	cac	gag	acc	gtg	aat	gtc	cca	ggt	act	ggt	tcc	ctc	tgg	2352
Asn	Ala	Lys	His	Glu	Thr	Val	Asn	Val	Pro	Gly	Thr	Gly	Ser	Leu	Trp	
		770					775					780				
cca	ctt	tct	gcc	caa	tct	ccc	att	ggg	aag	tgt	gga	gag	cct	aac	aga	2400
Pro	Leu	Ser	Ala	Gln	Ser	Pro	Ile	Gly	Lys	Cys	Gly	Glu	Pro	Asn	Arg	
				785			790			795					800	
tgc	gct	cca	cac	ctt	gag	tgg	aat	cct	gac	ttg	gac	tgc	tcc	tgc	agg	2448
Cys	Ala	Pro	His	Leu	Glu	Trp	Asn	Pro	Asp	Leu	Asp	Cys	Ser	Cys	Arg	
				805					810					815		
gat	ggc	gag	aag	tgt	gcc	cac	cat	tct	cat	cac	ttc	tcc	ttg	gac	atc	2496
Asp	Gly	Glu	Lys	Cys	Ala	His	His	Ser	His	His	Phe	Ser	Leu	Asp	Ile	
			820					825						830		

015908

gat gtg gga tgt act gac ctg aat gag gac ctc gga gtc tgg gtc atc Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile 835 840 845	2544
ttc aag atc aag acc caa gac gga cac gca aga ctt ggc aac ctt gag Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu 850 855 860	2592
ttt ctc gaa gag aaa cca ttg gtc ggt gaa gct ctc gct cgt gtg aag Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys 865 870 875 880	2640
aga gca gag aag aag tgg agg gac aaa cgt gag aaa ctc gaa tgg gaa Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu 885 890 895	2688
act aac atc gtt tac aag gag gcc aaa gag tcc gtg gat gct ttg ttc Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe 900 905 910	2736
gtg aac tcc caa tat gat cag ttg caa gcc gac acc aac atc gcc atg Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met 915 920 925	2784
atc cac gcc gca gac aaa cgt gtg cac agc att cgt gag gct tac ttg Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu 930 935 940	2832
cct gag ttg tcc gtg atc cct ggt gtg aac gct gcc atc ttc gag gaa Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu 945 950 955 960	2880
ctt gag gga cgt atc ttt acc gca ttc tcc ttg tac gat gcc aga aac Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn 965 970 975	2928
gtc atc aag aac ggt gac ttc aac aat ggc ctc agc tgc tgg aat gtg Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val 980 985 990	2976
aaa ggt cat gtg gac gtg gag gaa cag aac aat cag cgt tcc gtc ctg Lys Gly His Val Asp Val Glu Gln Asn Asn Gln Arg Ser Val Leu 995 1000 1005	3024
gtt gtg cct gag tgg gaa gct gaa gtg tcc caa gag gtt aga gtc Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val 1010 1015 1020	3069
tgt cca ggt aga ggc tac att ctc cgt gtg acc gct tac aag gag Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu 1025 1030 1035	3114
gga tac ggt gag ggt tgc gtg acc atc cac gag atc gag aac aac Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn 1040 1045 1050	3159
acc gac gag ctt aag ttc tcc aac tgc gtc gag gaa gaa atc tat Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr 1055 1060 1065	3204

015908

ccc aac aac acc gtt act tgc aac gac tac act gtg aat cag gaa 3249
Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu
1070 1075 1080
gag tac gga ggt gcc tac act agc cgt aac aga ggt tac aac gaa 3294
Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu
1085 1090 1095
gct cct tcc gtt cct gct gac tat gcc tcc gtg tac gag gag aaa 3339
Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys
1100 1105 1110
tcc tac aca gat ggc aga cgt gag aac cct tgc gag ttc aac aga 3384
Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg
1115 1120 1125
ggt tac agg gac tac aca cca ctt cca gtt ggc tat gtt acc aag 3429
Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys
1130 1135 1140
gag ctt gag tac ttt cct gag acc gac aaa gtg tgg atc gag atc 3474
Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile
1145 1150 1155
ggt gaa acc gag gga acc ttc atc gtg gac agc gtg gag ctt ctc 3519
Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu
1160 1165 1170
ttg atg gag gaa taa 3534
Leu Met Glu Glu
1175

<210> 2
<211> 1177
<212> Белок
<213> Искусственная последовательность

<220>
<223> Синтетический конструктор

<400> 2

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

015908

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 His 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 Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val
 225 230 235 240

Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr 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

Gly 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 Glu Tyr Tyr Trp Ser Gly His Gln
 305 310 315 320

015908

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

Thr Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro Gly Phe Thr Gly
485 490 495

Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr Ile
500 505 510

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

Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu
530 535 540

Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly Asp Pro
545 550 555 560

015908

Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr
565 570 575

Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe Ser
580 585 590

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

Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala Val
610 615 620

Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val
625 630 635 640

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser
645 650 655

Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
660 665 670

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn
675 680 685

Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr
690 695 700

Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
705 710 715 720

Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
725 730 735

Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg
740 745 750

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Ser Ile Arg Tyr
755 760 765

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
770 775 780

Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg
785 790 795 800

Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg

015908

805 810 815
 Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile
 820 825 830
 Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile
 835 840 845
 Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu
 850 855 860
 Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys
 865 870 875 880
 Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu
 885 890 895
 Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe
 900 905 910
 Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met
 915 920 925
 Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu
 930 935 940
 Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu
 945 950 955 960
 Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn
 965 970 975
 Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val
 980 985 990
 Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu
 995 1000 1005
 Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val
 1010 1015 1020
 Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu
 1025 1030 1035
 Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn
 1040 1045 1050

015908

Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr
 1055 1060 1065

Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu
 1070 1075 1080

Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu
 1085 1090 1095

Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys
 1100 1105 1110

Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg
 1115 1120 1125

Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys
 1130 1135 1140

Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile
 1145 1150 1155

Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu
 1160 1165 1170

Leu Met Glu Glu
 1175

<210> 3
 <211> 3534
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Синтетическая нуклеотидная последовательность, кодирующая
 аминокислотную последовательность Cry1A.105

<220>
 <221> Кодирующая последовательность
 <222> (1)..(3534)

<400> 3
 atg gac aac aac cca aac atc aac gag tgc atc ccg tac aac tgc ctc 48
 Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
 1 5 10 15

agc aac cct gag gtc gag gtg ctc gcc ggt gag cgc atc gag acc ggt 96
 Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly
 20 25 30

tac acc ccc atc gac atc tcc ctc tcc ctc acg cag ttc ctg ctc agc	144
Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser	
35 40 45	
gag ttc gtg cca ggc gct ggc ttc gtc ctg ggc ctc gtg gac atc atc	192
Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile	
50 55 60	
tgg ggc atc ttt ggc ccc tcc cag tgg gac gcc ttc ctg gtg caa atc	240
Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile	
65 70 75 80	
gag cag ctc atc aac cag agg atc gag gag ttc gcc agg aac cag gcc	288
Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala	
85 90 95	
atc agc cgc ctg gag ggc ctc agc aac ctc tac caa atc tac gct gag	336
Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu	
100 105 110	
agc ttc cgc gag tgg gag gcc gac ccc act aac cca gct ctc cgc gag	384
Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu	
115 120 125	
gag atg cgc atc cag ttc aac gac atg aac agc gcc ctg acc acc gcc	432
Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala	
130 135 140	
atc cca ctc ttc gcc gtc cag aac tac caa gtc ccg ctc ctg tcc gtg	480
Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val	
145 150 155 160	
tac gtc cag gcc gcc aac ctg cac ctc agc gtg ctg agg gac gtc agc	528
Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser	
165 170 175	
gtg ttt ggc cag agg tgg ggc ttc gac gcc gcc acc atc aac agc cgc	576
Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg	
180 185 190	
tac aac gac ctc acc agg ctg atc ggc aac tac acc gac cac gct gtc	624
Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp His Ala Val	
195 200 205	
cgc tgg tac aac act ggc ctg gag cgc gtc tgg ggc cct gat tct aga	672
Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg	
210 215 220	
gac tgg att cgc tac aac cag ttc agg cgc gag ctg acc ctc acc gtc	720
Asp Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val	
225 230 235 240	
ctg gac att gtg tcc ctc ttc ccg aac tac gac tcc cgc acc tac ccg	768
Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr Tyr Pro	
245 250 255	
atc cgc acc gtg tcc caa ctg acc cgc gaa atc tac acc aac ccc gtc	816
Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val	
260 265 270	
ctg gag aac ttc gac ggt agc ttc agg ggc agc gcc cag ggc atc gag	864

015908

Leu	Glu	Asn	Phe	Asp	Gly	Ser	Phe	Arg	Gly	Ser	Ala	Gln	Gly	Ile	Glu		
		275					280					285					
ggc	tcc	atc	agg	agc	cca	cac	ctg	atg	gac	atc	ctc	aac	agc	atc	act		912
Gly	Ser	Ile	Arg	Ser	Pro	His	Leu	Met	Asp	Ile	Leu	Asn	Ser	Ile	Thr		
	290					295					300						
atc	tac	acc	gat	gcc	cac	cgc	ggc	gag	tac	tac	tgg	tcc	ggc	cac	cag		960
Ile	Tyr	Thr	Asp	Ala	His	Arg	Gly	Glu	Tyr	Tyr	Trp	Ser	Gly	His	Gln		
305				310					315						320		
atc	atg	gcc	tcc	ccg	gtc	ggc	ttc	agc	ggc	ccc	gag	ttt	acc	ttt	cct		1008
Ile	Met	Ala	Ser	Pro	Val	Gly	Phe	Ser	Gly	Pro	Glu	Phe	Thr	Phe	Pro		
				325					330					335			
ctc	tac	ggc	acg	atg	ggc	aac	gcc	gct	cca	caa	caa	cgc	atc	gtc	gct		1056
Leu	Tyr	Gly	Thr	Met	Gly	Asn	Ala	Ala	Pro	Gln	Gln	Arg	Ile	Val	Ala		
			340				345						350				
cag	ctg	ggc	cag	ggc	gtc	tac	cgc	acc	ctg	agc	tcc	acc	ctg	tac	cgc		1104
Gln	Leu	Gly	Gln	Gly	Val	Tyr	Arg	Thr	Leu	Ser	Ser	Thr	Leu	Tyr	Arg		
	355						360					365					
agg	ccc	ttc	aac	atc	ggc	atc	aac	aac	cag	cag	ctg	tcc	gtc	ctg	gat		1152
Arg	Pro	Phe	Asn	Ile	Gly	Ile	Asn	Asn	Gln	Gln	Leu	Ser	Val	Leu	Asp		
	370					375					380						
ggc	act	gag	ttc	gcc	tac	ggc	acc	tcc	tcc	aac	ctg	ccc	tcc	gct	gtc		1200
Gly	Thr	Glu	Phe	Ala	Tyr	Gly	Thr	Ser	Ser	Asn	Leu	Pro	Ser	Ala	Val		
385				390						395				400			
tac	cgc	aag	agc	ggc	acg	gtg	gat	tcc	ctg	gac	gag	atc	cca	cca	cag		1248
Tyr	Arg	Lys	Ser	Gly	Thr	Val	Asp	Ser	Leu	Asp	Glu	Ile	Pro	Pro	Gln		
				405					410					415			
aac	aac	aat	gtg	ccc	ccc	agg	cag	ggt	ttt	tcc	cac	agg	ctc	agc	cac		1296
Asn	Asn	Asn	Val	Pro	Pro	Arg	Gln	Gly	Phe	Ser	His	Arg	Leu	Ser	His		
			420				425						430				
gtg	tcc	atg	ttc	cgc	tcc	ggc	ttc	agc	aac	tcg	tcc	gtg	agc	atc	atc		1344
Val	Ser	Met	Phe	Arg	Ser	Gly	Phe	Ser	Asn	Ser	Ser	Val	Ser	Ile	Ile		
		435				440						445					
aga	gct	cct	atg	ttc	tct	tgg	ata	cac	cgt	agt	gct	gag	ttc	aac	aac		1392
Arg	Ala	Pro	Met	Phe	Ser	Trp	Ile	His	Arg	Ser	Ala	Glu	Phe	Asn	Asn		
	450					455					460						
atc	att	gca	tcc	gac	agc	att	act	caa	ata	ccc	ttg	gtg	aaa	gca	cat		1440
Ile	Ile	Ala	Ser	Asp	Ser	Ile	Thr	Gln	Ile	Pro	Leu	Val	Lys	Ala	His		
465				470						475				480			
aca	ctt	cag	tca	ggc	act	act	ggt	gtc	aga	ggt	cca	ggg	ttt	aca	gga		1488
Thr	Leu	Gln	Ser	Gly	Thr	Thr	Val	Val	Arg	Gly	Pro	Gly	Phe	Thr	Gly		
				485				490						495			
gga	gac	att	ctt	cgt	cgc	aca	agt	gga	gga	ccc	ttt	gct	tac	act	att		1536
Gly	Asp	Ile	Leu	Arg	Arg	Thr	Ser	Gly	Gly	Pro	Phe	Ala	Tyr	Thr	Ile		
			500					505					510				
ggt	aac	atc	aat	ggc	caa	ttg	ccc	caa	agg	tat	cgt	gca	aga	atc	cgc		1584
Val	Asn	Ile	Asn	Gly	Gln	Leu	Pro	Gln	Arg	Tyr	Arg	Ala	Arg	Ile	Arg		

015908

515	520	525	
tat gcc tct act aca aat ctc agg atc tac gtg act gtt gca ggt gaa			1632
Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu			
530	535	540	
agg atc ttt gct ggt cag ttc aac aag act atg gat acc ggt gac cct			1680
Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly Asp Pro			
545	550	555	560
ttg aca ttc caa tct ttt agc tac gca act atc aac aca gct ttt aca			1728
Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr			
	565	570	575
ttc cca atg agc cag agt agc ttc aca gta ggt gct gac act ttc agc			1776
Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe Ser			
	580	585	590
tca ggg aat gaa gtt tac atc gac agg ttt gaa ttg att cca gtt act			1824
Ser Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu Ile Pro Val Thr			
	595	600	605
gca acc ctc gag gct gag tac aac ctt gag aga gcc cag aag gct gtg			1872
Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala Val			
610	615	620	
aac gcc ctc ttt acc tcc acc aat cag ctt ggc ttg aaa act aac gtt			1920
Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val			
625	630	635	640
act gac tat cac att gac caa gtg tcc aac ttg gtc acc tac ctt agc			1968
Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser			
	645	650	655
gat gag ttc tgc ctc gac gag aag cgt gaa ctc tcc gag aaa gtt aaa			2016
Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys			
	660	665	670
cac gcc aag cgt ctc agc gac gag agg aat ctc ttg caa gac tcc aac			2064
His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn			
	675	680	685
ttc aaa gac atc aac agg cag cca gaa cgt ggt tgg ggt gga agc acc			2112
Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr			
690	695	700	
ggg atc acc atc caa gga ggc gac gat gtg ttc aag gag aac tac gtc			2160
Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val			
705	710	715	720
acc ctc tcc gga act ttc gac gag tgc tac cct acc tac ttg tac cag			2208
Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln			
	725	730	735
aag atc gat gag tcc aaa ctc aaa gcc ttc acc agg tat caa ctt aga			2256
Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg			
	740	745	750
ggc tac atc gaa gac agc caa gac ctt gaa atc tac tcg atc agg tac			2304
Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Ser Ile Arg Tyr			
	755	760	765

015908

aat gcc aag cac gag acc gtg aat gtc cca ggt act ggt tcc ctc tgg	2352
Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp	
770 .775 780	
cca ctt tct gcc caa tct ccc att ggg aag tgt gga gag cct aac aga	2400
Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg	
785 790 795 800	
tgc gct cca cac ctt gag tgg aat cct gac ttg gac tgc tcc tgc agg	2448
Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg	
805 810 815	
gat ggc gag aag tgt gcc cac cat tct cat cac ttc tcc ttg gac atc	2496
Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile	
820 825 830	
gat gtg gga tgt act gac ctg aat gag gac ctc gga gtc tgg gtc atc	2544
Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile	
835 840 845	
ttc aag atc aag acc caa gac gga cac gca aga ctt ggc aac ctt gag	2592
Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu	
850 855 860	
ttt ctc gaa gag aaa cca ttg gtc ggt gaa gct ctc gct cgt gtg aag	2640
Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys	
865 870 875 880	
aga gca gag aag aag tgg agg gac aaa cgt gag aaa ctc gaa tgg gaa	2688
Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu	
885 890 895	
act aac atc gtt tac aag gag gcc aaa gag tcc gtg gat gct ttg ttc	2736
Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe	
900 905 910	
gtg aac tcc caa tat gat cag ttg caa gcc gac acc aac atc gcc atg	2784
Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met	
915 920 925	
atc cac gcc gca gac aaa cgt gtg cac agc att cgt gag gct tac ttg	2832
Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu	
930 935 940	
cct gag ttg tcc gtg atc cct ggt gtg aac gct gcc atc ttc gag gaa	2880
Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu	
945 950 955 960	
ctt gag gga cgt atc ttt acc gca ttc tcc ttg tac gat gcc aga aac	2928
Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn	
965 970 975	
gtc atc aag aac ggt gac ttc aac aat ggc ctc agc tgc tgg aat gtg	2976
Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val	
980 985 990	
aaa ggt cat gtg gac gtg gag gaa cag aac aat cag cgt tcc gtc ctg	3024
Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu	
995 1000 1005	

015908

gtt gtg cct gag tgg gaa gct gaa gtg tcc caa gag gtt aga gtc 3069
 Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val
 1010 1015 1020
 tgt cca ggt aga ggc tac att ctc cgt gtg acc gct tac aag gag 3114
 Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu
 1025 1030 1035
 gga tac ggt gag ggt tgc gtg acc atc cac gag atc gag aac aac 3159
 Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn
 1040 1045 1050
 acc gac gag ctt aag ttc tcc aac tgc gtc gag gaa gaa atc tat 3204
 Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr
 1055 1060 1065
 ccc aac aac acc gtt act tgc aac gac tac act gtg aat cag gaa 3249
 Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu
 1070 1075 1080
 gag tac gga ggt gcc tac act agc cgt aac aga ggt tac aac gaa 3294
 Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu
 1085 1090 1095
 gct cct tcc gtt cct gct gac tat gcc tcc gtg tac gag gag aaa 3339
 Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys
 1100 1105 1110
 tcc tac aca gat ggc aga cgt gag aac cct tgc gag ttc aac aga 3384
 Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg
 1115 1120 1125
 ggt tac agg gac tac aca cca ctt cca gtt ggc tat gtt acc aag 3429
 Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys
 1130 1135 1140
 gag ctt gag tac ttt cct gag acc gac aaa gtg tgg atc gag atc 3474
 Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile
 1145 1150 1155
 ggt gaa acc gag gga acc ttc atc gtg gac agc gtg gag ctt ctc 3519
 Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu
 1160 1165 1170
 ttg atg gag gaa taa 3534
 Leu Met Glu Glu
 1175

<210> 4
 <211> 1177
 <212> Белок
 <213> Искусственная последовательность

<220>
 <223> Синтетический конструктор

<400> 4

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

015908

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 His 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 Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val
 225 230 235 240

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

015908

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

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

Thr Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro Gly Phe Thr Gly
485 490 495

Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr Ile
 500 505 510

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

Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu
 530 535 540

Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly Asp Pro
 545 550 555 560

Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr
 565 570 575

Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe Ser
 580 585 590

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

Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala Val
 610 615 620

Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val
 625 630 635 640

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser
 645 650 655

Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
 660 665 670

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn
 675 680 685

Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr
 690 695 700

Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
 705 710 715 720

Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
 725 730 735

015908

Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg
740 745 750

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Ser Ile Arg Tyr
755 760 765

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
770 775 780

Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg
785 790 795 800

Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg
805 810 815

Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile
820 825 830

Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile
835 840 845

Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu
850 855 860

Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys
865 870 875 880

Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu
885 890 895

Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe
900 905 910

Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met
915 920 925

Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu
930 935 940

Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu
945 950 955 960

Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn
965 970 975

Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val

015908

		980						985						990	
Lys	Gly	His	Val	Asp	Val	Glu	Glu	Gln	Asn	Asn	Gln	Arg	Ser	Val	Leu
		995					1000							1005	
Val	Val	Pro	Glu	Trp	Glu	Ala	Glu	Val	Ser	Gln	Glu	Val	Arg	Val	
	1010					1015					1020				
Cys	Pro	Gly	Arg	Gly	Tyr	Ile	Leu	Arg	Val	Thr	Ala	Tyr	Lys	Glu	
	1025					1030					1035				
Gly	Tyr	Gly	Glu	Gly	Cys	Val	Thr	Ile	His	Glu	Ile	Glu	Asn	Asn	
	1040					1045					1050				
Thr	Asp	Glu	Leu	Lys	Phe	Ser	Asn	Cys	Val	Glu	Glu	Glu	Ile	Tyr	
	1055					1060					1065				
Pro	Asn	Asn	Thr	Val	Thr	Cys	Asn	Asp	Tyr	Thr	Val	Asn	Gln	Glu	
	1070					1075					1080				
Glu	Tyr	Gly	Gly	Ala	Tyr	Thr	Ser	Arg	Asn	Arg	Gly	Tyr	Asn	Glu	
	1085					1090					1095				
Ala	Pro	Ser	Val	Pro	Ala	Asp	Tyr	Ala	Ser	Val	Tyr	Glu	Glu	Lys	
	1100					1105					1110				
Ser	Tyr	Thr	Asp	Gly	Arg	Arg	Glu	Asn	Pro	Cys	Glu	Phe	Asn	Arg	
	1115					1120					1125				
Gly	Tyr	Arg	Asp	Tyr	Thr	Pro	Leu	Pro	Val	Gly	Tyr	Val	Thr	Lys	
	1130					1135					1140				
Glu	Leu	Glu	Tyr	Phe	Pro	Glu	Thr	Asp	Lys	Val	Trp	Ile	Glu	Ile	
	1145					1150					1155				
Gly	Glu	Thr	Glu	Gly	Thr	Phe	Ile	Val	Asp	Ser	Val	Glu	Leu	Leu	
	1160					1165					1170				
Leu	Met	Glu	Glu												
	1175														

<210> 5
 <211> 5480
 <212> ДНК
 <213> Искусственная последовательность
 <220>

<223> Экспрессионная кассета, кодирующая аминокислотную последовательность Cry1A.105

<220>
 <221> промотор
 <222> (1)..(981)
 <223> промотор FMV.e35S

<220>
 <221> 5'-нетранслируемый регион
 <222> (985)..(1080)

<220>
 <221> Транзитный пептид
 <222> (1090)..(1263)
 <223> Последовательность хлоропласт-таргетирующего пептида

<220>
 <221> Кодирующая последовательность
 <222> (1267)..(4797)
 <223> Кодирующая последовательность Cry1A.105

<220>
 <221> Терминатор
 <222> (4833)..(5480)
 <223> Сигналы терминации трансляции и полиаденилирования Ps.E9 rbcS

<400> 5
 aattctcagt cсаааgсctc аасааggтса gggтacаgаg tctccааacc attаgссааа 60
 аgсtаcаgga gatсаатgаа гаатсttсаа tсаааgтааа сtаctgттсc аgсacатgса 120
 tсаtggtсag тааgtttсag аааааgасат ссаccгаага сttаааgтта gtgggсатсt 180
 ttgаааgтаа tcttgтсаас atсgаgсagс tggcttgтgг gгаccаgаса аааааgгаат 240
 ggtгсагаат tgttaggсгс асctаccааа аgсатсtttg сctttattгс аааgатаааg 300
 сagattсctс tagтасааgt ggggaасааа атаасgtгга аааgаgсtgt сctгacаgсс 360
 сactсactаа тgсgtатgас гаасгсagтg асgaccасаа аагаатtagс ttgаgсtсag 420
 gatttagсag сattссagат tgggttсаат саасааggта сgаgссатат сactttattс 480
 ааattggtат сgссаааacc аагааgгаас tccсатсctс аааggттgt ааgгаагаат 540
 tctсagтсса ааgсctсаас ааggтсaggg тacаgаgtct ссааaccатт аgссааааgс 600
 тacаggаgат саатгаагаа tcttсаатса ааgтааactа ctgттссagс асатgсатса 660
 tggтсagтаа gtttсагааа ааgасатсса ссгааgаctt аааgttagтg ggcатсtttg 720
 аааgтаатct tgtсаасатс gаgсagсtгг сttgtgggга ссagасаааа ааgгаатggt 780
 гсгаааттgt taggсгсacc тaccааааgс атctttгсct ttattгсааа gатаааgсag 840
 attсctсtag тасааgtggg гаасаааата асgtggаааа gаgсtгtсct гacаgсссac 900
 тсactаатгс gtатgасgаа сгсagтgасг accасааааg аattссctct ататаагааg 960

015908

gcattcattc ccatttgaag gacacagaaa aatttgctac attgtttcac aaacttcaaa 1020
 tattattcat ttatttgtca gctttcaaac tctttgtttc ttgtttgttg attgagaata 1080
 tttaaaacaa tggcttcctc tatgctctct tccgctacta tggttgcctc tccggctcag 1140
 gccactatgg tcgctccttt caacggactt aagtcctccg ctgccttccc agccaccgcg 1200
 aaggctaaca acgacattac ttccatcaca agcaacggcg gaagagttaa ctgcatgcag 1260
 gccatg gac aac aac cca aac atc aac gaa tgc att cca tac aac tgc 1308
 Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys
 1 5 10
 ttg agt aac cca gaa gtt gaa gta ctt ggt gga gaa cgc att gaa acc 1356
 Leu Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr
 15 20 25 30
 ggt tac act ccc atc gac atc tcc ttg tcc ttg aca cag ttt ctg ctc 1404
 Gly Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu
 35 40 45
 agc gag ttc gtg cca ggt gct ggg ttc gtt ctc gga cta gtt gac atc 1452
 Ser Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile
 50 55 60
 atc tgg ggt atc ttt ggt cca tct caa tgg gat gca ttc ctg gtg caa 1500
 Ile Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln
 65 70 75
 att gag cag ttg atc aac cag agg atc gaa gag ttc gcc agg aac cag 1548
 Ile Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln
 80 85 90
 gcc atc tct agg ttg gaa gga ttg agc aat ctc tac caa atc tat gca 1596
 Ala Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala
 95 100 105 110
 gag agc ttc aga gag tgg gaa gcc gat cct act aac cca gct ctc cgc 1644
 Glu Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg
 115 120 125
 gag gaa atg cgt att caa ttc aac gac atg aac agc gcc ttg acc aca 1692
 Glu Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr
 130 135 140
 gct atc cca ttg ttc gca gtc cag aac tac caa gtt cct ctc ttg tcc 1740
 Ala Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser
 145 150 155
 gtg tac gtt caa gca gct aat ctt cac ctc agc gtg ctt cga gac gtt 1788
 Val Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val
 160 165 170
 agc gtg ttt ggg caa agg tgg gga ttc gat gct gca acc atc aat agc 1836
 Ser Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser
 175 180 185 190
 cgt tac aac gac ctt act agg ctg att gga aac tac acc gac cac gct 1884
 Arg Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp His Ala
 195 200 205

015908

ggt cgt tgg tac aac act ggc ttg gag cgt gtc tgg ggt cct gat tct	1932
Val Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser	
210 215 220	
aga gat tgg att aga tac aac cag ttc agg aga gaa ttg acc ctc aca	1980
Arg Asp Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr	
225 230 235	
ggt ttg gac att gtg tct ctc ttc ccg aac tat gac tcc aga acc tac	2028
Val Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr Tyr	
240 245 250	
cct atc cgt aca gtg tcc caa ctt acc aga gaa atc tat act aac cca	2076
Pro Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro	
255 260 265 270	
ggt ctt gag aac ttc gac ggt agc ttc cgt ggt tct gcc caa ggt atc	2124
Val Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile	
275 280 285	
gaa ggc tcc atc agg agc cca cac ttg atg gac atc ttg aac agc ata	2172
Glu Gly Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile	
290 295 300	
act atc tac acc gat gct cac aga gga gag tat tac tgg tct gga cac	2220
Thr Ile Tyr Thr Asp Ala His Arg Gly Glu Tyr Tyr Trp Ser Gly His	
305 310 315	
cag atc atg gcc tct cca gtt gga ttc agc ggg ccc gag ttt acc ttt	2268
Gln Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe	
320 325 330	
cct ctc tat gga act atg gga aac gcc gct cca caa caa cgt atc gtt	2316
Pro Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val	
335 340 345 350	
gct caa cta ggt cag ggt gtc tac aga acc ttg tct tcc acc ttg tac	2364
Ala Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr	
355 360 365	
aga aga ccc ttc aat atc ggt atc aac aac cag caa ctt tcc gtt ctt	2412
Arg Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu	
370 375 380	
gac gga aca gag ttc gcc tat gga acc tct tct aac ttg cca tcc gct	2460
Asp Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala	
385 390 395	
ggt tac aga aag agc gga acc gtt gat tcc ttg gac gaa atc cca cca	2508
Val Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro	
400 405 410	
cag aac aac aat gtg cca ccc agg caa gga ttc tcc cac agg ttg agc	2556
Gln Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser	
415 420 425 430	
cac gtg tcc atg ttc cgt tcc gga ttc agc aac agt tcc gtg agc atc	2604
His Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile	
435 440 445	

015908

atc aga gct cct atg ttc tct tgg ata cac cgt agt gct gag ttc aac	2652
Ile Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn	
450 455 460	
aac atc att gca tcc gac agc att act caa ata ccc ttg gtg aaa gca	2700
Asn Ile Ile Ala Ser Asp Ser Ile Thr Gln Ile Pro Leu Val Lys Ala	
465 470 475	
cat aca ctt cag tca ggt act act gtt gtc aga ggt cca ggg ttt aca	2748
His Thr Leu Gln Ser Gly Thr Val Val Arg Gly Pro Gly Phe Thr	
480 485 490	
gga gga gac att ctt cgt cgc aca agt gga gga ccc ttt gct tac act	2796
Gly Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr	
495 500 505 510	
att gtt aac atc aat ggc caa ttg ccc caa agg tat cgt gca aga atc	2844
Ile Val Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg Ala Arg Ile	
515 520 525	
cgc tat gcc tct act aca aat ctc agg atc tac gtg act gtt gca ggt	2892
Arg Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly	
530 535 540	
gaa agg atc ttt gct ggt cag ttc aac aag act atg gat acc ggt gac	2940
Glu Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly Asp	
545 550 555	
cct ttg aca ttc caa tct ttt agc tac gca act atc aac aca gct ttt	2988
Pro Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe	
560 565 570	
aca ttc cca atg agc cag agt agc ttc aca gta ggt gct gac act ttc	3036
Thr Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe	
575 580 585 590	
agc tca ggg aat gaa gtt tac atc gac agg ttt gaa ttg att cca gtt	3084
Ser Ser Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu Ile Pro Val	
595 600 605	
act gca acc ctc gag gct gag tac aac ctt gag aga gcc cag aag gct	3132
Thr Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala	
610 615 620	
gtg aac gcc ctc ttt acc tcc acc aat cag ctt ggc ttg aaa act aac	3180
Val Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn	
625 630 635	
gtt act gac tat cac att gac caa gtg tcc aac ttg gtc acc tac ctt	3228
Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu	
640 645 650	
agc gat gag ttc tgc ctc gac gag aag cgt gaa ctc tcc gag aaa gtt	3276
Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val	
655 660 665 670	
aaa cac gcc aag cgt ctc agc gac gag agg aat ctc ttg caa gac tcc	3324
Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser	
675 680 685	
aac ttc aaa gac atc aac agg cag cca gaa cgt ggt tgg ggt gga agc	3372

015908

Asn	Phe	Lys	Asp	Ile	Asn	Arg	Gln	Pro	Glu	Arg	Gly	Trp	Gly	Gly	Ser		
			690					695					700				
acc	ggg	atc	acc	atc	caa	gga	ggc	gac	gat	gtg	ttc	aag	gag	aac	tac		3420
Thr	Gly	Ile	Thr	Ile	Gln	Gly	Gly	Asp	Asp	Val	Phe	Lys	Glu	Asn	Tyr		
		705					710					715					
gtc	acc	ctc	tcc	gga	act	ttc	gac	gag	tgc	tac	cct	acc	tac	ttg	tac		3468
Val	Thr	Leu	Ser	Gly	Thr	Phe	Asp	Glu	Cys	Tyr	Pro	Thr	Tyr	Leu	Tyr		
	720						725					730					
cag	aag	atc	gat	gag	tcc	aaa	ctc	aaa	gcc	ttc	acc	agg	tat	caa	ctt		3516
Gln	Lys	Ile	Asp	Glu	Ser	Lys	Leu	Lys	Ala	Phe	Thr	Arg	Tyr	Gln	Leu		
	735				740					745					750		
aga	ggc	tac	atc	gaa	gac	agc	caa	gac	ctt	gaa	atc	tac	tcg	atc	agg		3564
Arg	Gly	Tyr	Ile	Glu	Asp	Ser	Gln	Asp	Leu	Glu	Ile	Tyr	Ser	Ile	Arg		
				755					760						765		
tac	aat	gcc	aag	cac	gag	acc	gtg	aat	gtc	cca	ggt	act	ggt	tcc	ctc		3612
Tyr	Asn	Ala	Lys	His	Glu	Thr	Val	Asn	Val	Pro	Gly	Thr	Gly	Ser	Leu		
			770					775						780			
tgg	cca	ctt	tct	gcc	caa	tct	ccc	att	ggg	aag	tgt	gga	gag	cct	aac		3660
Trp	Pro	Leu	Ser	Ala	Gln	Ser	Pro	Ile	Gly	Lys	Cys	Gly	Glu	Pro	Asn		
		785					790						795				
aga	tgc	gct	cca	cac	ctt	gag	tgg	aat	cct	gac	ttg	gac	tgc	tcc	tgc		3708
Arg	Cys	Ala	Pro	His	Leu	Glu	Trp	Asn	Pro	Asp	Leu	Asp	Cys	Ser	Cys		
	800					805						810					
agg	gat	ggc	gag	aag	tgt	gcc	cac	cat	tct	cat	cac	ttc	tcc	ttg	gac		3756
Arg	Asp	Gly	Glu	Lys	Cys	Ala	His	His	Ser	His	His	Phe	Ser	Leu	Asp		
	815				820					825					830		
atc	gat	gtg	gga	tgt	act	gac	ctg	aat	gag	gac	ctc	gga	gtc	tgg	gtc		3804
Ile	Asp	Val	Gly	Cys	Thr	Asp	Leu	Asn	Glu	Asp	Leu	Gly	Val	Trp	Val		
				835					840					845			
atc	ttc	aag	atc	aag	acc	caa	gac	gga	cac	gca	aga	ctt	ggc	aac	ctt		3852
Ile	Phe	Lys	Ile	Lys	Thr	Gln	Asp	Gly	His	Ala	Arg	Leu	Gly	Asn	Leu		
			850					855						860			
gag	ttt	ctc	gaa	gag	aaa	cca	ttg	gtc	ggt	gaa	gct	ctc	gct	cgt	gtg		3900
Glu	Phe	Leu	Glu	Glu	Lys	Pro	Leu	Val	Gly	Glu	Ala	Leu	Ala	Arg	Val		
		865					870							875			
aag	aga	gca	gag	aag	aag	tgg	agg	gac	aaa	cgt	gag	aaa	ctc	gaa	tgg		3948
Lys	Arg	Ala	Glu	Lys	Lys	Trp	Arg	Asp	Lys	Arg	Glu	Lys	Leu	Glu	Trp		
	880					885						890					
gaa	act	aac	atc	gtt	tac	aag	gag	gcc	aaa	gag	tcc	gtg	gat	gct	ttg		3996
Glu	Thr	Asn	Ile	Val	Tyr	Lys	Glu	Ala	Lys	Glu	Ser	Val	Asp	Ala	Leu		
	895				900					905					910		
ttc	gtg	aac	tcc	caa	tat	gat	cag	ttg	caa	gcc	gac	acc	aac	atc	gcc		4044
Phe	Val	Asn	Ser	Gln	Tyr	Asp	Gln	Leu	Gln	Ala	Asp	Thr	Asn	Ile	Ala		
				915					920						925		
atg	atc	cac	gcc	gca	gac	aaa	cgt	gtg	cac	agc	att	cgt	gag	gct	tac		4092
Met	Ile	His	Ala	Ala	Asp	Lys	Arg	Val	His	Ser	Ile	Arg	Glu	Ala	Tyr		

015908

930	935	940	
ttg cct gag	ttg tcc gtg atc cct ggt gtg aac gct gcc atc ttc gag		4140
Leu Pro Glu	Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu		
945	950	955	
gaa ctt gag	gga cgt atc ttt acc gca ttc tcc ttg tac gat gcc aga		4188
Glu Leu Glu	Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg		
960	965	970	
aac gtc atc	aag aac ggt gac ttc aac aat ggc ctc agc tgc tgg aat		4236
Asn Val Ile	Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn		
975	980	985	990
gtg aaa ggt	cat gtg gac gtg gag gaa cag aac aat cag cgt tcc gtc		4284
Val Lys Gly	His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val		
995	1000	1005	
ctg gtt gtg	cct gag tgg gaa gct gaa gtg tcc caa gag gtt aga		4329
Leu Val Val	Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg		
1010	1015	1020	
gtc tgt cca	ggt aga ggc tac att ctc cgt gtg acc gct tac aag		4374
Val Cys Pro	Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys		
1025	1030	1035	
gag gga tac	ggt gag ggt tgc gtg acc atc cac gag atc gag aac		4419
Glu Gly Tyr	Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn		
1040	1045	1050	
aac acc gac	gag ctt aag ttc tcc aac tgc gtc gag gaa gaa atc		4464
Asn Thr Asp	Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile		
1055	1060	1065	
tat ccc aac	aac acc gtt act tgc aac gac tac act gtg aat cag		4509
Tyr Pro Asn	Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln		
1070	1075	1080	
gaa gag tac	gga ggt gcc tac act agc cgt aac aga ggt tac aac		4554
Glu Glu Tyr	Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn		
1085	1090	1095	
gaa gct cct	tcc gtt cct gct gac tat gcc tcc gtg tac gag gag		4599
Glu Ala Pro	Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu		
1100	1105	1110	
aaa tcc tac	aca gat ggc aga cgt gag aac cct tgc gag ttc aac		4644
Lys Ser Tyr	Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn		
1115	1120	1125	
aga ggt tac	agg gac tac aca cca ctt cca gtt ggc tat gtt acc		4689
Arg Gly Tyr	Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr		
1130	1135	1140	
aag gag ctt	gag tac ttt cct gag acc gac aaa gtg tgg atc gag		4734
Lys Glu Leu	Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu		
1145	1150	1155	
atc ggt gaa	acc gag gga acc ttc atc gtg gac agc gtg gag ctt		4779
Ile Gly Glu	Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu		
1160	1165	1170	

015908

ctc ttg atg gag gaa taa tgagatccac gatatacctgc aggaattggc 4827
 Leu Leu Met Glu Glu
 1175

cggccagctt tcgttcgtat catcggtttc gacaacgttc gtcaagttca atgcatcagt 4887
 ttcattgctc acacaccaga atcctactga gtttgagtat tatggcattg ggaaaactgt 4947
 ttttcttgta ccatttggtg tgcttgtaat ttactgtggt ttttattcgg ttttcgctat 5007
 cgaactgtga aatggaaatg gatggagaag agttaatgaa tgatatggtc cttttgttca 5067
 ttctcaaatt aatattatgt gttttttctc ttatttggtg tgtggtgaa ttgaaattat 5127
 aagagatatg caaacatgtt gttttgagta aaaatgtgtc aaatcgtggc ctctaagac 5187
 cgaagtaat atgaggagta aaacacttgt agttgtacca ttatgcttat tcaactaggca 5247
 acaaataat tttcagacct agaaaagctg caaatgttac tgaatacaag tatgtcctct 5307
 tgtgttttag acatttatgg actttccttt atgtaatttt ccagaatcct tgcagattc 5367
 taatcattgc tttataatta tagttatact catggatttg tagttgagta tgaaaatatt 5427
 ttttaatgca ttttatgact tgccaattga ttgacaacat gcatcaatcg acc 5480

<210> 6
 <211> 1176
 <212> Белок
 <213> Искусственная последовательность

<220>
 <223> Синтетический конструктор

<400> 6

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

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

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

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

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

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

015908

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

015908

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

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

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

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

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

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

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

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

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

Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro Gly Phe Thr Gly Gly
 485 490 495

Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr Ile Val
 500 505 510

Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg Ala Arg Ile Arg Tyr
 515 520 525

Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu Arg
 530 535 540

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

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

Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe Ser Ser

015908

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

015908

Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe
 835 840 845

Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe
 850 855 860

Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg
 865 870 875 880

Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu Thr
 885 890 895

Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val
 900 905 910

Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile
 915 920 925

His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu Pro
 930 935 940

Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu Leu
 945 950 955 960

Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val
 965 970 975

Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val Lys
 980 985 990

Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu Val
 995 1000 1005

Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys
 1010 1015 1020

Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly
 1025 1030 1035

Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr
 1040 1045 1050

Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr Pro
 1055 1060 1065

015908

Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu Glu
 1070 1075 1080

Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu Ala
 1085 1090 1095

Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys Ser
 1100 1105 1110

Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg Gly
 1115 1120 1125

Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys Glu
 1130 1135 1140

Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile Gly
 1145 1150 1155

Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu Leu
 1160 1165 1170

Met Glu Glu
 1175

<210> 7
 <211> 4990
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Экспрессионная кассета, кодирующая аминокислотную последовательность Cry1A.105

<220>
 <221> промотор
 <222> (1)..(614)
 <223> Промотор CAMV.e35s

<220>
 <221> 5'-нетранслируемый регион
 <222> (650)..(710)

<220>
 <221> Интрон
 <222> (727)..(1206)
 <223> Интрон рисового гена актина RACT1

<220>
 <221> Кодирующая последовательность
 <222> (1216)..(4752)
 <223> Кодирующая последовательность Cry1A.105

<220>
 <221> Терминатор
 <222> (4781)..(4990)
 <223> Сигналы терминации трансляции и полиаденилирования Та.Hsp17

 <400> 7
 ggtccgatgt gagacttttc aacaaagggt aatatccgga aacctcctcg gattccattg 60
 cccagctatc tgtcacttta ttgtgaagat agtggaaaag gaaggtggct cctacaaatg 120
 ccatcattgc gataaaggaa aggccatcgt tgaagatgcc tctgccgaca gtgggtcccaa 180
 agatggaccc ccaccacga ggagcatcgt ggaaaaagaa gacgttccaa ccacgtcttc 240
 aaagcaagtg gattgatgtg atgggccgat gtgagacttt tcaacaaagg gtaatatccg 300
 gaaacctcct cggattccat tgcccagcta tctgtcactt tattgtgaag atagtggaaa 360
 aggaaggtgg ctctacaaa tgccatcatt gcgataaagg aaaggccatc gttgaagatg 420
 cctctgccga cagtgggtccc aaagatggac cccaccccac gaggagcatc gtggaaaaag 480
 aagacgttcc aaccacgtct tcaaagcaag tggattgatg tgatatctcc actgacgtaa 540
 gggatgacgc acaatcccac tctcttcgc aagaccttc ctctatataa ggaagttcat 600
 ttcatttga gaggacacgc tgacaagctg actctagcag atcctctaga accatcttcc 660
 acacactcaa gccacactat tggagaacac acagggacaa cacaccataa gatccaaggg 720
 aggcctccgc cgccgccggt aaccaccccg cccctctcct ctttctttct ccgttttttt 780
 ttccgtctcg gtctcgatct ttggccttgg tagtttgggt gggcgagagg cggcttcgtg 840
 cgcgcccaaga tcggtgocg gggggggcgg gatctocgcg ctggggctct cgccggcgtg 900
 gatccggccc ggatctocg gggaaatggg ctctcggatg tagatctcg atccgccggt 960
 gttgggggag atgatggggg gtttaaaatt tccgccgtgc taaacaagat caggaagagg 1020
 ggaaaaggc actatggttt atattttat atatttctgc tgcttcgtca ggcttagatg 1080
 tgctagatct ttctttcttc ttttgtggg tagaatttga atccctcagc attgttcatc 1140
 ggtagttttt cttttcatga tttgtgacaa atgcagcctc gtgcggagct tttttgtagg 1200

 tagaagtgat caacc atg gac aac aac cca aac atc aac gag tgc atc ccg 1251
 Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro
 1 5 10

 tac aac tgc ctc agc aac cct gag gtc gag gtg ctc ggc ggt gag cgc 1299
 Tyr Asn Cys Leu Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg
 15 20 25

 atc gag acc ggt tac acc ccc atc gac atc tcc ctc tcc ctc acg cag 1347
 Ile Glu Thr Gly Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln
 30 35 40

 ttc ctg ctc agc gag ttc gtg cca ggc gct ggc ttc gtc ctg ggc ctc 1395
 Phe Leu Leu Ser Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu

015908

45	50	55	60	
gtg gac atc atc tgg ggc atc ttt ggc ccc tcc cag tgg gac gcc ttc Val Asp Ile Ile Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe 65 70 75				1443
ctg gtg caa atc gag cag ctc atc aac cag agg atc gag gag ttc gcc Leu Val Gln Ile Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala 80 85 90				1491
agg aac cag gcc atc agc cgc ctg gag ggc ctc agc aac ctc tac caa Arg Asn Gln Ala Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln 95 100 105				1539
atc tac gct gag agc ttc cgc gag tgg gag gcc gac ccc act aac cca Ile Tyr Ala Glu Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro 110 115 120				1587
gct ctc cgc gag gag atg cgc atc cag ttc aac gac atg aac agc gcc Ala Leu Arg Glu Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala 125 130 135 140				1635
ctg acc acc gcc atc cca ctc ttc gcc gtc cag aac tac caa gtc ccg Leu Thr Thr Ala Ile Pro Leu Phe Ala Val Gln Asn Tyr Gln Val Pro 145 150 155				1683
ctc ctg tcc gtg tac gtc cag gcc gcc aac ctg cac ctc agc gtg ctg Leu Leu Ser Val Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu 160 165 170				1731
agg gac gtc agc gtg ttt ggc cag agg tgg ggc ttc gac gcc gcc acc Arg Asp Val Ser Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr 175 180 185				1779
atc aac agc cgc tac aac gac ctc acc agg ctg atc ggc aac tac acc Ile Asn Ser Arg Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr 190 195 200				1827
gac cac gct gtc cgc tgg tac aac act ggc ctg gag cgc gtc tgg ggc Asp His Ala Val Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly 205 210 215 220				1875
cct gat tct aga gac tgg att cgc tac aac cag ttc agg cgc gag ctg Pro Asp Ser Arg Asp Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu 225 230 235				1923
acc ctc acc gtc ctg gac att gtg tcc ctc ttc ccg aac tac gac tcc Thr Leu Thr Val Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser 240 245 250				1971
cgc acc tac ccg atc cgc acc gtg tcc caa ctg acc cgc gaa atc tac Arg Thr Tyr Pro Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr 255 260 265				2019
acc aac ccc gtc ctg gag aac ttc gac ggt agc ttc agg ggc agc gcc Thr Asn Pro Val Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala 270 275 280				2067
cag ggc atc gag ggc tcc atc agg agc cca cac ctg atg gac atc ctc Gln Gly Ile Glu Gly Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu 285 290 295 300				2115

015908

aac agc atc act atc tac acc gat gcc cac cgc ggc gag tac tac tgg	2163
Asn Ser Ile Thr Ile Tyr Thr Asp Ala His Arg Gly Glu Tyr Tyr Trp	
305 310 315	
tcc ggc cac cag atc atg gcc tcc ccg gtc ggc ttc agc ggc ccc gag	2211
Ser Gly His Gln Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu	
320 325 330	
ttt acc ttt cct ctc tac ggc acg atg ggc aac gcc gct cca caa caa	2259
Phe Thr Phe Pro Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln	
335 340 345	
cgc atc gtc gct cag ctg ggc cag ggc gtc tac cgc acc ctg agc tcc	2307
Arg Ile Val Ala Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser	
350 355 360	
acc ctg tac cgc agg ccc ttc aac atc ggt atc aac aac cag cag ctg	2355
Thr Leu Tyr Arg Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu	
365 370 375 380	
tcc gtc ctg gat ggc act gag ttc gcc tac ggc acc tcc tcc aac ctg	2403
Ser Val Leu Asp Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu	
385 390 395	
ccc tcc gct gtc tac cgc aag agc ggc acg gtg gat tcc ctg gac gag	2451
Pro Ser Ala Val Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu	
400 405 410	
atc cca cca cag aac aac aat gtg ccc ccc agg cag ggt ttt tcc cac	2499
Ile Pro Pro Gln Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His	
415 420 425	
agg ctc agc cac gtg tcc atg ttc cgc tcc ggc ttc agc aac tcg tcc	2547
Arg Leu Ser His Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser	
430 435 440	
gtg agc atc atc aga gct cct atg ttc tct tgg ata cac cgt agt gct	2595
Val Ser Ile Ile Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala	
445 450 455 460	
gag ttc aac aac atc att gca tcc gac agc att act caa ata ccc ttg	2643
Glu Phe Asn Asn Ile Ile Ala Ser Asp Ser Ile Thr Gln Ile Pro Leu	
465 470 475	
gtg aaa gca cat aca ctt cag tca ggt act act gtt gtc aga ggt cca	2691
Val Lys Ala His Thr Leu Gln Ser Gly Thr Thr Val Val Arg Gly Pro	
480 485 490	
ggg ttt aca gga gga gac att ctt cgt cgc aca agt gga gga ccc ttt	2739
Gly Phe Thr Gly Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe	
495 500 505	
gct tac act att gtt aac atc aat ggc caa ttg ccc caa agg tat cgt	2787
Ala Tyr Thr Ile Val Asn Ile Asn Gly Gln Leu Pro Gln Arg Tyr Arg	
510 515 520	
gca aga atc cgc tat gcc tct act aca aat ctc agg atc tac gtg act	2835
Ala Arg Ile Arg Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr	
525 530 535 540	

015908

ggt gca ggt gaa agg atc ttt gct ggt cag ttc aac aag act atg gat	2883
Val Ala Gly Glu Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp	
545 550 555	
acc ggt gac cct ttg aca ttc caa tct ttt agc tac gca act atc aac	2931
Thr Gly Asp Pro Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn	
560 565 570	
aca gct ttt aca ttc cca atg agc cag agt agc ttc aca gta ggt gct	2979
Thr Ala Phe Thr Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala	
575 580 585	
gac act ttc agc tca ggg aat gaa gtt tac atc gac agg ttt gaa ttg	3027
Asp Thr Phe Ser Ser Gly Asn Glu Val Tyr Ile Asp Arg Phe Glu Leu	
590 595 600	
att cca gtt act gca acc ctc gag gct gag tac aac ctt gag aga gcc	3075
Ile Pro Val Thr Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala	
605 610 615 620	
cag aag gct gtg aac gcc ctc ttt acc tcc acc aat cag ctt ggc ttg	3123
Gln Lys Ala Val Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu	
625 630 635	
aaa act aac gtt act gac tat cac att gac caa gtg tcc aac ttg gtc	3171
Lys Thr Asn Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val	
640 645 650	
acc tac ctt agc gat gag ttc tgc ctc gac gag aag cgt gaa ctc tcc	3219
Thr Tyr Leu Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser	
655 660 665	
gag aaa gtt aaa cac gcc aag cgt ctc agc gac gag agg aat ctc ttg	3267
Glu Lys Val Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu	
670 675 680	
caa gac tcc aac ttc aaa gac atc aac agg cag cca gaa cgt ggt tgg	3315
Gln Asp Ser Asn Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp	
685 690 695 700	
ggt gga agc acc ggg atc acc atc caa gga ggc gac gat gtg ttc aag	3363
Gly Gly Ser Thr Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys	
705 710 715	
gag aac tac gtc acc ctc tcc gga act ttc gac gag tgc tac cct acc	3411
Glu Asn Tyr Val Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr	
720 725 730	
tac ttg tac cag aag atc gat gag tcc aaa ctc aaa gcc ttc acc agg	3459
Tyr Leu Tyr Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg	
735 740 745	
tat caa ctt aga ggc tac atc gaa gac agc caa gac ctt gaa atc tac	3507
Tyr Gln Leu Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr	
750 755 760	
tcg atc agg tac aat gcc aag cac gag acc gtg aat gtc cca ggt act	3555
Ser Ile Arg Tyr Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr	
765 770 775 780	
ggt tcc ctc tgg cca ctt tct gcc caa tct ccc att ggg aag tgt gga	3603

015908

Gly	Ser	Leu	Trp	Pro	Leu	Ser	Ala	Gln	Ser	Pro	Ile	Gly	Lys	Cys	Gly		
				785					790					795			
gag	cct	aac	aga	tgc	gct	cca	cac	ctt	gag	tgg	aat	cct	gac	ttg	gac		3651
Glu	Pro	Asn	Arg	Cys	Ala	Pro	His	Leu	Glu	Trp	Asn	Pro	Asp	Leu	Asp		
			800					805					810				
tgc	tcc	tgc	agg	gat	ggc	gag	aag	tgt	gcc	cac	cat	tct	cat	cac	ttc		3699
Cys	Ser	Cys	Arg	Asp	Gly	Glu	Lys	Cys	Ala	His	His	Ser	His	His	Phe		
		815					820					825					
tcc	ttg	gac	atc	gat	gtg	gga	tgt	act	gac	ctg	aat	gag	gac	ctc	gga		3747
Ser	Leu	Asp	Ile	Asp	Val	Gly	Cys	Thr	Asp	Leu	Asn	Glu	Asp	Leu	Gly		
	830					835					840						
gtc	tgg	gtc	atc	ttc	aag	atc	aag	acc	caa	gac	gga	cac	gca	aga	ctt		3795
Val	Trp	Val	Ile	Phe	Lys	Ile	Lys	Thr	Gln	Asp	Gly	His	Ala	Arg	Leu		
845					850				855					860			
ggc	aac	ctt	gag	ttt	ctc	gaa	gag	aaa	cca	ttg	gtc	ggt	gaa	gct	ctc		3843
Gly	Asn	Leu	Glu	Phe	Leu	Glu	Glu	Lys	Pro	Leu	Val	Gly	Glu	Ala	Leu		
				865				870					875				
gct	cgt	gtg	aag	aga	gca	gag	aag	aag	tgg	agg	gac	aaa	cgt	gag	aaa		3891
Ala	Arg	Val	Lys	Arg	Ala	Glu	Lys	Lys	Trp	Arg	Asp	Lys	Arg	Glu	Lys		
			880					885					890				
ctc	gaa	tgg	gaa	act	aac	atc	gtt	tac	aag	gag	gcc	aaa	gag	tcc	gtg		3939
Leu	Glu	Trp	Glu	Thr	Asn	Ile	Val	Tyr	Lys	Glu	Ala	Lys	Glu	Ser	Val		
		895					900					905					
gat	gct	ttg	ttc	gtg	aac	tcc	caa	tat	gat	cag	ttg	caa	gcc	gac	acc		3987
Asp	Ala	Leu	Phe	Val	Asn	Ser	Gln	Tyr	Asp	Gln	Leu	Gln	Ala	Asp	Thr		
	910					915					920						
aac	atc	gcc	atg	atc	cac	gcc	gca	gac	aaa	cgt	gtg	cac	agc	att	cgt		4035
Asn	Ile	Ala	Met	Ile	His	Ala	Ala	Asp	Lys	Arg	Val	His	Ser	Ile	Arg		
925					930				935					940			
gag	gct	tac	ttg	cct	gag	ttg	tcc	gtg	atc	cct	ggt	gtg	aac	gct	gcc		4083
Glu	Ala	Tyr	Leu	Pro	Glu	Leu	Ser	Val	Ile	Pro	Gly	Val	Asn	Ala	Ala		
				945					950					955			
atc	ttc	gag	gaa	ctt	gag	gga	cgt	atc	ttt	acc	gca	ttc	tcc	ttg	tac		4131
Ile	Phe	Glu	Glu	Leu	Glu	Gly	Arg	Ile	Phe	Thr	Ala	Phe	Ser	Leu	Tyr		
			960					965					970				
gat	gcc	aga	aac	gtc	atc	aag	aac	ggt	gac	ttc	aac	aat	ggc	ctc	agc		4179
Asp	Ala	Arg	Asn	Val	Ile	Lys	Asn	Gly	Asp	Phe	Asn	Asn	Gly	Leu	Ser		
		975					980					985					
tgc	tgg	aat	gtg	aaa	ggt	cat	gtg	gac	gtg	gag	gaa	cag	aac	aat	cag		4227
Cys	Trp	Asn	Val	Lys	Gly	His	Val	Asp	Val	Glu	Glu	Gln	Asn	Asn	Gln		
	990					995					1000						
cgt	tcc	gtc	ctg	gtt	gtg	cct	gag	tgg	gaa	gct	gaa	gtg	tcc	caa			4272
Arg	Ser	Val	Leu	Val	Val	Pro	Glu	Trp	Glu	Ala	Glu	Val	Ser	Gln			
1005					1010					1015							
gag	gtt	aga	gtc	tgt	cca	ggt	aga	ggc	tac	att	ctc	cgt	gtg	acc			4317
Glu	Val	Arg	Val	Cys	Pro	Gly	Arg	Gly	Tyr	Ile	Leu	Arg	Val	Thr			

015908

1020		1025		1030		
gct tac aag gag gga	tac ggt gag ggt tgc	gtg acc atc cac gag				4362
Ala Tyr Lys Glu Gly	Tyr Gly Glu Gly Cys	Val Thr Ile His Glu				
1035	1040	1045				
atc gag aac aac acc	gac gag ctt aag ttc	tcc aac tgc gtc gag				4407
Ile Glu Asn Asn Thr	Asp Glu Leu Lys Phe	Ser Asn Cys Val Glu				
1050	1055	1060				
gaa gaa atc tat ccc	aac aac acc gtt act	tgc aac gac tac act				4452
Glu Glu Ile Tyr Pro	Asn Asn Thr Val Thr	Cys Asn Asp Tyr Thr				
1065	1070	1075				
gtg aat cag gaa gag	tac gga ggt gcc tac	act agc cgt aac aga				4497
Val Asn Gln Glu Glu	Tyr Gly Gly Ala Tyr	Thr Ser Arg Asn Arg				
1080	1085	1090				
ggt tac aac gaa gct	cct tcc gtt cct gct	gac tat gcc tcc gtg				4542
Gly Tyr Asn Glu Ala	Pro Ser Val Pro Ala	Asp Tyr Ala Ser Val				
1095	1100	1105				
tac gag gag aaa tcc	tac aca gat ggc aga	cgt gag aac cct tgc				4587
Tyr Glu Glu Lys Ser	Tyr Thr Asp Gly Arg	Arg Glu Asn Pro Cys				
1110	1115	1120				
gag ttc aac aga ggt	tac agg gac tac aca	cca ctt cca gtt ggc				4632
Glu Phe Asn Arg Gly	Tyr Arg Asp Tyr Thr	Pro Leu Pro Val Gly				
1125	1130	1135				
tat gtt acc aag gag	ctt gag tac ttt cct	gag acc gac aaa gtg				4677
Tyr Val Thr Lys Glu	Leu Glu Tyr Phe Pro	Glu Thr Asp Lys Val				
1140	1145	1150				
tgg atc gag atc ggt	gaa acc gag gga acc	ttc atc gtg gac agc				4722
Trp Ile Glu Ile Gly	Glu Thr Glu Gly Thr	Phe Ile Val Asp Ser				
1155	1160	1165				
gtg gag ctt ctc ttg	atg gag gaa taa tga	gatctatcga ttctagaagg				4772
Val Glu Leu Leu Leu	Met Glu Glu					
1170	1175					
cctgaattct gcatgcgttt	ggacgtatgc tcattcaggt	tggagccaat ttggttgatg				4832
tgtgtgcgag ttcttgcgag	tctgatgaga catctctgta	ttgtgtttct ttccccagtg				4892
ttttctgtac ttgtgtaatc	ggctaatcgc caacagattc	ggcgatgaat aaatgagaaa				4952
taaattgttc tgattttgag	tgcaaaaaaa aaggaatt					4990
<210>	8					
<211>	1177					
<212>	Белок					
<213>	Искусственная последовательность					
<220>						
<223>	Синтетический конструктор					
<400>	8					

015908

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 His 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 Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val
 225 230 235 240

Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr Tyr Pro

015908

Gly Asp Ile Leu Arg Arg Thr Ser Gly Gly Pro Phe Ala Tyr Thr Ile
500 505 510

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

Tyr Ala Ser Thr Thr Asn Leu Arg Ile Tyr Val Thr Val Ala Gly Glu
530 535 540

Arg Ile Phe Ala Gly Gln Phe Asn Lys Thr Met Asp Thr Gly Asp Pro
545 550 555 560

Leu Thr Phe Gln Ser Phe Ser Tyr Ala Thr Ile Asn Thr Ala Phe Thr
565 570 575

Phe Pro Met Ser Gln Ser Ser Phe Thr Val Gly Ala Asp Thr Phe Ser
580 585 590

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

Ala Thr Leu Glu Ala Glu Tyr Asn Leu Glu Arg Ala Gln Lys Ala Val
610 615 620

Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asn Val
625 630 635 640

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Thr Tyr Leu Ser
645 650 655

Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
660 665 670

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Ser Asn
675 680 685

Phe Lys Asp Ile Asn Arg Gln Pro Glu Arg Gly Trp Gly Gly Ser Thr
690 695 700

Gly Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
705 710 715 720

Thr Leu Ser Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
725 730 735

015908

Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg
740 745 750

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Ser Ile Arg Tyr
755 760 765

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
770 775 780

Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg
785 790 795 800

Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg
805 810 815

Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile
820 825 830

Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile
835 840 845

Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu
850 855 860

Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys
865 870 875 880

Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu
885 890 895

Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe
900 905 910

Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met
915 920 925

Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu
930 935 940

Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu
945 950 955 960

Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn
965 970 975

Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val
 980 985 990

Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu
 995 1000 1005

Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val
 1010 1015 1020

Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu
 1025 1030 1035

Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn
 1040 1045 1050

Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr
 1055 1060 1065

Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu
 1070 1075 1080

Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu
 1085 1090 1095

Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys
 1100 1105 1110

Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg
 1115 1120 1125

Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys
 1130 1135 1140

Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile
 1145 1150 1155

Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu
 1160 1165 1170

Leu Met Glu Glu
 1175

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Выделенный полинуклеотид, кодирующий инсектицидный белок *B. thuringiensis* Cry1A.105, включающий аминокислоты 10-600 последовательности SEQ ID NO:2.

2. Выделенный полинуклеотид по п.1, выбранный из группы, состоящей из нуклеотидных последовательностей SEQ ID NO:1 и SEQ ID NO:3.

3. Применение выделенного полинуклеотида по п.2 для экспрессии инсектицидного белка, определенного в п.1, в сельскохозяйственных культурах.

4. Применение по п.3, где сельскохозяйственная культура выбрана из группы, состоящей из однодольных и двудольных растений.

5. Применение по п.4, где однодольное растение выбрано из группы, состоящей из кукурузы, пшеницы, овса, риса, сорго, гречихи, ржи, овсяницы, тимофеевки луговой, костера, ежи сборной, августиновой травы, свиноры пальчатого, полевицы и ячменя.

6. Применение по п.4, где двудольное растение выбрано из группы, состоящей из люцерны, спаржи, фасоли, кофе, ежевики, черники, канолы, моркови, цветной капусты, сельдерея, нута, хлопчатника, вигны китайской, клюквы, огурца, баклажана, винограда, салата-латука, льна, дыни, горчицы, орехоплодных

деревьев, бамии, окры, гороха, арахиса, картофеля, сои, тыквы, крупноплодной тыквы, земляники, сахарной свеклы, подсолнуха, батата, табака, томата, репы, плодовых деревьев, в частности яблони, абрикоса, персика, груши, сливы, вишни, цитрусовых, в частности лимона, апельсина, и других овощных и ягодных культур.

7. Выделенный и очищенный инсектицидный белок *B. thuringiensis* Cry1A.105, включающий аминокислоты 10-600 последовательности SEQ ID NO:2.

8. Композиция, включающая инсектицидно-эффективное количество индивидуального очищенного белка по п.7.

9. Экспрессионная кассета для экспрессии инсектицидного белка *B. thuringiensis* Cry1A.105, имеющего аминокислотную последовательность SEQ ID NO:2, в клетке-хозяине, содержащая оперативно связанные функциональную в данной клетке-хозяине промоторную последовательность и полинуклеотид, кодирующий вышеуказанный инсектицидный белок.

10. Экспрессионная кассета по п.9, где клетка-хозяин выбрана из группы, состоящей из клеток бактерий, грибов, млекопитающих и растений.

11. Экспрессионная кассета по п.10, где

(a) указанная бактериальная клетка выбрана из группы, состоящей из клеток бактерий родов *Bacillus*, *Enterobacteriaceae*, *Pseudomonas*, *Clostridium*, *Rhizobium* и *Agrobacterium*;

(b) указанная клетка выбрана из группы, состоящей из клеток однодольных и двудольных растений, причем двудольное растение выбрано из группы, состоящей из люцерны, спаржи, фасоли, кофе, ежевики, черники, канолы, моркови, цветной капусты, сельдерея, нута, хлопчатника, вигны китайской, клюквы, огурца, баклажана, винограда, салата-латука, льна, дыни, горчицы, орехоплодных деревьев, бамии, окры, гороха, арахиса, картофеля, сои, тыквы, крупноплодной тыквы, земляники, сахарной свеклы, подсолнуха, батата, табака, томата, репы, плодовых деревьев, в частности яблони, абрикоса, персика, груши, сливы, вишни, цитрусовых, в частности лимона, апельсина, и других овощных и ягодных культур, а однодольное растение выбрано из группы, состоящей из кукурузы, пшеницы, овса, риса, сорго, гречихи, ржи, овсяницы, тимофеевки луговой, костера, ежи сборной, августиновой травы, свинороя пальчатого, полевицы и ячменя.

12. Экспрессионная кассета по п.9, где клетка-хозяин представляет собой клетку растения и указанная экспрессионная кассета дополнительно содержит полинуклеотидную последовательность, выбранную из группы, состоящей из энхансерной последовательности экспрессии, нетранслируемой лидерной последовательности, интронной последовательности, последовательности, кодирующей хлоропласт-направляющий пептид, а также сигналов терминации транскрипции и полиаденилирования.

13. Экспрессионная кассета по п.12, в которой полинуклеотид, кодирующий инсектицидный белок, выбран из группы, состоящей из последовательности SEQ ID NO:5 и SEQ ID NO:7.

14. Вектор, содержащий экспрессионную кассету по любому из пп.9-13.

15. Трансгенное растение или растительная клетка, устойчивые к заражению чешуекрылыми насекомыми, включающие полинуклеотидную последовательность, кодирующую инсектицидный белок *B. thuringiensis* Cry1A.105, включающий аминокислоты 10-600 последовательности SEQ ID NO:2, который сохраняет инсектицидную активность будучи экспрессированным в указанном растении или растительной клетке.

16. Трансгенное растение или растительная клетка по п.15, где трансгенное растение выбрано из группы, состоящей из однодольных и двудольных растений, где двудольное растение выбрано из группы, состоящей из люцерны, спаржи, фасоли, кофе, ежевики, черники, канолы, моркови, цветной капусты, сельдерея, нута, хлопчатника, вигны китайской, клюквы, огурца, баклажана, винограда, салата-латука, льна, дыни, горчицы, орехоплодных деревьев, бамии, окры, гороха, арахиса, картофеля, сои, тыквы, крупноплодной тыквы, земляники, сахарной свеклы, подсолнуха, батата, табака, томата, репы, плодовых деревьев, в частности яблони, абрикоса, персика, груши, сливы, вишни, цитрусовых, в частности лимона, апельсина, и других овощных и ягодных культур, а однодольное растение выбраны из группы, состоящей из кукурузы, пшеницы, овса, риса, сорго, гречихи, ржи, овсяницы, тимофеевки луговой, костера, ежи сборной, августиновой травы, свинороя пальчатого, полевицы и ячменя.

17. Трансгенное растение или растительная клетка по п.16, устойчивые к заражению чешуекрылыми насекомыми, выбранными из группы, состоящей из листоверток, гусениц озимой совки, "походных червей", точильщиков, мешочницы поденкоподобной и кормовых гусениц.

18. Трансгенное растение или растительная клетка по п.16, устойчивая к заражению чешуекрылыми насекомыми, выбранными из группы, состоящей из совки травяной, мотылька стеблевого кукурузного, гусеницы совки хлопковой американской, огневки кукурузной юго-западной и совки-ипсилон.

19. Потомство трансгенного растения по п.16 или потомство трансгенного растения, полученного из растительной клетки по п.16, содержащие полинуклеотид, определенный в п.15.

20. Семена трансгенного растения по п.16 или семена трансгенного растения, полученного из растительной клетки по п.16, содержащие полинуклеотид, определенный в п.15.

21. Способ борьбы с чешуекрылыми насекомыми-вредителями растений, включающий обеспечение поедания насекомыми одного или более типов растительных клеток, трансформированных нуклеиновой

кислотой, содержащей функциональный в растениях промотор, оперативно связанный с полинуклеотидом, кодирующим инсектицидный белок *B. thuringiensis* Cry1A.105, включающий аминокислоты 10-600 последовательности SEQ ID NO:2, обладающий инсектицидной активностью по отношению к чешуекрылам насекомым-вредителям.

22. Способ обнаружения в биологической пробе полинуклеотида, кодирующего инсектицидный белок *B. thuringiensis* Cry1A.105, включающий аминокислоты в положении от приблизительно 10 до приблизительно 600 последовательности SEQ ID NO:2, предусматривающий контактирование биологической пробы с полинуклеотидным зондом, гибридизующимся с указанным полинуклеотидом в жестких гибридизационных условиях, при этом детекция связывания является диагностичной в отношении присутствия указанного олигонуклеотида в указанной пробе.

23. Способ обнаружения в биологической пробе инсектицидного белка *B. thuringiensis* Cry1A.105, включающего аминокислоты в положении от приблизительно 10 до приблизительно 600 последовательности SEQ ID NO:2, включающий контактирование биологической пробы с антителом, специфично связывающимся с данным белком, с последующей детекцией этого связывания, при этом детекция связывания является диагностичной в отношении присутствия указанного белка в указанной пробе.

24. Выделенный полинуклеотид, кодирующий инсектицидный белок *B. thuringiensis* Cry1A.105, включающий аминокислоты от приблизительно 1 до приблизительно 612 последовательности SEQ ID NO:2.

25. Выделенный полинуклеотид, кодирующий инсектицидный белок *B. thuringiensis* Cry1A.105, включающий аминокислоты от приблизительно 1 до приблизительно 610 последовательности SEQ ID NO:2.

26. Выделенный полинуклеотид по п.24, нуклеотидная последовательность которого по меньшей мере на приблизительно 90% идентична последовательности SEQ ID NO:1.

27. Выделенный полинуклеотид по п.24, нуклеотидная последовательность которого по меньшей мере на приблизительно 90% идентична последовательности SEQ ID NO:3.

28. Гибридный инсектицидный белок, содержащий аминокислотный сегмент, включающий от приблизительно 500 до приблизительно 600 аминокислот, составляющих непрерывную последовательность в составе сегмента, состоящего из аминокислот в положении от приблизительно 10 до приблизительно 600 последовательности SEQ ID NO:2 инсектицидного белка *B. thuringiensis* Cry1A.105.

29. Композиция, включающая растительные, бактериальные и/или грибковые клетки, в которых накоплено инсектицидно-эффективное количество инсектицидного белка *B. thuringiensis* Cry1A.105, включающего аминокислоты в положении от приблизительно 10 до приблизительно 600 последовательности SEQ ID NO:2, представленная в форме коллоида, эмульсии или порошка, причем указанная композиция предназначена для использования в качестве покрытия семян или в качестве приманки.

30. Композиция по п.29, где белок Cry1A.105 в растительных, бактериальных или грибковых клетках накоплен в количестве от приблизительно 0,5 до приблизительно 200 частей на миллион (PPM).

31. Композиция по п.30, в которой белок Cry1A.105 в растительных, бактериальных или грибковых клетках накоплен в количестве от приблизительно 0,5 до приблизительно 20 PPM.

32. Композиция по п.30, содержащая растительные клетки, включающие указанный инсектицидный белок.

33. Композиция по п.32, в которой инсектицидно-эффективное количество указанного инсектицидного белка можно обнаружить при помощи зонда, последовательность которого представляет собой сегмент SEQ ID NO:1 с нуклеотидами в положении 1401-1420 или сегмент SEQ ID NO:1 с нуклеотидами в положении 1821-1840, либо последовательность, комплементарную этим сегментам.

34. Композиция по п.32, в которой инсектицидно-эффективное количество указанного белка при поедании чешуекрылыми насекомыми-вредителями достаточно для контроля численности последних; при этом чешуекрылые насекомые-вредители выбраны из группы, состоящей из родов *Anticarsia*, *Pseudoplusia*, *Rachiplusia*, *Heliothis*, *Helicoverpa*, *Spodoptera*, *Epinotia* и *Armigera*.

35. Способ защиты засеянной в поле сельскохозяйственной культуры от заражения чешуекрылыми насекомыми, включающий выращивание трансгенной сельскохозяйственной культуры, содержащей инсектицидно-эффективное количество инсектицидного белка *B. thuringiensis* Cry1A.105 в качестве инсектицидного агента, и попадающей в пищу чешуекрылых насекомых, таким образом снижая выживаемость этих насекомых на поверхности трансгенного растения; при этом указанный инсектицидный белок Cry1A.105 включает аминокислоты в положении от приблизительно 10 до приблизительно 600 последовательности SEQ ID NO:2.

36. Способ по п.35, в котором указанная культура дополнительно содержит инсектицидно-эффективное количество дополнительного инсектицидного агента, токсичного по отношению к тем же видам насекомых, что и инсектицидный белок Cry1A.105, и где указанный дополнительный инсектицидный агент выбран из группы, состоящей из токсинов бактерий родов *Bacillus*, *Xenorhabdus* и *Photorhabdus* toxin, и двухцепочечной РНК, специфично супрессирующей один или более генов, необходимых для выживания указанных насекомых.

37. Способ по п.36, где указанный дополнительный инсектицидный агент представляет собой ток-

син бактерий родов *Bacillus*, выбранный из группы белков, состоящей из токсинов Cry1, Cry2, Cry5, Cry9 и белка VIP.

38. Способ по любому из пп.35-37, в результате применения которого урожай указанной сельскохозяйственной культуры превышает таковой, собранный с изогенной культуры, лишенной указанного(ых) инсектицидного(ых) агента(ов), определенных в пп.35-37.

39. Способ задержки развития устойчивости к инсектицидам у чешуекрылых насекомых, включающий предоставление насекомым первого инсектицидного белка Cry1A.105, и по меньшей мере второго инсектицидного белка, отличного от первого, где указанный первый инсектицидный белок Cry1A.105 включает аминокислотную последовательность из аминокислот в положении от приблизительно 10 до приблизительно 600 последовательности SEQ ID NO:2 и второй инсектицидный белок выбран из группы, состоящей из токсинов Cry1, Cry2, Cry5, Cry9 и белка VIP.

