

(19)



**Евразийское
патентное
ведомство**

(21) **202091560** (13) **A1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ**

(43) Дата публикации заявки
2020.10.02

(51) Int. Cl. **C12N 15/82 (2006.01)**
C07K 14/415 (2006.01)

(22) Дата подачи заявки
2019.01.09

(54) **ГЕНЫ РЕЗИСТЕНТНОСТИ К ОСЫПАНИЮ СЕМЯН И ИХ МУТАЦИИ**

(31) **62/615,409; 62/732,397**

(72) Изобретатель:

(32) **2018.01.09; 2018.09.17**

Гокал Грегори Ф.В. (US)

(33) **US**

(74) Представитель:

(86) **PCT/US2019/012938**

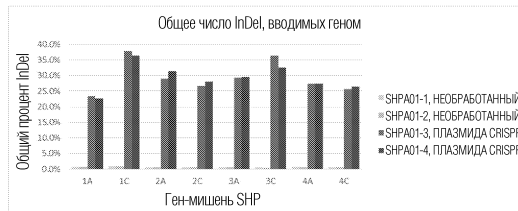
Медведев В.Н. (RU)

(87) **WO 2019/140009 2019.07.18**

(71) Заявитель:

**САЙБАС ЮЭс ЛЛС (US); САЙБАС
ЮРОП, Б.В. (NL)**

(57) В настоящем изобретении раскрываются гены резистентности к осыпанию семян (SHP) и растения и/или клетки растения, несущие одну или более мутаций в гене резистентности к осыпанию семян; а также способы выращивания и применения таких растений. В некоторых вариантах осуществления изобретения растение или клетка растения являются резистентными к растрескиванию плодов до сбора урожая.



A1

202091560

202091560

A1

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-563710EA/011

ГЕНЫ РЕЗИСТЕНТНОСТИ К ОСЫПАНИЮ СЕМЯН И ИХ МУТАЦИИ

Перекрестная ссылка на родственные заявки

[0001] В настоящей заявке испрашивается преимущество Предварительной заявки на патент США No. 62/615409, поданной 9 января 2018, и Предварительной заявки на патент США No. 62/732397, поданной 17 сентября 2018, каждая из которых в полном объеме включена в настоящее описание посредством ссылки.

Ссылка на список последовательностей, поданный в виде текстового файла ASCII

[0002] Содержание нижеследующего рассмотренного документа, поданного в виде текстового файла ASCII, включено в настоящее описание посредством ссылки в полном объеме: электронная форма (CRF) списка последовательностей (имя файла: 165072000140SEQLIST.txt, дата регистрации: 9 января 2019 г., размер: 90 кБ).

Область изобретения

[0003] В настоящем изобретении раскрываются композиции и способы, относящиеся к новым генам растений и генным продуктам, а также к растениям, имеющим одну или более генных мутаций. В частности, настоящее изобретение относится к генам резистентности к осыпанию семян (SHP) и к растениям и/или клеткам растений, несущим одну или более мутаций в гене резистентности к осыпанию семян, а также к способам выращивания и использования таких растений. В некоторых вариантах осуществления изобретения, растение или клетка растения являются резистентными к растрескиванию плодов до сбора урожая.

Предпосылки создания изобретения

[0004] Растрескивание стручков канолы до сбора урожая представляет собой серьезную агрономическую проблему, которая связана со значительной потерей урожая, а также с переносом культуры на последующий вегетационный период. В соответствии с этим, существует потребность в разработке усовершенствованных способов снижения или предотвращения растрескивания стручков с семенами до сбора урожая, а также в выращивании растений с улучшенными свойствами, а именно, с повышенной резистентностью или с пониженной чувствительностью к растрескиванию плодов до сбора урожая.

Краткое описание сущности изобретения

[0005] Настоящее изобретение основано, по меньшей мере частично, на обнаружении того факта, что растения Brassica имеют восемь генов резистентности к осыпанию семян; и что введение мутаций в один или более таких генов может снижать растрескивание плодов до сбора урожая сельскохозяйственных культур, таких как культура Brassica.

[0006] Используемый здесь термин «ген резистентности к осыпанию семян (SHP)» означает ген, имеющий последовательность, представленную последовательностями

SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A, SHP4C *Brassica napus*, как описано в настоящей заявке, или, в некоторых вариантах осуществления изобретения, их гомологи, варианты или мутанты. Термин «гомолог гена резистентности к осыпанию семян» или любой его вариант означает ген резистентности к осыпанию семян или его продукт, обнаруженные у растений других видов, которые обладают такой же или, по существу, такой же биологической функцией, как и гены *Brassica*, и описанные здесь продукты этого гена, где последовательности нуклеиновой кислоты кодирующей области или полипептидные последовательности (продукта гена SHP) считаются «идентичными» или по меньшей мере на 50%, или по меньшей мере на 60%, или по меньшей мере на 70%, или по меньшей мере на 75%, или по меньшей мере на 80%, или по меньшей мере на 85%, или по меньшей мере на 90%, или по меньшей мере на 92%, или по меньшей мере на 95%, или по меньшей мере на 96%, или по меньшей мере на 97%, или по меньшей мере на 98% или по меньшей мере на 99% аналогичными (этот термин также упоминается как «процент идентичности» или «по существу идентичен») одной или более описанным здесь последовательностям SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A, SHP4C.

[0007] В своем первом аспекте, настоящее изобретение относится к способу предотвращения или уменьшения растрескивания плодов растения до сбора урожая, где указанный способ включает введение мутации по меньшей мере в один эндогенный ген резистентности к осыпанию семян в клетках указанного растения. В некоторых вариантах осуществления изобретения, способ включает (1) введение в клетки растения олигонуклеосоединения для репарации генов с получением клеток растений с мутантным геном SHP; и (2) регенерацию нетрансгенного растения, имеющего мутированный ген SHP, из указанной выбранной клетки растения. В некоторых вариантах осуществления изобретения, способ включает (1) введение в клетки растения ДНК-разрезающего фермента, структура которого обеспечивает специфический одноцепочечный разрыв или разрезание гена SHP для продуцирования клеток растений с мутантным геном SHP; и (2) регенерацию нетрансгенного растения, имеющего мутированный ген SHP, из указанной выбранной клетки растения. В своем родственном варианте, настоящее изобретение относится к способу, включающему контактирование клетки с ДНК-разрезающим ферментом, структура которого обеспечивает специфический одноцепочечный разрыв или разрезание гена резистентности к осыпанию семян. В своем родственном аспекте, настоящее изобретение относится к способам введения мутации в ген SHP. В некоторых вариантах осуществления изобретения, описанные здесь способ или способы могут включать обработку клетки ДНК-разрезающим ферментом и GRON. В некоторых вариантах осуществления изобретения, способы включают обработку клетки ДНК-разрезающим ферментом и GRON, где указанный GRON модифицируют одной или более группами Су3, группами 3PS и 2'О-метильными группами. В некоторых вариантах осуществления изобретения, способ или способы могут включать обработку клетки ДНК-разрезающим ферментом без обработки клетки GRON. В некоторых вариантах

осуществления изобретения, которые включают обработку ДНК-разрезающим ферментом, ДНК-разрезающий фермент специфически нацелен на ген SHP. В некоторых вариантах осуществления изобретения, ДНК-разрезающий фермент представляет собой один или более ферментов, выбранных из CRISPR, которые включают, но не ограничиваются ими, Cas9, Cpf1 и их соответствующие гомологи, ортологи и/или паралоги; редактор оснований; TALEN; «цинковый палец»; мегануклеазу и ДНК-разрезающий антибиотик. В некоторых вариантах осуществления изобретения, ДНК-разрезающим ферментом может быть плаزمид (ДНК), РНК и/или белок. В некоторых вариантах осуществления изобретения, способы не включают контактирование растения или клетки растения с любым трансгеном. В некоторых вариантах любых рассматриваемых здесь аспектов и вариантов, растение или клетка растения являются нетрансгенными. В некоторых аспектах, мутация, изменение или модификация в гене SHP включает инсерцию или делецию. В некоторых вариантах осуществления изобретения, мутация, изменение или модификация представляет собой или включает нуклеотидную модификацию или замену. В некоторых вариантах осуществления этого способа, изменение, мутация или модификация вводят преждевременный стоп-кодон. В некоторых вариантах осуществления изобретения, изменение, мутация или модификация вводят мутацию со сдвигом рамки считывания. В некоторых вариантах описанных здесь композиций и способов, мутация по сравнению с геном SHP дикого типа представляет собой инсерцию или делецию нуклеотида +1, -1, -2 (InDel). В некоторых вариантах описанных здесь композиций и способов, мутация, по сравнению с геном SHP дикого типа, представляет собой инсерцию или делецию нуклеотида +1, -1, -2 (InDel), введенную путем целевой мутации. В некоторых вариантах описанных здесь способов, мутация, модификация или изменение в гене SHP приводят к уменьшению или устранению активности или экспрессии гена SHP. В некоторых вариантах описанных здесь способов, по меньшей мере один ген SHP; или по меньшей мере два гена SHP; или по меньшей мере три гена SHP; или по меньшей мере четыре гена SHP; или по меньшей мере пять генов SHP; или по меньшей мере шесть генов SHP; или по меньшей мере семь генов SHP; или восемь генов SHP являются модифицированными. В некоторых аспектах изобретения, мутация, изменение или модификация включают инсерцию или делецию. В некоторых вариантах осуществления изобретения, мутация, изменение или модификация включают нуклеотидную модификацию или замену. В некоторых вариантах этого способа, изменение, мутация или модификация вводит преждевременный стоп-кодон. В некоторых вариантах описанных здесь способов, мутация, модификация или изменение в гене SHP уменьшает или устраняет активность или экспрессию гена SHP. В некоторых вариантах осуществления изобретения, растение или клетка растения представляет собой растение Brassica. В некоторых вариантах осуществления изобретения, описанные растения или клетки растения получают описанными здесь способами.

[0008] В одном из своих аспектов, настоящее изобретение относится к выделенной нуклеиновой кислоте, имеющей раскрытую здесь последовательность SHP1A, SHP1C,

SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C или ее фрагмент. В некоторых вариантах осуществления изобретения, фрагмент одной или более из вышеупомянутых последовательностей гена SHP включает по меньшей мере 80%; или по меньшей мере 85%, или по меньшей мере 90%, или по меньшей мере 92%, или по меньшей мере 95%, или по меньшей мере 96%, или по меньшей мере 97%, или по меньшей мере 98% или по меньшей мере 99% от всей последовательности гена. В своем родственном аспекте, настоящее изобретение относится к выделенной аминокислотной последовательности, кодируемой раскрытой здесь последовательностью нуклеиновой кислоты SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C или ее фрагментом.

[0009] В другом своем аспекте, настоящее изобретение относится к растению или к клетке растения, имеющим по меньшей мере три, или по меньшей мере четыре, или по меньшей мере пять, или по меньшей мере семь или восемь генов резистентности к осыпанию семян, имеющих последовательность, отличающуюся от последовательности любого встречающегося в природе гена резистентности к осыпанию семян.

[0010] В одном своем аспекте, настоящее изобретение относится к растению или к клетке растения, имеющим по меньшей мере три, или по меньшей мере четыре, или по меньшей мере пять, или по меньшей мере семь или восемь эндогенных генов резистентности к осыпанию семян, имеющих последовательность, отличающуюся от последовательности любого встречающегося в природе гена резистентности к осыпанию семян.

[0011] В другом своем аспекте, настоящее изобретение относится к растению канолы или к клетке растения канолы, имеющим, по меньшей мере один, по меньшей мере два, или по меньшей мере три, или по меньшей мере четыре, или по меньшей мере пять, или по меньшей мере семь или восемь генов резистентности к осыпанию семян, имеющих последовательность, отличающуюся от последовательности любого встречающегося в природе гена резистентности к осыпанию семян.

[0012] В некоторых аспектах и вариантах осуществления изобретения, желательным, чтобы происходило растрескивание (но не преждевременное), и, таким образом, в некоторых вариантах осуществления изобретения, может сохраняться определенное количество активности продукта одного или более генов/локусов SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C. В соответствии с этим, в одном из своих вариантов, настоящее изобретение относится к растению или к клетке растения, имеющим от трех до семи генов SHP, имеющих последовательность, которая отличается от последовательности любого встречающегося в природе гена резистентности к осыпанию семян. В другом своем варианте, настоящее изобретение относится к растению или к клетке растения, имеющим от трех до шести генов SHP, имеющих последовательность, которая отличается от последовательности любого встречающегося в природе гена резистентности к осыпанию семян. В другом своем варианте, настоящее изобретение относится к растению или к клетке растения, имеющим от трех до пяти генов SHP, имеющих последовательность, которая отличается от последовательности любого

встречающегося в природе гена резистентности к осыпанию семян. В другом своем варианте, настоящее изобретение относится к растению или к клетке растения, имеющим от четырех до шести генов SHP, имеющих последовательность, которая отличается от последовательность любого встречающегося в природе гена резистентности к осыпанию семян. В другом своем варианте, настоящее изобретение относится к растению или к клетке растения, имеющим четыре или пять генов SHP, имеющих последовательность, которая отличается от последовательность любого встречающегося в природе гена резистентности к осыпанию семян. В другом своем варианте, настоящее изобретение относится к растению или к клетке растения, имеющим три или четыре гена SHP, имеющих последовательность, которая отличается от последовательность любого встречающегося в природе гена резистентности к осыпанию семян. В другом своем варианте, настоящее изобретение относится к растению или к клетке растения, имеющим три гена SHP, имеющих последовательность, которая отличается от последовательность любого встречающегося в природе гена резистентности к осыпанию семян. В другом своем варианте, настоящее изобретение относится к растению или к клетке растения, имеющим четыре гена SHP, имеющих последовательность, которая отличается от последовательность любого встречающегося в природе гена резистентности к осыпанию семян. В другом своем варианте, настоящее изобретение относится к растению или к клетке растения, имеющим пять генов SHP, имеющих последовательность, которая отличается от последовательность любого встречающегося в природе гена резистентности к осыпанию семян. В другом своем варианте, настоящее изобретение относится к растению или к клетке растения, имеющим шесть генов SHP, имеющих последовательность, которая отличается от последовательность любого встречающегося в природе гена резистентности к осыпанию семян. В другом своем варианте, настоящее изобретение относится к растению или к клетке растения, имеющим семь генов SHP, имеющих последовательность, которая отличается от последовательность любого встречающегося в природе гена резистентности к осыпанию семян.

[0013] В некоторых своих аспектах, настоящее изобретение относится к растению или к клетке растения, имеющим мутацию в гене SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C. В некоторых вариантах этого аспекта, ген SHP представляет собой эндогенный ген SHP.

[0014] В другом своем аспекте, настоящее изобретение относится к растению или к его части, включающим по меньшей мере одну мутацию по меньшей мере в одной, по меньшей мере в двух, по меньшей мере в трех, по меньшей мере в четырех, по меньшей мере в пяти, по меньшей мере в шести, по меньшей мере в семи или в восьми последовательностях нуклеиновых кислот, кодирующих гены резистентности к осыпанию семян (SHATTERPROOF (SHP)). В некоторых вариантах осуществления изобретения, последовательности нуклеиновой кислоты по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 98% или по меньшей мере на 99% идентичны последовательностям нуклеиновой кислоты, выбранным из группы последовательностей

SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 и SEQ ID NO: 8. В некоторых вариантах осуществления изобретения, последовательности нуклеиновых кислот выбраны из группы последовательностей SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 и SEQ ID NO: 8. В некоторых вариантах, которые могут быть объединены с любыми предшествующими вариантами, мутация представляет собой мутацию со сдвигом рамки считывания. В некоторых вариантах осуществления изобретения, мутация со сдвигом рамки считывания приводит к одной или более инсерциям или делециям нуклеотидов по сравнению с соответствующим эндогенным геном без мутации со сдвигом рамки считывания. В некоторых вариантах, которые могут быть объединены с любыми предшествующими вариантами, мутация со сдвигом рамки считывания приводит к введению преждевременного стоп-кодона. В некоторых вариантах, которые могут быть объединены с любыми предшествующими вариантами, мутация уменьшает или устраняет экспрессию гена SHP и/или полипептида SHP. В некоторых вариантах, которые могут быть объединены с любыми предшествующими вариантами, растение обладает пониженной чувствительностью к растрескиванию плодов до сбора урожая. В некоторых вариантах, которые могут быть объединены с любыми предшествующими вариантами, растение выбирают из группы растений видов *Brassica napus*, *Brassica rapa*, *Brassica oleracea*, *Brassica juncea*, *Brassica*, *Raphanus sativus*, *Pisum sativum*, *Phaseolus vulgaris*, *Lens culinaris*, *Glycine max* и *Fabaceae*.

[0015] В другом своем аспекте, настоящее изобретение относится к способу получения растения или его части по любому из предшествующих вариантов осуществления изобретения, где указанный способ включает стадии: а) введения мутаций в клетки растения, где мутации представляют собой по меньшей мере одну мутацию по меньшей мере в одной, по меньшей мере в двух, по меньшей мере в трех, по меньшей мере в четырех, по меньшей мере в пяти, по меньшей мере в шести, по меньшей мере в семи или в восьми последовательностях нуклеиновых кислот, кодирующих гены SHP; б) отбора клеток растения, содержащих мутации; и с) регенерации растения, имеющего мутации, где растение обладает пониженной восприимчивостью к растрескиванию плодов до сбора урожая. В некоторых вариантах осуществления изобретения, мутации вводят с использованием одного или более векторов, где векторы включают компоненты редактирования генов, выбранные из группы, состоящей из системы CRISPR/Cas9, TALEN, «цинкового пальца» и мегануклеазы, предназначенных для нацеливания на последовательность нуклеиновой кислоты, кодирующую ген SHP. В некоторых вариантах осуществления изобретения, мутации вводят с использованием системы GRON, предназначенной для нацеливания на последовательность нуклеиновой кислоты, кодирующую ген SHP. В некоторых вариантах осуществления изобретения, система GRON включает одну или более модификаций, выбранных из группы, состоящей из группы Су3, группы ZPS и 2'О-метильной группы. В некоторых вариантах осуществления, которые могут быть объединены с любыми предшествующими вариантами,

последовательности нуклеиновой кислоты по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 98% или по меньшей мере на 99% идентичны последовательностям нуклеиновой кислоты, выбранным из группы последовательностей SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 и SEQ ID NO: 8. В некоторых вариантах осуществления изобретения, которые могут быть объединены с любыми предшествующими вариантами, последовательности нуклеиновых кислот выбраны из группы последовательностей SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 и SEQ ID NO: 8. В некоторых вариантах, которые могут быть объединены с любыми предшествующими вариантами, мутация выбрана из группы, состоящей из мутации со сдвигом рамки считывания, мутации со сдвигом рамки считывания, которая дает одну или более инсерций или делеций нуклеотидов по сравнению с соответствующим эндогенным геном без мутации со сдвигом рамки считывания, и мутации со сдвигом рамки считывания, которая приводит к введению преждевременного стоп-кодона, и где такая мутация снижает или устраняет экспрессию гена SHP и/или полипептида SHP. В некоторых вариантах, которые могут быть объединены с любыми предшествующими вариантами, растение выбирают из группы растений видов *Brassica napus*, *Brassica rapa*, *Brassica oleracea*, *Brassica juncea*, *Brassica*, *Raphanus sativus*, *Pisum sativum*, *Phaseolus vulgaris*, *Lens culinaris*, *Glycine max* и *Fabaceae*.

[0016] В другом своем аспекте, настоящее изобретение относится к растению F₁, где растение F₁ представляет собой родительское растение согласно любому из предшествующих вариантов осуществления изобретения. В другом своем аспекте, настоящее изобретение относится к способу получения семян растений, включающему скрещивание растений согласно любому из предшествующих вариантов осуществления изобретения с другим растением и сбор семян этого растения. В другом своем аспекте, настоящее изобретение относится к способу выращивания растения согласно любому из предшествующих вариантов осуществления изобретения, где указанный способ включает отбор семян после скрещивания растения согласно любому из предшествующих вариантов осуществления изобретения с растением согласно любому из предшествующих вариантов осуществления изобретения и выращивание растения согласно любому из предшествующих вариантов осуществления изобретения. В другом своем аспекте, настоящее изобретение относится к растению, полученному путем выращивания семян согласно любому из предшествующих вариантов осуществления изобретения, где растение обладает всеми физиологическими и морфологическими свойствами растения согласно любому из предшествующих вариантов осуществления изобретения.

Краткое описание чертежей

[0017] Файл патента или заявки содержит по меньшей мере один чертеж, выполненный в цвете. Копии этого патента или публикации патентной заявки с цветным(и) чертежом(ами) будут предоставлены Ведомством по запросу и после уплаты необходимой пошлины.

[0018] На фиг. 1A-1B проиллюстрировано выравнивание множества частичных нуклеотидных последовательностей и выведенных аминокислотных последовательностей генов SHATTERPROOF (BnSHP) *Brassica napus* и генов SHP *Arabidopsis thaliana* (AtSHP1 и AtSHP2) с помощью CLUSTAL. На фиг. 1A проиллюстрировано выравнивание множества частичных нуклеотидных последовательностей BnSHP1A (SEQ ID NO: 30), BnSHP1C (SEQ ID NO: 31), BnSHP2A (SEQ ID NO: 23), BnSHP2C (SEQ ID NO: 22), BnSHP3A (SEQ ID NO: 28), BnSHP3C (SEQ ID NO: 29), BnSHP4A (SEQ ID NO: 25), BnSHP4C (SEQ ID NO: 26), AtSHP1 (SEQ ID NO: 24) и AtSHP2 (SEQ ID NO: 27), начиная со старт-кодона (нуклеотидные последовательности *B. napus*, полученные из гДНК линии BN2-SU) с помощью CLUSTAL. На фиг. 1B проиллюстрировано выравнивание множества выведенных аминокислотных последовательностей BnSHP1A (SEQ ID NO: 34), BnSHP1C (SEQ ID NO: 35), BnSHP2A (SEQ ID NO: 36), BnSHP2C (SEQ ID NO: 37), BnSHP3A (SEQ ID NO: 39), BnSHP3C (SEQ ID NO: 38), BnSHP4A (SEQ ID NO: 40), BnSHP4C (SEQ ID NO: 41), AtSHP1 (SEQ ID NO: 32) и AtSHP2 (SEQ ID NO: 33) с помощью CLUSTAL.

[0019] На фиг. 2A-2B проиллюстрирован анализ на экспрессию гена SHP путем секвенирования следующего поколения (NGS). На фиг. 2A проиллюстрированы стадии развития образцов плодов, взятых для анализа на экспрессию гена SHP: стадия развития 13=цветение, когда цветки раскрываются и самоопыляются; стадии развития 17-1, 17-2, 17-3, 17-4 и 17-5=плоды увеличиваются в размерах на стадии удлинения (Roeder and Yanofsky, 2006). На фиг. 2B показан процент общего числа считываний, идентифицированных для каждого гена SHP на каждой стадии развития плода после анализа с помощью ОТ-ПЦР и анализа с помощью NGS (стадии развития показаны слева направо для каждого гена SHP в порядке 13, 17-1, 17-2, 17-3, 17-4 и 17-5).

[0020] На фиг. 3 проиллюстрирован общий процент инсерций и делеций (InDel), идентифицированных с помощью NGS в каждом из генов SHP в побегах, регенерированных из контрольных необработанных протопластов (SHPA01-1 и SHPA01-2, более светлые серые столбики показаны слева для каждого гена SHP) и протопластов, обработанных плазмидой CRISPR/Cas9 (SHPA01-3 и SHPA01-4, более темные серые столбики показаны справа для каждого гена SHP). Процедуры проводили с двумя повторностями.

[0021] На фиг. 4 проиллюстрировано окрашивание флороглюцинолом слоев одревесневших клеток в плодах канолы (в стручках). На левой панели проиллюстрирован стручок (продольное изображение полного стручка). Стручки происходят от двух плодолистиков, которые образуют две доли, разделенные перегородкой. Стенки плодов представляют собой створки, содержащие семена, которые прикрепляются к перегородкам, образуя ложбинку. На верхней средней панели проиллюстрирован поперечный разрез створок стручка, содержащих семена, прикрепленные к перегородке. На нижней средней панели показан поперечный разрез створок стручка, содержащих семена, прикрепленные к перегородке, окрашенной флороглюцинолом, и показаны слои одревесневших клеток в области прикрепления створки стручка к перегородке и слой

одревесневших клеток эндокарпия-b. На верхней правой панели проиллюстрированы слои одревесневших клеток в поперечном сечении стручка дикого типа (оценка окрашивания флороглюцинолом=1). На нижней правой панели проиллюстрировано отсутствие слоя одревесневших клеток в поперечном сечении стручка от полной мутантной линии с нокаутом SHP (КО) (оценка окрашивания флороглюцинолом=5). Отсутствие слоя одревесневших клеток в области прикрепления кремнеземного створки стручка к перегородке ассоциируется с высоким уровнем резистентности к осыпанию плодов у масличных растений Brassicas.

[0022] На фиг. 5 проиллюстрирован тест на растрескивание стручков отобранных КО-растений (C_0) с различным количеством КО-генов BnSHP. Высушенные и невысушенные зрелые стручки были получены из растений дикого типа (0 КО) и растений с КО 2, 3, 5, 6 или 7 генов SHP (КО показаны слева направо для каждой частоты, показанной в следующем порядке: 0 КО (невысушенные), 0 КО (высушенные), 2 КО (высушенные), 3 КО (высушенные), 5 КО (высушенные), 6 КО (высушенные) и 7 КО (высушенные)).

[0023] На фиг. 6 проиллюстрирована корреляция между частотой события осыпания семян, как было определено с помощью TissueLyser (Гц), и частотой события осыпания семян, определенной с помощью Geno/Grinder (RPM). Значение $R=0,88$; значение $p=0,00016$.

[0024] На Фиг. 7 представлены данные о фенотипе и генотипе линий и контрольных линий (контроль) канолы с КО гена резистентности к осыпанию семян.

Подробное описание изобретения

[0025] Различные аспекты и варианты осуществления изобретения относятся к растению, имеющему одну или более мутаций SHP и/или комбинации таких мутаций, к способам получения такого растения и к способам снижения растрескивания плодов до сбора урожая.

[0026] Для специалиста в данной области очевидно, что раскрытие настоящего изобретения хорошо адаптировано для выполнения задач и достижения вышеупомянутых целей, а также присущих ему преимуществ. Приведенные здесь примеры представляют собой репрезентативные предпочтительные варианты осуществления изобретения, которые приводятся лишь в иллюстративных целях и не рассматриваются как ограничение объема изобретения.

[0027] Специалисту в данной области будет очевидно, что в настоящее изобретение могут быть внесены различные замены и модификации, не выходящие за рамки существа и объема изобретения.

[0028] Описанное здесь изобретение может быть осуществлено на практике в отсутствие какого-либо элемента или элементов, ограничения или ограничений, которые конкретно не раскрыты в настоящей заявке. Так, например, в каждом конкретном случае, любой из терминов «содержащий», «состоящий по существу из» и «состоящий из» может быть заменен любым из двух других терминов. Используемые здесь термины и

выражения употребляются в качестве описательных, а не ограничительных терминов, и при этом не предполагается, что при использовании таких терминов и выражений исключаются какие-либо эквиваленты представленных и описанных здесь признаков или их частей, но при этом предусматривается, что в объем заявленного изобретения могут входить различные модификации. Таким образом, следует отметить, что хотя настоящее изобретение было конкретно раскрыто путем представления предпочтительных вариантов осуществления изобретения и необязательных признаков, однако, специалистам в данной области очевидно, что могут быть введены модификации и варианты раскрытых здесь концепций, и что такие модификации и варианты не выходят за рамки объема данного раскрытия, определенного в прилагаемой формуле изобретения.

[0029] Таким образом, следует отметить, что, хотя настоящее изобретение было конкретно раскрыто путем представления предпочтительных вариантов осуществления изобретения и необязательных признаков, однако, специалистам в данной области очевидно, что могут быть введены модификации, усовершенствования и варианты раскрытого здесь изобретения, и что такие модификации, усовершенствования и варианты не выходят за рамки объема данного раскрытия. Описанные здесь материалы, методы и примеры представлены путем иллюстрации предпочтительных вариантов осуществления изобретения, и не должны рассматриваться как ограничение объема изобретения.

[0030] Настоящее изобретение было раскрыто в широком смысле и в общих чертах. Каждый из более узких видов и субродовых групп, входящих в общее раскрытие изобретения, также является частью такого раскрытия. Настоящее изобретение включает общее описание раскрытия с условием или отрицательным ограничением, удаляющим любой предмет из рода, независимо от того, был ли упомянут здесь исключенный материал или нет.

[0031] Кроме того, если признаки или аспекты раскрытия описаны на примере «групп Маркуша», то следует отметить, что настоящее изобретение также описано на примере любых отдельных членов или подгрупп членов группы Маркуша.

[0032] Все цитируемые здесь публикации, патентные заявки, патенты и другие документы включены в настоящее описание посредством ссылки в полном объеме, так, как если бы каждый из них был введен отдельно посредством ссылки. В случае каких-либо разночтений, следует отдать предпочтение определениям, используемым в настоящей заявке.

Растрескивание плодов перед сбором урожая

[0033] Плоды или стручки растений Brassica высвобождают свои семена в процессе, называемом растрескиванием плодов. Опадание семян (также называемое «осыпанием семян» или «осыпанием стручков») из зрелых стручков до или во время сбора урожая является универсальным феноменом для культур, у которых развиваются сухие растрескивающиеся плоды. Преждевременное осыпание семян приводит к снижению урожая семян, что представляет проблему для культур, которые выращиваются в основном для получения семян, таких как масличные растения Brassica, а в частности,

масличный рапс. Другая проблема, связанная с преждевременным осыпанием семян, заключается в увеличении роста растений-самосевов (сорняков) в последующем году сбора урожая.

[0034] Растрескивание стручков канолы до сбора урожая представляет собой серьезную агрономическую проблему, которая вызывает значительную потерю урожая, а также перенос культуры на последующий вегетационный период. В растениях канолы, осыпание стручков приводит к ежегодной потере урожая на 20% и может приводить к 50% потере при задержке сбора урожая и при неблагоприятных погодных условиях. Растрескивание семян происходит у сельскохозяйственных культур в период зрелости во время жаркого и ветреного лета, из-за воздействия других растений и во время валковой жатвы, осуществляемой с помощью уборочной техники (MacLeod, 1981; Child and Evans, 1989). В действительности, осыпание семян приводит к потере урожая, оцениваемой в 20-25 долларов на акр (ежегодная потеря урожая в США оценивается в 39 миллионов долларов), а затраты на валку добавляют еще 6 долларов на акр (8,7 миллионов долларов дополнительных затрат в год; по оценкам Барри Кольмана, Северная Ассоциация Фермеров по выращиванию канолы).

[0035] Используемая здесь культура Brassicaceae развивается из гинцея, состоящего из двух совмещенных плодолистиков, которые после оплодотворения растут и превращаются в стручки с двумя углублениями (створками), которые содержат развивающиеся семена (см. Фиг. 4, Пример 4). Стенки плода представляют собой створки, которые прикрепляются к перегородке (персистентной перегородке завязи) с образованием ложбинки, также называемой зоной растрескивания (DZ), расположенной по краям створки. DZ обычно состоит из тонкого слоя клеток паренхимы, который действует как разделительный слой при созревании плодов, если секретируются ферменты, разрушающие клеточную стенку, такие как целлюлазы и полигалатуроназы, и снижающие слипание клеток, а также приводящие к осыпанию стручков под действием внешних механических сил (Meakin & Roberts, 1990a, b). Отсутствие разделительного слоя в области прикрепления створки стручка к перегородке ассоциируется с высоким уровнем резистентности к осыпанию семян у масличных культур Brassicas (Kadkol et al., 1986; Meakin and Roberts, 1990a, b; Liljegren et al., 2000).

[0036] Используемый здесь термин «одревесневший клеточный» означает другой слой специализированных клеток вдоль краев створки, который, помимо разделительного слоя, способствует открытию плода (см. Фиг. 4). По достижении зрелости, слой одревесневших клеток в створке и в перегородке перестает удерживать неодревесневший разделительный слой по краям створки. Было высказано предположение, что жесткость клеточных стенок в процессе одревеснения пограничного слоя и внутреннего одревесневшего слоя створки эндокарпия b сообщает механическую силу, способствующую раскрытию плодов (Spence et al., 1996). Предполагается, что по мере высыхания плодов, дифференциальная усадка оставшихся тонкостенных клеток створок относительно жесткого одревесневшего края и слоев створки создает внутреннее

натяжение, что приводит к растрескиванию плода, характерному для осыпания семян.

[0037] Используемый здесь термин «гены резистентности к осыпанию семян (SHATTERPROOF (SHP))» означает члены семейства MADS-боксов, включающего факторы транскрипции, участвующие в дифференцировке DZ в развивающихся стручках Brassicaceae (Liljegrén et al., 2000). Исследования потери функции показали, что SHP способствует одревеснению клеток краев створки (то есть, одревесневшего клеточного слоя) у плодов *Arabidopsis*. Двойные мутанты *Arabidopsis shp1shp2* образуют нефункциональную DZ, которая полностью не дифференцирует ни слой клеток с одревесневшими клеточными стенками, ни разделительный слой клеток, и, как следствие, плоды не растрескиваются и не раскрываются в конце развития (Liljegrén et al., 2000). Гены, кодирующие SHP, экспрессируются на краях створки на ранних стадиях развития гинецея, где они активируют экспрессию факторов bHLH INDEHISCENT (IND), необходимых для развития разделительного слоя и одревесневшего слоя, а ALCATRAZ (ALC) необходим только для образования разделительного слоя (Roeder and Yanofsky, 2006).

[0038] Используемый здесь масличный рапс (синоним канола, рапс, *Brassica napus* L., подвид *Oleifera*; геномы AACС, $2n=4 \times = 38$), также являющийся членом семейства Brassicaceae, представляет собой аллополиплоидное растение, формирующееся посредством спонтанной гибридизации между турнепсом (*Brassica rapa* L.; геном AA, $2n=2 \times = 20$) и капустой (*Brassica oleracea* L.; геном CC, $2n=2 \times = 18$) (Chalhoub et al., 2014). Гомологи SHP1/2 *Arabidopsis* (а также другие функционально родственные факторы транскрипции IND и ALC) были обнаружены в каноле, а молекулярно-генетические исследования, проведенные ранее, указывали на наличие нескольких локусов количественных признаков (QTL), ассоциированных с осыпанием семян с эпистатическими отношениями между ними (Gururaj, 2009; Raman et al., 2014).

[0039] Используемый здесь термин «повышенная резистентность к осыпанию стручков», означает снижение степени осыпания семян зрелых (высушенных) плодов вследствие воздействия внешних механических сил в лаборатории и в полевых условиях. Лабораторные испытания имитируют процесс осыпания стручков, который происходит в естественных полевых условиях, и результаты обычно коррелируют с полевыми измерениями. Одна только полевая оценка резистентности к осыпанию семян может быть неточной из-за меняющихся погодных условий во время сбора урожая в разные сезоны и в различных регионах.

[0040] Используемые здесь анатомические признаки плодов ассоциируются с резистентностью к осыпанию стручков. Дифференциация одревесневших клеток края створки и разделительного клеточного слоя определяет степень осыпания семян. У *Arabidopsis*, потеря одревесневших клеток края створки, окрашиваемых флороглюцинолом, позитивно коррелирует с более высокой резистентностью стручков к механическому осыпанию. Флороглюцинол представляет собой распространенный краситель, используемый для окрашивания лигнина клеточной стенки в растительной

ткани. После окрашивания флороглюцинолом, одревесневшие клеточные стенки выглядят красно-фиолетовыми, а интенсивность окраски (цвета) позитивно коррелирует с уровнем отложения лигнина и дифференцировки клеток. Одревесневшие клетки края створки легко окрашиваются флороглюцинолом в поперечных срезах семян *Arabidopsis* и масличного рапса (OSR) дикого типа. Клетки разделительного слоя DZ не содержат лигнина, а поэтому они не окрашиваются флороглюцинолом. В плодах растений с двойным мутантом *Arabidopsis shp1shp2* не дифференцируются клетки края створки с одревесневшими клеточными стенками, а поэтому они не окрашиваются флороглюцинолом (Liljegren et al., 2000).

[0041] Используемый здесь тест на разрушение стручка означает лабораторный тест, в котором используется устройство для лизиса ткани для оценки резистентности к осыпанию зрелых, полностью высушенных стручков. Фенотип резистентности к осыпанию семян был определен по уровню разделения створок, обнаруживаемому при регулируемом встряхивании стручков. Для этого теста, отдельные стручки помещают в глубокий 96-луночный контейнер с поддоном и закрепляют на кронштейнах TissueLyser II (Qiagen, Germany). Образцы одного стручка помещают на TissueLyser на 30 секунд при частотах 22, 23, 24, 25, 26, 27, 28, 29 и 30 Гц. Четыре отдельных стручка на растение тестируют на каждой частоте. Фенотип оценивают по шкале от 1 до 5 (см., например, Пример 4). Неповрежденному стручку присваивается балл 1, частично разделенному стручка с соединительными створками присваивается балл 2, стручку с одной разделенной створкой присваивается балл 3, а стручкам, в которых обе створки разделены перегородкой, присваивается балл 4. Высокая корреляция была обнаружена между частотой встряхивания и показателями осыпания высушенных стручков и показателями окрашивания флороглюцинолом одревесневших слоев развивающихся стручков ($r=0,797$). Это продемонстрировало, что окрашивание одревесневшего слоя плодов и частота встряхивания в тесте на осыпание плодов могут быть использованы для эффективной оценки признака резистентности к осыпанию семян.

[0042] Осыпание стручков может быть также определено с помощью Geno/Grinder 2010 (SPEX Sample Prep, США). В этом случае, фенотип резистентности к осыпанию семян определяют по уровню разделения створок, обнаруживаемому при регулируемом встряхивании стручков. Для оценки разделения створок, 12-24 стручка помещают в глубокий 96-луночный контейнер с поддоном и закрепляют на кронштейнах Geno/Grinder 2010. Контейнеры, содержащие образцы стручков, центрифугировали в течение 20 секунд с различными скоростями вращения (например, при 720, 750, 780, 810, 840, 870, 900, 930, 960, 990, 1020, 1050, 1080 об/мин). По окончании этого цикла, контейнер вынимают из устройства, и каждому стручку присваивают баллы по уровню их осыпания в соответствии с таблицей баллов (см. Пример 4). Когда средний показатель осыпания при определенных скоростях вращения превышает 2,5 оборотов в минуту, то такая величина может рассматриваться как величина осыпания стручков для данной линии.

Гены резистентности к осыпанию семян (SHP)

[0043] В общих чертах, настоящее изобретение относится к растениям, имеющим мутации в генах резистентности к осыпанию семян (SHP). В некоторых вариантах осуществления изобретения, одна или более мутаций в одном или более генах SHP приводят к повышению резистентности/снижению чувствительности к растрескиванию плодов до сбора урожая.

[0044] В некоторых аспектах изобретения, растениями согласно изобретению являются растения *Brassica napus* L., ssp. *oleifera* (канола, масличный рапс). Растения канолы содержат восемь генов SHATTERPROOF (SHP), обозначенных BnSHP1A, BnSHP1C, BnSHP2A, BnSHP2C, BnSHP3A, BnSHP3C, BnSHP4A и BnSHP4C. В некоторых аспектах изобретения, растения согласно изобретению имеют по меньшей мере одну мутацию по меньшей мере в одном гене SHP.

[0045] В некоторых своих аспектах, настоящее изобретение относится к BnSHP1A. Нуклеотидная кодирующая последовательность BnSHP1A представлена в SEQ ID NO: 1. Настоящее изобретение также относится к гомологам и ортологам BnSHP1A. В некоторых вариантах осуществления изобретения, гомолог или ортолог BnSHP1A имеет кодирующую последовательность нуклеиновой кислоты, которая по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 91%, по меньшей мере на 92%, по меньшей мере на 93%, по меньшей мере на 94%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98% или по меньшей мере на 99% идентична SEQ ID NO: 1. В некоторых вариантах осуществления изобретения, последовательность нуклеиновой кислоты, кодирующая гомолог или ортолог BnSHP1A, может также иметь одну или более мутаций.

[0046] В некоторых своих аспектах, настоящее изобретение относится к BnSHP1C. Нуклеотидная кодирующая последовательность BnSHP1C представлена в SEQ ID NO: 2. Настоящее изобретение также относится к гомологам и ортологам BnSHP1C. В некоторых вариантах осуществления изобретения, гомолог или ортолог BnSHP1C имеет кодирующую последовательность нуклеиновой кислоты, которая по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 91%, по меньшей мере на 92%, по меньшей мере на 93%, по меньшей мере на 94%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98% или по меньшей мере на 99% идентична SEQ ID NO: 2. В некоторых вариантах осуществления изобретения, последовательность нуклеиновой кислоты, кодирующая гомолог или ортолог BnSHP1C, может также иметь одну или более мутаций.

[0047] В некоторых своих аспектах, настоящее изобретение относится к BnSHP2A. Нуклеотидная кодирующая последовательность BnSHP2A представлена в SEQ ID NO: 3. Настоящее изобретение также относится к гомологам и ортологам BnSHP2A. В некоторых

вариантах осуществления изобретения, гомолог или ортолог BnSHP2A имеет кодирующую последовательность нуклеиновой кислоты, которая по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 91%, по меньшей мере на 92%, по меньшей мере на 93%, по меньшей мере на 94%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98% или по меньшей мере на 99% идентична SEQ ID NO: 3. В некоторых вариантах осуществления изобретения, последовательность нуклеиновой кислоты, кодирующая гомолог или ортолог BnSHP2A, может также иметь одну или более мутаций.

[0048] В некоторых своих аспектах, настоящее изобретение относится к BnSHP2C. Нуклеотидная кодирующая последовательность BnSHP2C представлена в SEQ ID NO: 4. Настоящее изобретение также относится к гомологам и ортологам BnSHP2C. В некоторых вариантах осуществления изобретения, гомолог или ортолог BnSHP2C имеет кодирующую последовательность нуклеиновой кислоты, которая по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 91%, по меньшей мере на 92%, по меньшей мере на 93%, по меньшей мере на 94%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98% или по меньшей мере на 99% идентична SEQ ID NO: 4. В некоторых вариантах осуществления изобретения, последовательность нуклеиновой кислоты, кодирующая гомолог или ортолог BnSHP2C, может также иметь одну или более мутаций.

[0049] В некоторых своих аспектах, настоящее изобретение относится к BnSHP3A. Нуклеотидная кодирующая последовательность BnSHP3A представлена в SEQ ID NO: 5. Настоящее изобретение также относится к гомологам и ортологам BnSHP3A. В некоторых вариантах осуществления изобретения, гомолог или ортолог BnSHP3A имеет кодирующую последовательность нуклеиновой кислоты, которая по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 91%, по меньшей мере на 92%, по меньшей мере на 93%, по меньшей мере на 94%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98% или по меньшей мере на 99% идентична SEQ ID NO: 5. В некоторых вариантах осуществления изобретения, последовательность нуклеиновой кислоты, кодирующая гомолог или ортолог BnSHP3A, может также иметь одну или более мутаций.

[0050] В некоторых своих аспектах, настоящее изобретение относится к BnSHP3C. Нуклеотидная кодирующая последовательность BnSHP3C представлена в SEQ ID NO: 6. Настоящее изобретение также относится к гомологам и ортологам BnSHP3C. В некоторых вариантах осуществления изобретения, гомолог или ортолог BnSHP3C имеет

кодирующую последовательность нуклеиновой кислоты, которая по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 91%, по меньшей мере на 92%, по меньшей мере на 93%, по меньшей мере на 94%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98% или по меньшей мере на 99% идентична SEQ ID NO: 6. В некоторых вариантах осуществления изобретения, последовательность нуклеиновой кислоты, кодирующая гомолог или ортолог *BnSHP3C*, может также иметь одну или более мутаций.

[0051] В некоторых своих аспектах, настоящее изобретение относится к *BnSHP4A*. Нуклеотидная кодирующая последовательность *BnSHP4A* представлена в SEQ ID NO: 7. Настоящее изобретение также относится к гомологам и ортологам *BnSHP4A*. В некоторых вариантах осуществления изобретения, гомолог или ортолог *BnSHP4A* имеет кодирующую последовательность нуклеиновой кислоты, которая по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 91%, по меньшей мере на 92%, по меньшей мере на 93%, по меньшей мере на 94%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98% или по меньшей мере на 99% идентична SEQ ID NO: 7. В некоторых вариантах осуществления изобретения, последовательность нуклеиновой кислоты, кодирующая гомолог или ортолог *BnSHP4A*, может также иметь одну или более мутаций.

[0052] В некоторых своих аспектах, настоящее изобретение относится к *BnSHP4C*. Нуклеотидная кодирующая последовательность *BnSHP4C* представлена в SEQ ID NO: 8. Настоящее изобретение также относится к гомологам и ортологам *BnSHP4C*. В некоторых вариантах осуществления изобретения, гомолог или ортолог *BnSHP4C* имеет кодирующую последовательность нуклеиновой кислоты, которая по меньшей мере на 50%, по меньшей мере на 55%, по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 91%, по меньшей мере на 92%, по меньшей мере на 93%, по меньшей мере на 94%, по меньшей мере на 95%, по меньшей мере на 96%, по меньшей мере на 97%, по меньшей мере на 98% или по меньшей мере на 99% идентична SEQ ID NO: 8. В некоторых вариантах осуществления изобретения, последовательность нуклеиновой кислоты, кодирующая гомолог или ортолог *BnSHP4C*, может также иметь одну или более мутаций.

[0053] В некоторых аспектах изобретения, растения согласно изобретению имеют мутацию в *BnSHP1A*. В некоторых вариантах осуществления изобретения, эти растения могут также иметь мутации в одном или более генах SHP, выбранных из *BnSHP1C*, *BnSHP2A*, *BnSHP2C*, *BnSHP3A*, *BnSHP3C*, *BnSHP4A* и *BnSHP4C*.

[0054] В некоторых аспектах изобретения, растения согласно изобретению имеют

мутацию в *BnSHP1C*. В некоторых вариантах осуществления изобретения, эти растения могут также иметь мутации в одном или более генах SHP, выбранных из *BnSHP1A*, *BnSHP2A*, *BnSHP2C*, *BnSHP3A*, *BnSHP3C*, *BnSHP4A* и *BnSHP4C*.

[0055] В некоторых аспектах изобретения, растения согласно изобретению имеют мутацию в *BnSHP2A*. В некоторых вариантах осуществления изобретения, эти растения могут также иметь мутации в одном или более генах SHP, выбранных из *BnSHP1A*, *BnSHP1C*, *BnSHP2C*, *BnSHP3A*, *BnSHP3C*, *BnSHP4A* и *BnSHP4C*.

[0056] В некоторых аспектах изобретения, растения согласно изобретению имеют мутацию в *BnSHP2C*. В некоторых вариантах осуществления изобретения, эти растения могут также иметь мутации в одном или более генах SHP, выбранных из *BnSHP1A*, *BnSHP1C*, *BnSHP2A*, *BnSHP3A*, *BnSHP3C*, *BnSHP4A* и *BnSHP4C*.

[0057] В некоторых аспектах изобретения, растения согласно изобретению имеют мутацию в *BnSHP3A*. В некоторых вариантах осуществления изобретения, эти растения могут также иметь мутации в одном или более генах SHP, выбранных из *BnSHP1A*, *BnSHP1C*, *BnSHP2A*, *BnSHP2C*, *BnSHP3C*, *BnSHP4A* и *BnSHP4C*.

[0058] В некоторых аспектах изобретения, растения согласно изобретению имеют мутацию в *BnSHP3C*. В некоторых вариантах осуществления изобретения, эти растения могут также иметь мутации в одном или более генах SHP, выбранных из *BnSHP1A*, *BnSHP1C*, *BnSHP2A*, *BnSHP2C*, *BnSHP3A*, *BnSHP4A* и *BnSHP4C*.

[0059] В некоторых аспектах изобретения, растения согласно изобретению имеют мутацию в *BnSHP4A*. В некоторых вариантах осуществления изобретения, эти растения могут также иметь мутации в одном или более генах SHP, выбранных из *BnSHP1A*, *BnSHP1C*, *BnSHP2A*, *BnSHP2C*, *BnSHP3A*, *BnSHP3C* и *BnSHP4C*.

[0060] В некоторых аспектах изобретения, растения согласно изобретению имеют мутацию в *BnSHP4C*. В некоторых вариантах осуществления изобретения, эти растения могут также иметь мутации в одном или более генах SHP, выбранных из *BnSHP1A*, *BnSHP1C*, *BnSHP2A*, *BnSHP2C*, *BnSHP3A*, *BnSHP3C* и *BnSHP4A*.

[0061] В некоторых аспектах изобретения, растения согласно изобретению имеют мутацию по меньшей мере в одном, по меньшей мере в двух, по меньшей мере в трех, по меньшей мере в четырех, по меньшей мере в пяти, по меньшей мере в шести, по меньшей мере в семи или в восьми генах SHP. В некоторых аспектах изобретения, растения согласно изобретению имеют мутацию по меньшей мере в пяти, по меньшей мере в шести, по меньшей мере в семи или в восьми генах SHP.

[00621] В некоторых аспектах изобретения, мутация может представлять собой мутацию со сдвигом рамки считывания; мутацию со сдвигом рамки считывания, приводящую к одной или более инсерциям или делециям нуклеотидов по сравнению с соответствующим эндогенным геном без мутации со сдвигом рамки считывания, или мутацию со сдвигом рамки считывания, приводящую к встраиванию преждевременного стоп-кодона, где такая мутация снижает или устраняет экспрессию гена SHP и/или полипептида SHP.

Методы идентификации сходства последовательностей

[0063] Два полинуклеотида или полипептида являются идентичными, если последовательность нуклеотидов или аминокислотных остатков, соответственно, в двух последовательностях является одинаковыми при их выравнивании на максимальное соответствие, как описано ниже. Термины «идентичный» или «процент идентичности» в отношении двух или более последовательностей нуклеиновых кислот или полипептидов относятся к двум или более последовательностям или подпоследовательностям, которые являются одинаковыми или имеют определенный процент аминокислотных остатков или нуклеотидов, которые являются одинаковыми при их сравнении и выравнивании на максимальное соответствие по окну сравнения, как было определено с использованием одного из нижеприведенных алгоритмов сравнения последовательностей или путем выравнивания вручную и визуального наблюдения. Для полипептидов, последовательности которых отличаются по консервативным заменам, процент идентичности последовательностей может быть скорректирован в сторону его увеличения с учетом консервативной природы замен. Способы осуществления такой коррекции хорошо известны специалистам в данной области. Обычно, они включают оценку консервативных замен как частичного, но не полного несоответствия, что повышает процентную идентичность последовательностей. Так, например, если идентичной аминокислоте присваивают оценку 1, а неконсервативной замене присваивают оценку 0, то консервативная замена получает оценку от 0 до 1. Оценка консервативных замен вычисляют, например, в соответствии с алгоритмом Meyers & Miller, *Computer Applic. Biol. Sci.* 4: 11-17 (1988), например, как это было реализовано в программе PC/GENE (Intelligenetics, Mountain View, CA, USA).

[0064] Термины «по существу, идентичный» и «процент идентичности» в отношении двух нуклеиновых кислот или полипептидов относятся к последовательностям или подпоследовательностям, которые по меньшей мере на 50%, преимущественно на 60%, предпочтительно на 70%, более предпочтительно на 80%, а наиболее предпочтительно на 90-95% идентичны по своим нуклеотидным или аминокислотным остаткам при их выравнивании на максимальное соответствие по окну сравнения, как было определено с использованием одного из нижеприведенных алгоритмов сравнения последовательностей или путем выравнивания вручную и визуального наблюдения. В это определение также входит комплемент тестируемой последовательности, который имеет значительную комплементарность последовательности или подпоследовательности, если тестируемая последовательность, по существу, идентична эталонной последовательности.

[0065] Для специалиста в данной области очевидно, что два полипептида могут быть также «по существу, идентичными», если два полипептида являются иммунологически сходными. Таким образом, общая структура белка может быть сходной, в то время как первичная структура двух полипептидов демонстрирует значительные различия. Следовательно, способ определения того, являются ли два полипептида по существу идентичными, включает оценку уровня связывания моноклональных или

поликлональных антител с каждым полипептидом. Два полипептида являются по существу идентичными, если антитела, специфичные к первому полипептиду, связываются со вторым полипептидом с аффинностью, составляющей по меньшей мере одну треть от аффинности к первому полипептиду. Для сравнения последовательностей, обычно одна последовательность служит в качестве эталонной последовательности, с которой сравнивают тестируемые последовательности. При использовании алгоритма сравнения последовательностей, тестируемые и эталонные последовательности вводят в компьютер, при необходимости выбирают координаты подпоследовательности и вводят параметры программы алгоритма последовательностей. Затем с помощью алгоритма сравнения последовательностей вычисляют процент идентичности последовательностей для тестируемой последовательности по сравнению с эталонной последовательностью исходя из выбранных параметров программы.

[0066] Процент «сходства последовательностей» представляет собой процент аминокислот или нуклеотидов, которые являются либо идентичными, либо отличающимися друг от друга («сходство последовательностей» = процент идентичности последовательности+процент изменений). Таким образом, если используется термин «сходство последовательностей», то он охватывает определенный процент «идентичности» последовательностей и их «модификаций». В некоторых вариантах осуществления изобретения, модификации в последовательностях, допустимые по эталонным процентам идентичности последовательностей, представляют собой все или почти все консервативные замены, то есть, в тех вариантах осуществления изобретения, в которых последовательности являются идентичными на 90%, а оставшиеся 10% представляют собой все или почти все консервативные замены. Термин «почти все» в данном контексте относится по меньшей мере к 75%, более предпочтительно, по меньшей мере 85%, еще более предпочтительно по меньшей мере 90%, а наиболее предпочтительно, по меньшей мере 95% допустимых изменений в последовательностях, которые представляют собой консервативные замены.

[0067] Оптимальное выравнивание последовательностей для сравнения может быть осуществлено, например, с помощью алгоритма выравнивания локальной гомологии Smith & Waterman, 0.4dv. Appl. Math. 2:482 (I 98 I); с помощью алгоритма выравнивания гомологии Needleman & Wunsch, J. Mol. Biol. 48:443 (1970); путем поиска сходства с применением метода Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 5 85:2444 (1988); с помощью компьютерных реализаций этих алгоритмов (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.); с помощью компьютерной программы для выравнивания, такой как VECTOR NTI Version #11.5 by Life Technologies, Carlsbad, CA, USA; № с помощью процедур, описанных ClustalW, Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research, 22:4673-4680); или путем визуального наблюдения (см., в общих чертах

Protocols in Molecular Biology, F. M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1995 Supplement) (Ausubel)).

[0068] Примерами алгоритмов, которые являются подходящими для определения процента идентичности и сходства последовательностей, являются алгоритмы BLAST и BLAST 2.0, которые описаны Altschul et al. (1990) *J. Mol. Biol.* 215: 403-410 и Altschul et al. (1977) *Nucleic Acids Res.* 25: 33 89-3402, соответственно. Программа осуществления анализов BLAST является общедоступной и имеется в Национальном Центре Биотехнологической Информации (<http://www.ncbi.nlm.nih.gov/>). Этот алгоритм включает сначала идентификацию пар последовательностей с высокой оценкой (HSP) путем идентификации коротких слов длиной W в запрашиваемой последовательности, которые либо совпадают, либо удовлетворяют некоторому положительному значению пороговой оценки T при выравнивании со словом той же длины в последовательности базы данных. T означает порог оценки соседних слов (Altschul et al., см. выше). Эти начальные совпадения соседних слов служат в качестве отправной точки для начала поиска более длинных HSP, содержащих эти слова. Затем совпадения слов распространяются в обоих направлениях вдоль каждой последовательности настолько, насколько может быть увеличена совокупная оценка выравнивания. Суммарные баллы рассчитываются с использованием, для нуклеотидных последовательностей, параметров M (балл вознаграждения за пару совпадающих остатков; всегда >0) и N (штрафной балл за несовпадающие остатки; всегда <0). Для аминокислотных последовательностей, для подсчета совокупного балла используется оценочная матрица. Расширение совпадений слов в каждом направлении останавливается, когда: совокупный показатель выравнивания падает на величину X от своего максимального достигнутого значения; суммарный балл падает до нуля или ниже из-за накопления одного или нескольких выравниваний остатков с отрицательным баллом; или когда достигается конец любой последовательности. Параметры алгоритма BLAST W , T и X определяют чувствительность и скорость выравнивания. Программа BLASTN (для нуклеотидных последовательностей) использует по умолчанию длину слова (W), равную 11, математическое ожидание (E), равное 10, $M=5$, $N = -4$, и сравнение обеих цепей. Для аминокислотных последовательностей, программа BLASTP использует по умолчанию длину слова (W), равную 3; математическое ожидание (E), равное 10; и оценочную матрицу BLOSUM62 (см. Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89:10915 (1989)). Помимо вычисления процента идентичности последовательностей, с помощью алгоритма BLAST может быть также проведен статистический анализ сходства между двумя последовательностями (см., например, Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90: 5873-5787 (1993)). Одним из показателей сходства, обеспечиваемого алгоритмом BLAST, является наименьшая сумма вероятностей ($P(N)$), которая указывает на вероятность случайного совпадения между двумя нуклеотидными или аминокислотными последовательностями. Так, например, нуклеиновая кислота считается аналогичной эталонной последовательности, если

наименьшая сумма вероятностей при сравнении тестируемой нуклеиновой кислоты с эталонной нуклеиновой кислотой составляет менее чем приблизительно 0,1, более предпочтительно менее, чем приблизительно 0,01, а наиболее предпочтительно менее, чем приблизительно 0,001.

Нуклеиновые кислоты и их доставка в клетки

[0069] Некоторые аспекты настоящего раскрытия включают нуклеиновые кислоты (например, гены SHP), а также нуклеиновые кислоты, имеющие одну или несколько мутаций. Существуют различные способы индуцирования мутаций в нуклеиновой кислоте, как описано в настоящей заявке. В некоторых вариантах осуществления изобретения, одна или более нуклеиновых кислот могут быть доставлены в клетку, как описано в настоящей заявке.

Олигонуклеоснования

[0070] Используемый здесь термин «олигонуклеоснования» означает полимер из нуклеоснований, которые могут гибридизоваться посредством спаривания оснований Уотсона-Крика с ДНК, имеющей комплементарную последовательность.

[0071] Нуклеоснования включают основание, которое может представлять собой пурин, пиримидин или его производное или аналог. Нуклеоснования включают пептидные нуклеоснования, субъединицы пептид-содержащих нуклеиновых кислот и морфолиновые нуклеоснования, а также нуклеозиды и нуклеотиды. Нуклеозиды представляют собой нуклеоснования, которые содержат пентозофуранозильную группу, например, необязательно замещенный рибозид или 2'-дезоксирибозид. Нуклеозиды могут быть связаны одной из нескольких связывающих групп, которые могут содержать, а могут и не содержать, фосфор. Нуклеозиды, которые связаны незамещенными фосфодифирными связями, называются нуклеотидами.

[0072] Цепь олигонуклеоснований может иметь один 5'- и 3'-конец, которые представляют собой конечные нуклеоснования полимера. Конкретная цепь олигонуклеоснований может содержать нуклеоснования всех типов. Соединение олигонуклеоснований представляет собой соединение, содержащее одну или более цепей олигонуклеоснований, которые являются комплементарными и гибридизуются посредством спаривания оснований Уотсона-Крика. Нуклеоснования представляют собой основания либо дезоксирибо-типа, либо рибо-типа. Нуклеоснования рибо-типа представляют собой пентозофуранозил-содержащие нуклеоснования, где 2'-углерод представляет собой метилен, замещенный гидроксилом, алкилокси или галогеном. Нуклеоснования дезоксирибо-типа представляют собой нуклеоснования, отличающиеся от нуклеоснований рибо-типа, и включают все нуклеоснования, которые не содержат пентозофуранозильную группу.

[0073] Цепь олигонуклеоснований обычно включает цепи олигонуклеоснований и сегменты или области цепей олигонуклеоснований. Цепь олигонуклеоснований имеет 3'-конец и 5'-конец. Когда нить олигонуклеоснований является также и цепью, то 3'- и 5'-концы нити также являются 3'- и 5'-концами цепи.

[0074] Олигонуклеосодержащее вещество может быть также введено в клетку растения любым способом, обычно применяемым в данной области, включая, но не ограничиваясь ими, способ с использованием микроносителей (биобаллистическая доставка) и микроволокон (усов), электропорацию, нуклеофекцию, ПЭГ- опосредуемую доставку, прямое поглощение ДНК и микроинъекцию. Иллюстративные примеры олигонуклеосодержащих веществ описаны ниже.

[0075] Настоящее изобретение может быть осуществлено на практике с использованием олигонуклеосодержащих веществ, имеющих конформации и химические группы, описанные в патентах Kmiec I и Kmiec II, которые включены в настоящее описание посредством ссылки. В Kmiec I описан метод введения специфических генетических модификаций в ген-мишень. Олигонуклеосодержащие вещества в Kmiec I и/или Kmiec II содержат две комплементарные цепи, одна из которых содержит по меньшей мере один сегмент нуклеотидов РНК-типа («сегмент РНК»), основания которого спарены с нуклеотидами ДНК-типа другой цепи.

[0076] В Kmiec II описано, что нуклеотиды могут быть заменены не-нуклеотидами, содержащими пуриновое и пиримидиновое основание. В патентах США № 5756325; 5871984; 5760012; 5888983; 5795972; 5780296; 5945339; 6004804; и 6010907; в Международном патенте No. PCT/US00/23457; и в Международных патентных публикациях WO 98/49350; WO 99/07865; WO 99/58723; WO 99/58702; WO 99/40789; в патенте США No. 6870075; и в опубликованной заявке на патент США 20030084473, каждая из которых включена в настоящее описание в полном объеме посредством ссылки, раскрыты дополнительные молекулы, которые можно использовать в настоящем изобретении. Используемый здесь термин «олигонуклеосодержащее вещество» означает молекулы, которые могут быть использованы в способах согласно изобретению, и включают смешанные дуплексные олигонуклеотиды; молекулы, содержащие не-нуклеотиды, как описано в Kmiec II; одноцепочечные олигодезоксинуклеотиды и другие молекулы, описанные в вышеупомянутых патентах и патентных публикациях.

[0077] В одном варианте осуществления изобретения, олигонуклеосодержащее вещество представляет собой смешанный дуплексный олигонуклеотид, в котором нуклеотиды РНК-типа смешанного дуплексного олигонуклеотида становятся резистентными к РНКазе в результате замены 2'-гидроксила на функциональные группы фтора, хлора или брома, или в результате введения заместителя в положение 2'-О. Подходящими заместителями являются заместители, описанные в патенте Kmiec II. Альтернативными заместителями являются заместители, описанные в патенте США No. 5334711 (Sproat), и заместители, описанные в патентных публикациях EP 629387 и EP 679657 (совместно поданные заявки Martin Applications), которые включены в настоящее описание посредством ссылки. Используемое здесь 2'-фтор-, хлор- или бром-содержащее производное рибонуклеотида или рибонуклеотида, имеющего 2'-ОН, замещенный заместителем, описанным в заявках Martin Applications или Sproat, называется «2'-замещенным рибонуклеотидом». Используемый здесь термин «нуклеотид РНК-типа» означает 2'-гидроксил или 2'-

замещенный нуклеотид, который связан с другими нуклеотидами смешанного дуплексного олигонуклеотида посредством незамещенной фосфодиэфирной связи или любой из не-природных связей, описанных в Kmiec I или Kmiec II. Используемый здесь термин «нуклеотид дезоксирибо-типа» означает нуклеотид, имеющий 2'-Н, который может быть связан с другими нуклеотидами MDON посредством незамещенной фосфодиэфирной связи или любых не-природных связей, описанных в Kmiec I или Kmiec II.

[0078] В одном варианте осуществления изобретения, олигонуклеосооснование или GRON представляет собой смешанный дуплексный олигонуклеотид, который связан исключительно незамещенными фосфодиэфирными связями. В альтернативных вариантах осуществления изобретения, связь представляет собой замещенные фосфодиэфиры, производные фосфодиэфира и нефосфорные связи, как описано в Kmiec II. В еще одном варианте осуществления изобретения, каждый нуклеотид РНК-типа в смешанном дуплексном олигонуклеотиде представляет собой 2'-замещенный нуклеотид. Особенно предпочтительными вариантами 2'-замещенных рибонуклеотидов являются 2'-фтор-, 2'-метокси-, 2'-пропилокси-, 2'-аллилокси-, 2'-гидроксилэтилокси-, 2'-метоксиэтилокси-, 2'-фторпропилокси- и 2'-трифторпропилокси-замещенные рибонуклеотиды. Более предпочтительными вариантами 2'-замещенных рибонуклеотидов являются 2'-фтор-, 2'-метокси-, 2'-метоксиэтилокси- и 2'-аллилокси-замещенные нуклеотиды. В другом варианте осуществления изобретения, смешанный дуплексный олигонуклеотид связан незамещенными фосфодиэфирными связями.

[0079] Хотя смешанный дуплексный олигонуклеотид, имеющий 2'-замещенный нуклеотид РНК-типа только одного типа, является более удобным для синтеза, однако, способы согласно изобретению могут быть осуществлены с использованием смешанных дуплексных олигонуклеотидов, имеющих нуклеотиды РНК-типа двух или более типов. На функцию сегмента РНК может не влиять разрыв, вызванный введением дезоксинуклеотида между двумя тринуклеотидами РНК-типа, и, соответственно, термин «сегмент РНК» охватывает такой «сегмент РНК с разрывом». Сегмент РНК без разрывов называется непрерывным сегментом РНК. В альтернативном варианте осуществления изобретения, сегмент РНК может содержать чередующиеся резистентные к РНКазе и незамещенные 2'-ОН-нуклеотиды. Смешанные дуплексные олигонуклеотиды предпочтительно имеют менее 100 нуклеотидов, а более предпочтительно менее, чем 85 нуклеотидов, но более, чем 50 нуклеотидов. Первая и вторая нити представляют собой пары оснований Уотсона-Крика. В одном варианте осуществления изобретения, нити смешанного дуплексного олигонуклеотида ковалентно связаны посредством линкера, такого как одноцепочечный гекса-, пента- или тетрануклеотид, а поэтому первая и вторая нити представляют собой сегменты одной олигонуклеотидной цепи, имеющей один 3'-конец и один 5'-конец. 3'- и 5'-концы могут быть защищены путем введения «шпилечного кэпа», а поэтому 3'- и 5'-концевые нуклеотиды являются нуклеотидами Уотсона-Крика, спаренными со смежными нуклеотидами. Второй шпилечный кэп может быть

дополнительно введен на стыке между первой и второй цепями, удаленными от 3'- и 5'-концов, что позволяет стабилизировать пары Уотсона-Крика между первой и второй цепями.

[0080] Первая и вторая цепи содержат две области, которые гомологичны двум фрагментам гена-мишени SHP, то есть, они имеют такую же последовательность, как и ген-мишень. Гомологичная область содержит нуклеотиды сегмента РНК и может содержать один или более нуклеотидов ДНК-типа, соединяющих сегмент ДНК, а также она может содержать нуклеотиды ДНК-типа, которые не находятся внутри промежуточного сегмента ДНК. Две области гомологии разделены областью, и каждая из них является смежной с областью, имеющей последовательность, которая отличается от последовательности гена-мишени, и эти области называются «гетерологичными областями». Гетерологичная область может содержать один, два или три ошибочно спаренных нуклеотида. Ошибочно спаренные нуклеотиды могут быть смежными или, альтернативно, они могут быть разделены одним или двумя нуклеотидами, которые гомологичны гену-мишени. Альтернативно, гетерологичная область может также содержать инсерцию одного, двух, трех или пяти или нескольких нуклеотидов. Альтернативно, последовательность смешанного дуплексного олигонуклеотида может отличаться от последовательности гена-мишени только делецией одного, двух, трех или пяти или нескольких нуклеотидов из смешанного дуплексного олигонуклеотида. В этом случае, длина и положение гетерологичной области, предположительно считается длиной делеции, даже если в гетерологичной области отсутствуют нуклеотиды смешанного дуплексного олигонуклеотида. Расстояние между фрагментами гена-мишени, которые комплементарны двум гомологичным областям, идентично длине гетерологичной области, когда предполагается введение замены или замен. Если гетерологичная область содержит инсерцию, то гомологичные области, таким образом, разделяются в смешанном дуплексном олигонуклеотиде дальше, чем их комплементарные гомологичные фрагменты в гене, и наоборот, если гетерологичная область кодирует делецию.

[0081] Каждый из сегментов РНК смешанных дуплексных олигонуклеотидов является частью гомологичной области, то есть области, которая по своей последовательности идентична фрагменту гена-мишени, причем, эти сегменты вместе предпочтительно содержат по меньшей мере 13 нуклеотидов РНК-типа и предпочтительно от 16 до 25 нуклеотидов РНК-типа или еще более предпочтительно 18-22 нуклеотидов РНК-типа или наиболее предпочтительно 20 нуклеотидов. В одном варианте осуществления изобретения, сегменты РНК областей гомологии разделены внутренним сегментом ДНК, и являются смежными с этим сегментом, то есть «сединены» с ним. В одном варианте осуществления изобретения, каждый нуклеотид гетерологичной области представляет собой нуклеотид промежуточного сегмента ДНК. Промежуточный сегмент ДНК, который содержит гетерологичную область смешанного дуплексного олигонуклеотида, называется «сегментом-мутатором».

[0082] Замена, которая должна быть введена в ген-мишень, кодируется

гетерологичной областью. Замена, которая должна быть введена в ген SHP, может представлять собой замену одного или более оснований последовательности гена-мишени, которые заменяют нативную аминокислоту в нужном положении на желаемую аминокислоту.

[0083] В другом варианте осуществления изобретения, олигонуклеосодержащее представляет собой одноцепочечный олигодезоксинуклеотидный мутантный вектор или SSOMV, которые раскрываются в Международной патентной заявке PCT/US00/23457, включенной в настоящее описание в полном объеме посредством ссылки. Последовательность SSOMV основана на тех же принципах, что и мутантные векторы, описанные в патентах США No. 5756325; 5871984; 5760012; 5888983; 5795972; 5780296; 5945339; 6004804; и 6010907 и в публикациях Международных заявок WO 98/49350; WO 99/07865; WO 99/58723; WO 99/58702; WO 99/40789; US 6870075; и в опубликованной заявке на патент США 20030084473. Последовательность SSOMV содержит две области, которые гомологичны последовательности-мишени, разделенной областью, которая содержит желаемую генетическую модификацию, называемую областью-мутатором. Область-мутатор может иметь последовательность, которая имеет такую же длину, как и последовательность, которая разделяет гомологичные области в последовательности-мишени, но имеет другую последовательность. Такая область-мутатор приводит к замене.

[0084] Нуклеотиды SSOMV представляют собой дезоксирибонуклеотиды, которые связаны немодифицированными фосфодиэфирными связями, за исключением того, что 3'-концевая и/или 5'-концевая межнуклеотидная связь или, альтернативно, две 3'-концевые и/или 5'-концевые межнуклеотидные связи могут представлять собой фосфориоат или фосфоамидат. Используемый здесь термин межнуклеотидная связь означает связь между нуклеотидами SSOMV и не включает связь между 3'-концевым нуклеотидом или 5'-концевым нуклеотидом и блокирующим заместителем, см. выше. В конкретном варианте осуществления изобретения, длина SSOMV составляет от 21 до 55 дезоксинуклеотидов; области гомологии имеют, соответственно, общую длину по меньшей мере 20 дезоксинуклеотидов; а по меньшей мере каждая из двух областей гомологии должна иметь длину по меньшей мере 8 дезоксинуклеотидов.

[0085] SSOMV может быть сконструирован так, чтобы он был комплементарен либо кодирующей, либо некодирующей цепи гена-мишени. Если желаемая мутация представляет собой замену одного основания, то предпочтительно, чтобы оба нуклеотида-мутатора представляли собой пиримидин. Для достижения желаемого функционального результата, предпочтительно, чтобы нуклеотид-мутатор и целевой нуклеотид в комплементарной цепи представляли собой пиримидины. Особенно предпочтительными являются SSOMV, которые кодируют мутации трансверсии, то есть, нуклеотид-мутатор С или Т, который ошибочно спаривается, соответственно, с нуклеотидом С или Т в комплементарной цепи.

[0086] Помимо олигодезоксинуклеотида, SSOMV может содержать 5'-блокирующий заместитель, который присоединен к 5'-концевым атомам углерода

посредством линкера. Химическая природа линкера не имеет важного значения, за исключением того, что его длина должна предпочтительно составлять по меньшей мере 6 атомов, и этот линкер должен быть гибким. При этом могут быть использованы различные нетоксичные заместители, такие как биотин, холестерин или другие стероиды или неинтеркалирующий катионный флуоресцентный краситель. Особенно предпочтительными реагентами для получения SSOMV являются реагенты, поставляемые как Cy3TM и Cy5TM компанией Glen Research, Sterling VA (ныне называемой GE Healthcare), которые представляют собой блокированные фосфоамидиты, которые при включении в олигонуклеотид дают 3,3,3',3'-тетраметил-N, N'-изопропил-замещенные индомонокрбозианиновые и индодикарбозианиновые красители, соответственно. Cy3 является наиболее предпочтительным. Если индокарбозианин замещен N-оксиалкилом, то он может легко связываться с 5'-концом олигодезоксинуклеотида посредством фосфодизфира, имеющего 5'-концевой фосфат. Химическая природа линкерного красителя, соединяющего краситель и олигодезоксинуклеотид, не имеет важного значения и была выбрана лишь для удобства синтеза. Если непосредственно используется коммерчески доступный Cy3-фосфоамидит, то полученная 5'-модификация состоит из блокирующего заместителя и линкера, которые вместе представляют собой N-гидроксипропил, N'-фосфатидилпропил-3,3,3',3'-тетраметилиндомонокрбозианин.

[0087] В предпочтительном варианте осуществления изобретения, индокарбозианиновый краситель является тетразамещенным в 3- и 3'-положениях индольных колец. Не ограничиваясь какой-либо конкретной теорией, авторы лишь отмечают, что эти заместители не позволяют красителю быть интеркалирующим красителем. Идентичность заместителей в этих положениях не играет важной роли. Кроме того, SSOMV может иметь 3'-блокирующий заместитель. И в этом случае, химическая природа 3'-блокирующего заместителя не играет важной роли.

[0088] В другом варианте осуществления изобретения, олигонуклеотид может представлять собой одноцепочечный олигодезоксинуклеотид, имеющий 3'-концевой нуклеотид, 5'-концевой нуклеотид, и длину по меньшей мере 25 дезоксинуклеотидов и не более, чем 65 дезоксинуклеотидов, а также имеющий последовательность, содержащую по меньшей мере две области, каждая из которых имеет длину по меньшей мере 8 дезоксинуклеотидов, каждый из которых, соответственно, идентичен по меньшей мере двум областям целевого хромосомного гена, где указанные области, взятые вместе, имеют длину по меньшей мере 24 нуклеотида и разделены по меньшей мере одним нуклеотидом в последовательности целевого хромосомного гена или в последовательности олигодезоксинуклеотида или в обеих последовательностях, а поэтому последовательность олигодезоксинуклеотида не идентична последовательности целевого хромосомного гена. См. патент США No. 6271360, который включен в настоящее описание посредством ссылки.

[0089] Описанные здесь мутации также могут быть получены путем мутагенеза (случайного, соматического или направленного) и другого редактирования ДНК или

посредством нуклеаз с использованием репаративной матрицы, включая, но не ограничиваясь ими, таргетинг гена с использованием нуклеаз «цинковый палец», с использованием эффекторных нуклеаз, подобных активатору транскрипции (TALEN), и с использованием регулярно кластеризованных перемежающихся коротких палиндромных повторов (CRISPR). Эти нуклеазы могут быть получены на основе плазмиды (ДНК), РНК и/или белка.

Микроносители и микроволокна

[0090] Использование металлических микроносителей (микросфер) для введения крупных фрагментов ДНК в клетки растений, имеющих целлюлозные клеточные стенки, посредством биобаллистической пенетрации, хорошо известно специалистам в данной области (далее называемой биобаллистической доставкой). В патентах США № 4945050; 5100792 и 5204253 описаны общие методы отбора микроносителей и устройств для их инъекции. В патентах США № 5484956 и 5489520 описано получение оплодотворенной трансгенной кукурузы посредством бомбардировки микрочастицами ткани каллуса кукурузы. Биобаллистические методы также применяются для трансформации незрелых эмбрионов кукурузы.

[0091] Конкретные условия использования микроносителей в способах согласно изобретению описаны в публикации Международной заявки WO 99/07865. В репрезентативном методе, охлажденные льдом микроносители (60 мг/мл), смешанный дуплексный олигонуклеотид (60 мг/мл), 2,5 М CaCl₂ и 0,1 М спермидина добавляют в указанном порядке; а затем смесь осторожно перемешивают, например, путем встряхивания в течение 10 минут, и оставляют на 10 минут при комнатной температуре, после чего микроносители разводят в 5 объемах этанола, центрифугируют и ресуспендируют в 100% этаноле. Хорошие результаты могут быть получены в адгезивном растворе при концентрациях 8-10 мкг/мкл микроносителей, 14-17 мкг/мл смешанного дуплексного олигонуклеотида, 1,1-1,4М CaCl₂ и 18-22 мМ спермидина. Оптимальные результаты наблюдались при концентрациях 8 мкг/мкл микроносителей, 16,5 мкг/мл смешанного дуплексного олигонуклеотида, 1,3 М CaCl₂ и 21 мМ спермидина.

[0092] Олигонуклеос основания могут быть также введены в клетки растений для практического применения настоящего изобретения с использованием микроволокон для проникновения через клеточную стенку и клеточную мембрану. В патенте США № 5302523, выданном Coffee R. и Dunwell J.M. (1994), описано использование волокон карбида кремния размером 30 × 0,5 мкм и 10 × 0,3 мкм для облегчения трансформации суспензионных культур кукурузы сорта Black Mexican Sweet. Любая механическая методика, которая может быть применена для введения ДНК для трансформации клетки растения с использованием микроволокон, может быть применена для доставки олигонуклеос оснований для их использования в целях получения мутантов SHP согласно изобретению. Способ, описанный Coffee R. и Dunwell J.M. (1994) в патенте США № 5302523, может быть применен с использованием регенерируемых материалов клеток растений для введения олигонуклеос оснований согласно изобретению в целях достижения

мутации гена SHP.

[0093] Иллюстративный способ доставки олигонуклеосложения с использованием микроволокон заключается в следующем: стерильные микроволокна (2 мкг) суспендируют в 150 мкл среды для культивирования растений, содержащей приблизительно 10 мкг смешанного дуплексного олигонуклеотида. Суспензионную культуру оставляют для осаждения, и равные объемы упакованных клеток и стерильной суспензии волокна/нуклеотида встряхивают в течение 10 минут и высевают. Селективные среды наносят немедленно или с задержкой до 120 часов, в зависимости от конкретного признака.

Электропорация

[0094] В альтернативном варианте осуществления изобретения, олигонуклеосложения могут быть доставлены в клетку растения путем электропорации протопласта, происходящего от части растения, в соответствии с методами, хорошо известными специалисту в данной области. См., например, Gallois et al., 1996, в *Methods in Molecular Biology* 55: 89-107, Humana Press, Totowa, N.J.; Kipp et al., 1999, в *Methods in Molecular Biology* 133: 213-221, Humana Press, Totowa, N.J.

[0095] Олигонуклеосложения могут быть также введены в микроспоры путем электропорации. После высвобождения тетрады, микроспора становится одноядерной и имеет тонкие стенки. Затем она начинает увеличиваться и развивается в зародышевую оболочку до того, как образуется экзин. Микроспора на этой стадии потенциально более подвержена трансформации экзогенной ДНК, чем другие клетки растений. Кроме того, развитие микроспор может быть изменено *in vitro* для получения гаплоидных эмбрионов или эмбрионного каллуса, которые могут быть регенерированы в растения (Coumans et al., *Plant Cell Rep.* 7: 618-621, 1989; Datta et al., *Plant Sci.* 67: 83-88, 1990; Maheshwari et al., *Am. J. Bot.* 69: 865-879, 1982; Schaeffer, *Adv. In Cell Culture* 7: 161-182, 1989; Swanson et al., *Plant Cell Rep.* 6: 94-97, 1987). Таким образом, трансформированные микроспоры могут быть регенерированы непосредственно в гаплоидные растения или дигаплоидные фертильные растения при удвоении хромосом стандартными методами. См. также совместно поданную заявку США рег. № 09/680858, озаглавленную «Композиции и способы генетической модификации растений», где указанная заявка включена в настоящее описание посредством ссылки.

[0096] Электропорация микроспор может быть проведена для растений любых видов, для которых возможно культивирование микроспор, включая, но не ограничиваясь ими, растения семейства Graminae, Leguminosae, Cruciferae, Solanaceae, Cucurbitaceae, Rosaceae, Poaceae, Lilaceae, Rutaceae, Vitaceae, включая такие виды, как кукуруза (*Zea mays*), пшеница (*Triticum aestivum*), рис (*Oryza sativa*), овес, ячмень, канола (*Brassica napus*, *Brassica rapa*, *Brassica oleracea* и *Brassica juncea*), хлопчатник (*Gossypium hirsutum* L.), различные виды бобовых (например, соя (*Glycine max*), горох (*Pisum sativum*) и т.п.), виноград (*Vitis vinifera*) и множество других важных сельскохозяйственных культур. Эмбриогенез микроспор, как из культуры пыльников, так и из микроспор, был

описан для более, чем 170 видов, относящихся к 68 родам и 28 семействам двудольных и однодольных растений (Raghavan, Embryogenesis in Angiosperms: A Developmental and Experimental Study, Cambridge University Press, Cambridge, England, 1986; Raghavan, Cell Differentiation 21:213-226, 1987; Raemakers et al., Euphytica 81:93-107, 1995). Подробное обсуждение выделения микроспор, культивирования и регенерации двойных гаплоидных растений из зародышей, происходящих от микроспор (MDE) в *Brassica napus* L., можно найти в публикации Nehlin, The Use of Rapeseed (*Brassica napus* L.) Microspores as a Tool for Biotechnological Applications, doctoral thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden, 1999; also Nehlin et al., Plant Sci. 111:219-227, 1995, and Nehlin et al., Plant Sci. 111:219-227, 1995). Удвоение хромосом из культуры микроспор или пыльников является общепризнанным методом получения двойных гаплоидных гомозиготных линий растений у нескольких культур (Heberle-Bors et al., In vitro pollen cultures: Progress and perspectives. In: Pollen Biotechnology. Gene expression and allergen characterization, vol. 85-109, ed. Mohapatra, S. S., and Knox, R. B., Chapman and Hall, New York, 1996).

[0097] Методы электропорации микроспор описаны в публикациях Jardinaud et al., Plant Sci. 93: 177-184, 1993, и Fennell and Hauptman, Plant Cell Reports 11: 567-570, 1992. Методы электропорации MDON в протопласты растений могут быть также адаптированы для использования в электропорации микроспор.

Метод усов

[0098] В еще одном альтернативном варианте осуществления изобретения, олигонуклеосодержащее основание может быть доставлено в клетку растения с помощью усов или микроинъекции. Так называемый метод усов осуществляют, в основном, как описано в Frame et al., 1994, Plant J. 6: 941-948. Олигонуклеосодержащие основания добавляют в усы и используют для трансформации клеток растения. Олигонуклеосодержащее основание может быть совместно инкубировано с плазмидами, содержащими последовательности, кодирующие белки, способные образовывать рекомбиназные комплексы в клетках растений, так, чтобы такая рекомбинация катализировалась между олигонуклеотидом и последовательностью-мишенью в гене SHP.

Другие способы доставки

[0099] В альтернативном варианте осуществления изобретения, нуклеиновые кислоты внедряют в микросферы, состоящие из альгината кальция, и эти нуклеиновые кислоты поглощаются растительными протопластами в присутствии агента полиэтиленгликоля, модифицирующего мембрану (см., например, Sone et al., 2002; Liu et al., 2004).

[0100] В альтернативном варианте осуществления изобретения, нуклеиновые кислоты замораживают в воде и вводят в клетки растения путем бомбардировки микрочастицами (см., например, Gilmore, 1991, патент США 5219746; Brinegar et al.).

[0101] В альтернативном варианте осуществления изобретения, нуклеиновые кислоты, присоединенные к наночастицам, вводят в интактные клетки растений путем инкубирования клеток в суспензии, содержащей наночастицы (см., например, Pasupathy et

al., 2008), или путем доставки их в интактные клетки путем бомбардировки частицами или в протопласты путем совместного инкубирования (см., например, Torney et al., 2007).

[0102] В альтернативном варианте осуществления изобретения, нуклеиновые кислоты, образующие комплекс с проникающими пептидами, доставляются в клетки путем совместного инкубирования (см., например, Chugh et al., 2008, WO 2008148223 A1; Eudes и Chugh).

[0103] В альтернативном варианте осуществления изобретения, нуклеиновые кислоты вводят в интактные клетки посредством электропорации (см., например, He et al., 1998, U.S. 2003/0115641 A1, Dobres et al.).

[0104] В альтернативном варианте осуществления изобретения, нуклеиновые кислоты доставляются в клетки сухих эмбрионов путем их смачивания в растворе с нуклеиновыми кислотами (путем смачивания сухих эмбрионов (см., например, Töpfer et al., 1989, Senaratna et al., 1991)).

Целевая модификация гена

[0105] Целевая генетическая модификация, опосредуемая олигонуклеотидами, является ценным методом, применяемым для специфического изменения коротких фрагментов ДНК и введения делеций, коротких вставок и точковых мутаций, и этот метод может быть применен в комбинации с раскрываемыми здесь способами, например, для введения одной или более рассматриваемых здесь мутаций SHP. В некоторых вариантах осуществления изобретения, эти способы включают спаривание/отжиг ДНК с последующим событием репарации ДНК. Сначала, нуклеиновую кислоту гибридизуют с комплементарной ей цепью в двухцепочечной ДНК в способе на основе клеточных белковых факторов. Эта гибридизация приводит к образованию центрально расположенной пары ошибочно спаренных оснований (в случае точковой мутации), и тем самым к структурной перестановке, которая, с большой вероятностью, будет стимулировать запуск эндогенным белком механизма инициации второй стадии процесса репарации, то есть, сайт-специфической модификации хромосомной последовательности и/или последовательности в органеллах (например, в митохондриях и хлоропластах). Это только что введенное ошибочное спаривание индуцирует механизм репарации ДНК с последующей инициацией второго события репарации, что приводит к конечной репарации сайта-мишени. В различных описанных здесь аспектах и вариантах, способы и композиции согласно изобретению могут быть усовершенствованы благодаря разработке новых подходов, позволяющих повысить доступность компонентов репарации ДНК, и тем самым повысить эффективность и воспроизводимость модификаций, опосредуемых репарацией генов в нуклеиновых кислотах-мишенях.

[0106] Эффективные методы сайт-направленной геномной модификации являются желательными для исследований, клинической генотерапии, промышленной микробиологии и сельского хозяйства. В одном из таких подходов используются триплекс-образующие олигонуклеотиды (TFO), которые связываются в качестве третьих цепей с дуплексной ДНК по последовательность-специфическому механизму и

опосредуют прямой мутагенез. Такие TFO могут действовать либо посредством доставки связанного мутагена, такого как псорален или хлорамбуцил (Havre et al., Proc Nat'l Acad Sci, USA 90: 7879-7883, 1993; Havre et al., J Virol 67: 7323-7331, 1993; Wang et al., Mol Cell Biol 15: 1759-1768, 1995; Takasugi et al., Proc Nat'l Acad Sci, USA 88: 5602-5606, 1991; Belousov и др., Nucleic Acids Res 25: 3440-3444, 1997) или посредством связывания с аффинностью, достаточной для инициации репарации с вероятностью ошибки (Wang et al., Science 271: 802-805, 1996).

[0107] Другая стратегия геномной модификации, которая может быть применена в комбинации с описанными здесь композициями и способами, включает индуцирование гомологичной рекомбинации между экзогенным фрагментом ДНК и целевым геном. Этот подход был успешно применен для нацеливания и разрушения выбранных генов в клетках млекопитающих и позволил получить трансгенных мышей, несущих гены со специфическим нокаутом (Carpeschi et al., Science 244: 1288-1292, 1989; Wagner, патент США № 4873191). Этот подход включает перенос селективных маркеров для выделения желаемых рекомбинантов. Без отбора, отношение гомологичной и негомологичной интеграции трансфицированной ДНК в типичных экспериментах по переносу генов является низким, а обычно составляет в пределах 1:1000 или менее (Sedivy et al., Gene Targeting, WH Freeman and Co., New York, 1992). Эта низкая эффективность гомологичной интеграции ограничивает эффективность переноса гена для экспериментального использования или генотерапии. Частота целевой мутации может быть увеличена путем повреждения сайта-мишени УФ-облучением и выбранными канцерогенами (Wang et al., Mol Cell Biol. 8: 196-202, 1988), а также сайт-специфическими эндонуклеазами (Sedivy et al., Gene Targeting, WH Freeman and Co., New York, 1992; Rouet et al., Proc Nat'l Acad Sci, USA 91: 6064-6068, 1994; Segal et al., Proc Nat'l Acad Sci, USA 92: 806-810, 1995). Кроме того, повреждение ДНК, индуцированное триплекс-опосредуемыми фотоаддуктами псоралена, может стимулировать рекомбинацию во внехромосомных векторах и между ними (Segal et al., Proc Nat'l Acad Sci, USA 92: 806-810, 1995; Faruqi et al., Mol Cell Biol 16: 6820-6828, 1996; Glazer, патент США № 5962426).

[0108] Линейные донорные фрагменты являются более эффективными для целевой мутации, чем их кольцевые аналоги (Folger et al., Mol Cell Biol 2: 1372-1387, 1982). В некоторых вариантах осуществления изобретения, рекомбинация может также зависеть от длины непрерывной гомологии между донорными сайтами и сайтами-мишенями, причем, короткие фрагменты часто оказываются неэффективными субстратами (Rubnitz et al., Mol Cell Biol 4: 2253-2258, 1984). Тем не менее, использование коротких фрагментов ДНК или гибридов ДНК/РНК для коррекции генов находится в центре внимания различных стратегий (Kunzelmann et al., Gene Ther 3: 859-867, 1996).

[0109] Используемые здесь термины «последовательность нуклеиновой кислоты», «нуклеотидная последовательность» и «полинуклеотидная последовательность» означают олигонуклеотид или полинуклеотид и их фрагменты или части, а также ДНК или РНК геномного или синтетического происхождения, которые могут быть одноцепочечными

или двухцепочечными и представляют собой смысловые или бессмысловые цепи.

[0110] Используемые здесь термины «олигонуклеотид» и «олигомер» означают полимер из нуклеосахаридов. В некоторых вариантах осуществления изобретения, «олигонуклеотид» или «олигомер» может иметь длину по меньшей мере приблизительно 8 нуклеосахаридов или может иметь приблизительно до 1500 нуклеосахаридов или более. В некоторых вариантах осуществления изобретения, «олигонуклеотид» или «олигомер» может иметь любую длину, рассматриваемую в настоящей заявке.

[0111] Используемые здесь термины «ДНК-модифицирующая молекула» и «ДНК-модифицирующий реагент» означают молекулу, способную распознавать последовательность нуклеиновой кислоты и специфически связываться с этой последовательностью в геноме клетки, и модифицировать нуклеотидную последовательность-мишень в геноме, где распознавание и специфическое связывание ДНК-модифицирующей молекулы с последовательностью нуклеиновой кислоты не зависит от белка. Используемый здесь термин «независимый от белка», если он относится к ДНК-модифицирующей молекуле, означает, что ДНК-модифицирующая молекула не требует присутствия и/или активности белка и/или фермента для распознавания и/или специфического связывания с последовательностью нуклеиновой кислоты. Репрезентативными ДНК-модифицирующими молекулами являются, но не ограничиваются ими, триплекс-образующие олигонуклеотиды, пептид-содержащие нуклеиновые кислоты, полиамиды и олигонуклеотиды, которые предназначены для стимуляции конверсии генов. В некоторых вариантах осуществления изобретения, ДНК-модифицирующие молекулы согласно изобретению отличаются от известных последовательностей нуклеиновых кислот, используемых для гомологичной рекомбинации (Wong & Capeschi, *Molec. Cell. Biol.* 7: 2294-2295, 1987) тем, что известные последовательности нуклеиновой кислоты, используемые для гомологичной рекомбинации, зависят от белка. Используемый здесь термин «зависимый от белка», если он относится к молекуле, означает, что эта молекула требует присутствия и/или активности белка и/или фермента для распознавания и/или для специфического связывания этой молекулы с последовательностью нуклеиновой кислоты. Способы определения зависимости ДНК-модифицирующей молекулы от присутствия и/или активности белка и/или фермента для распознавания и/или специфического связывания с последовательностью нуклеиновой кислоты, хорошо известны специалистам в данной области (см., например, Dennis и др. *Nucl. Acids Res.* 27: 4734-4742, 1999). Так, например, ДНК-модифицирующая молекула может быть инкубирована *in vitro* с последовательностью нуклеиновой кислоты в отсутствие каких-либо белков и/или ферментов. Детектирование специфического связывания между ДНК-модифицирующей молекулой и последовательностью нуклеиновой кислоты показало, что ДНК-модифицирующая молекула не зависит от белка. С другой стороны, отсутствие специфического связывания между ДНК-модифицирующей молекулой и последовательностью нуклеиновой кислоты показало, что ДНК-модифицирующая

молекула зависит от белка и/или требует присутствия дополнительных факторов.

[0112] «Триплекс-образующий олигонуклеотид» (ТФО) определяют как последовательность ДНК или РНК, которая способна связываться в основной бороздке дуплексной спиральной ДНК или РНК с образованием тройной спирали. Хотя ТФО не ограничен какой-либо конкретной длиной, однако, предпочтительная длина ТФО составляет 250 нуклеотидов или менее, 200 нуклеотидов или менее, или 100 нуклеотидов или менее, или от 5 до 50 нуклеотидов, или от 10 до 25 нуклеотидов, или от 15 до 25 нуклеотидов. Хотя между ТФО и дуплексной ДНК необходима определенная степень специфичности последовательности для образования тройной спирали, однако, какой-либо конкретной степени специфичности не требуется, при условии, что может образовываться тройная спираль. Аналогичным образом, между ТФО и дуплексной спиралью не требуется какой-либо конкретной степени авидности или аффинности, при условии, что может образовываться тройная спираль. Хотя в одном варианте осуществления изобретения, длина нуклеотидной последовательности, с которой специфически связывается ТФО, не имеет конкретных ограничений, однако, длина нуклеотидной последовательности, с которой специфически связывается ТФО, может составлять от 1 до 100, в некоторых вариантах, от 5 до 50, в некоторых вариантах, от 10 до 25, а в других вариантах, от 15 до 25 нуклеотидов. Кроме того, «тройную спираль» определяют как двухспиральную нуклеиновую кислоту с олигонуклеотидом, связанным с последовательностью-мишенью в нуклеиновой кислоте с двойной спиралью. «Двухспиральная» нуклеиновая кислота может представлять собой любую двухцепочечную нуклеиновую кислоту, включая двухцепочечную ДНК, двухцепочечную РНК и смешанные дуплексы ДНК и РНК. Двухцепочечная нуклеиновая кислота не ограничена какой-либо конкретной длиной. Однако, в предпочтительных вариантах осуществления изобретения, она имеет длину более 500 п.о., в некоторых вариантах, более, чем 1 т.п.о., а в других вариантах, более, чем приблизительно 5 т.п.о.. Во многих применениях, двухспиральной нуклеиновой кислотой является клеточная геномная нуклеиновая кислота. Триплекс-образующий олигонуклеотид может связываться с последовательностью-мишенью параллельным или антипараллельным образом.

[0113] «Пептид-содержащие нуклеиновые кислоты», «полиамиды» или «PNA» представляют собой нуклеиновые кислоты, в которых фосфатный остов заменен полиамидной структурой на основе N-аминоэтилглицина. PNA обладают более высокой аффинностью к комплементарным нуклеиновым кислотам, чем их природные аналоги, в соответствии с правилами спаривания оснований Уотсона-Крика. PNA могут образовывать в высокой степени стабильные трехспиральные структуры с ДНК в соответствии с нижеследующей стехиометрией: (PNA)₂.ДНК. Хотя пептид-содержащие нуклеиновые кислоты и полиамиды не ограничены какой-либо конкретной длиной, однако, предпочтительная длина пептид-содержащих нуклеиновых кислот и полиамидов составляет 200 нуклеотидов или менее, в некоторых вариантах, 100 нуклеотидов или менее, а в других вариантах, от 5 до 50 нуклеотидов. Длина нуклеотидной

последовательности, с которой специфически связывается пептид-содержащая нуклеиновая кислота и полиамид, не имеет конкретных ограничений, однако, в одном варианте осуществления изобретения, нуклеотидная последовательность, с которой специфически связываются пептид-содержащая нуклеиновая кислота и полиамид, имеет длину от 1 до 100, в некоторых вариантах, от 5 до 50, в некоторых вариантах, от 5 до 25, и в других вариантах, от 5 до 20 нуклеотидов.

[0114] Термин «способный модифицировать ДНК» или «средство для модификации ДНК» относится к процедурам, а также к эндогенным или экзогенным агентам или реагентам, которые могут индуцировать или могут способствовать индуцированию модификаций в нуклеотидной последовательности целевого сегмента ДНК. Такие модификации могут быть внесены путем делеции, добавления или замены одного или более оснований на целевом сегменте ДНК. При этом, нет необходимости в том, чтобы изменения последовательности ДНК вызывали функциональные изменения любого гена, кодируемого целевой последовательностью. Кроме того, нет необходимости вносить изменения в ДНК в какой-либо конкретной части или в каком-либо процентном количестве клеток.

[0115] Термин «представляющая интерес нуклеотидная последовательность» означает любую нуклеотидную последовательность, которую специалисту в данной области желательно модифицировать по любой причине. Такие нуклеотидные последовательности включают, но не ограничиваются ими, кодирующие последовательности структурных генов (например, репортерных генов, генов селективных маркеров, онкогенов, генов резистентности к лекарственным средствам, факторов роста и т.п.) и не кодирующие регуляторные последовательности, которые не кодируют мРНК или белковый продукт (например, промоторная последовательность, энхансерная последовательность, последовательность полиаденилирования, последовательность терминации, регуляторные РНК, такие как миРНК и т.п.).

[0116] Используемые здесь термины «аминокислотная последовательность», «полипептидная последовательность», «пептидная последовательность» и «пептид», употребляемые для обозначения последовательности аминокислот, являются синонимами.

[0117] Используемый здесь термин «последовательность-мишень» означает двухспиральную нуклеиновую кислоту, содержащую представляющую интерес последовательность. В некоторых вариантах осуществления изобретения, последовательность-мишень может иметь длину более, чем 8 нуклеотидов, а в некоторых вариантах - менее, чем 1500 нуклеотидов. В некоторых вариантах осуществления изобретения, последовательность-мишень имеет от 8 до 30 оснований. В некоторых вариантах осуществления изобретения, длина последовательности-мишени может иметь длину приблизительно от 75 до 250 оснований. В некоторых вариантах осуществления изобретения, последовательность-мишень может иметь длину, комплементарную длине рассматриваемого здесь олигонуклеотида. Последовательность-мишень обычно определяют как нуклеотидную последовательность на одной из цепей двухспиральной

нуклеиновой кислоты.

[0118] Используемый здесь термин «богатая пурином последовательность» или «последовательность полипурина», если он относится к нуклеотидной последовательности на одной из цепей двухспиральной последовательности нуклеиновой кислоты, определяют как непрерывную нуклеотидную последовательность, в которой более, чем 50% нуклеотидов последовательности-мишени содержат пуриновое основание. Однако предпочтительно, чтобы богатая пурином последовательность-мишень содержала более 60% пуриновых нуклеотидов, в некоторых вариантах, более 75% пуриновых нуклеотидов, в некоторых вариантах, более 90% пуриновых нуклеотидов, а в других вариантах, 100% пуриновых нуклеотидов.

[0119] Используемый здесь термин «богатая пиримидином последовательность» или «последовательность полипиримидина», если он относится к нуклеотидной последовательности на одной из цепей двухспиральной последовательности нуклеиновой кислоты, определяют как непрерывную нуклеотидную последовательность, в которой более, чем 50% нуклеотидов последовательности-мишени содержат пиримидиновое основание. Однако предпочтительно, чтобы богатая пиримидином последовательность-мишень содержала более 60% пиримидиновых нуклеотидов, а в некоторых вариантах, более 75% пиримидиновых нуклеотидов. В некоторых вариантах осуществления изобретения, последовательность содержит более 90% пиримидиновых нуклеотидов, а в других вариантах, 100% пиримидиновых нуклеотидов.

[0120] «Вариант» первой нуклеотидной последовательности определяют как нуклеотидную последовательность, которая отличается от первой нуклеотидной последовательности (например, наличием одной или более делеций, инсерций или замен, которые могут быть детектированы с помощью анализов на гибридизацию или путем секвенирования ДНК). В это определение входит детектирование изменений или модификаций геномной последовательности первой нуклеотидной последовательности. Так, например, анализы на гибридизацию могут быть проведены для детектирования (1) изменений паттерна фрагментов рестриктирующих ферментов, способных гибридизоваться с первой нуклеотидной последовательностью, если она содержится в геноме (то есть, в анализе RFLP), (2) неспособности выбранной части первой нуклеотидной последовательности гибридизоваться с образцом геномной ДНК, которая содержит первую нуклеотидную последовательность (например, с использованием аллель-специфических олигонуклеотидных зондов), (3) неправильной или случайной гибридизации, такой как гибридизация с локусом, отличающимся от нормального хромосомного локуса для первой нуклеотидной последовательности (например, с использованием флуоресцентной гибридизации *in situ* (FISH) с областями хромосом на стадии метафазы и т.п.). Одним из примеров такого варианта является мутированная последовательность дикого типа.

[0121] Используемые здесь термины «нуклеиновая кислота» и «немодифицированная нуклеиновая кислота» означают любое из известных четырех

оснований дезоксирибонуклеиновой кислоты (то есть, гуанин, аденин, цитозин и тимин). Термин «модифицированная нуклеиновая кислота» означает нуклеиновую кислоту, структура которой была изменена по сравнению со структурой немодифицированной нуклеиновой кислоты. Репрезентативным примером таких модификаций являются замены посредством ковалентных модификаций оснований, таких как алкилирование аминогрупп и атомов азота на кольце, а также насыщение двойных связей.

[0122] Используемые здесь термины «мутация» и «модификация» и их грамматические эквиваленты, если они употребляются в отношении последовательности нуклеиновой кислоты, являются синонимами и означают делецию, инсерцию, замену, разрыв цепи и/или введение аддукта. «Делецию» определяют как изменение последовательности нуклеиновой кислоты, в которой отсутствуют один или более нуклеотидов. Используемые здесь термины «инсерция» или «добавление» означают модификацию последовательности нуклеиновой кислоты, которая способствует добавлению одного или более нуклеотидов. «Замена» является результатом замены одного или более нуклеотидов молекулой, которая представляет собой другую молекулу по сравнению с замененными одно или более нуклеотидами. Так, например, нуклеиновая кислота может быть заменена другой нуклеиновой кислотой, например, путем замены тимина на цитозин, аденин, гуанин или уридин. Замена пиримидин на пиримидин (например, нуклеотидные замены С на Т или Т на С) или пурина на пурин (например, нуклеотидные замены G на A или A на G) называются транзициями, а замены пиримидина на пурин или пурина на пиримидин (например, G на Т или G на С или А на Т или А на С) называются трансверсиями. Альтернативно, нуклеиновая кислота может быть заменена модифицированной нуклеиновой кислотой, например, путем замены тимина на тимингликоль. Мутации могут приводить к ошибочному спариванию. Термин «ошибочное спаривание» означает нековалентное взаимодействие между двумя нуклеиновыми кислотами, каждая из которых находится на различных последовательностях полинуклеиновой кислоты, и которые не подчиняются правилам спаривания оснований. Так, например, для частично комплементарных последовательностей 5'-AGT-3' и 5'-AAT-3' наблюдается ошибочное спаривание G-A (транзиция). Термины «введение аддукта» или «образование аддукта» означают ковалентное или нековалентное связывание молекулы с одним или более нуклеотидами в последовательности ДНК, так, чтобы такое связывание приводило к снижению (в некоторых вариантах, от 10% до 100%, в других вариантах, от 50% до 100%, а в некоторых вариантах, от 75% до 100%) уровня репликации и/или транскрипции ДНК.

[0123] Термин «ДНК-разрезающий фермент» означает молекулу, которая вызывает разрыв цепи. Неограничивающими примерами являются мегануклеазы, TALE/TALEN, антибиотики, «цинковые пальцы» и CRISPR или системы CRISPR/Cas.

[0124] Термин «разрыв цепи» применительно к последовательности двухцепочечной нуклеиновой кислоты включает разрыв одной цепи и/или разрыв двойной цепи. Одноцепочечный разрыв (ник) означает прерывание в одной из двух цепей

двухцепочечной последовательности нуклеиновой кислоты. Этот разрыв отличается от двухцепочечного разрыва, который означает прерывание в обеих цепях двухцепочечной последовательности нуклеиновой кислоты и может приводить к образованию тупых или ступенчатых концов. Разрывы цепи могут быть введены в двухцепочечную последовательность нуклеиновой кислоты либо непосредственно (например, посредством ионизирующего излучения или обработки определенными химическими веществами), либо опосредованно (например, путем ферментативного вырезания в определенном положении основания нуклеиновой кислоты). В некоторых вариантах осуществления изобретения, ДНК-разрезающий фермент может обладать селективностью к некоторым специфическим последовательностям, например, в случае CRISPR, цинкового пальца, мегануклеазы, TALEN, как описано в настоящей заявке.

[0125] Термины «мутантная клетка» и «модифицированная клетка» означают клетку, которая содержит по меньшей мере одну модификацию в геномной последовательности клетки.

[0126] Термин «часть», если он относится к нуклеотидной последовательности, означает фрагменты этой нуклеотидной последовательности. Эти фрагменты могут иметь размер от 5 нуклеотидных остатков до полноразмерной нуклеотидной последовательности минус один остаток нуклеиновой кислоты.

[0127] Считается, что молекулы ДНК имеют «5'-концы» и «3'-концы», если мононуклеотиды реагируют с образованием олигонуклеотидов таким образом, что 5'-фосфат пентозного кольца одного мононуклеотида присоединяется к 3'-кислороду соседней группы в одном направлении посредством фосфодиэфирной связи. Следовательно, конец олигонуклеотида называют «5'-концом», если его 5'-фосфат не связан с 3'-кислородом пентозного кольца мононуклеотида. Конец олигонуклеотида называют «3'-концом», если его 3'-кислород не связан с 5'-фосфатом пентозного кольца другого мононуклеотида. Используемая здесь последовательность нуклеиновой кислоты, даже если она находится внутри более крупного олигонуклеотида, может также иметь 5'- и 3'-концы. В линейной или кольцевой молекуле ДНК, дискретные элементы называются «вышерасположенными элементами» или «5'-элементами» или «нижерасположенными элементами» или «3'-элементами». Эта терминология отражает направление транскрипции от 5'- до 3' вдоль всей цепи ДНК. Промоторные и энхансерные элементы, которые направляют транскрипцию связанного гена, обычно расположены у 5'-конца или выше кодирующей области. Однако энхансерные элементы могут оказывать свое влияние, даже если они расположены у 3'-конца промоторного элемента и кодирующей области. Сигналы терминации транскрипции и полиаденилирования расположены у 3'-конца или ниже кодирующей области.

[0128] Используемый здесь термин «рекомбинантная молекула ДНК» означает молекулу ДНК, которая состоит из сегментов ДНК, связанных вместе с применением методов молекулярной биологии.

[0129] Используемый здесь термин «рекомбинантный белок» или

«рекомбинантный полипептид» означает молекулу белка, которая экспрессируется из молекулы рекомбинантной ДНК.

[0130] Используемые здесь термины «вектор» и «носитель» являются синонимами и означают молекулы нуклеиновой кислоты, которые переносят один или более сегментов ДНК из одной клетки в другую.

[0131] Используемые здесь термины «в функциональной комбинации», «в функциональном порядке» и «функционально присоединенный» относятся к связыванию последовательностей нуклеиновых кислот таким образом, что молекула нуклеиновой кислоты способна регулировать транскрипцию данного гена и/или синтез желаемой белковой молекулы. Эти термины также относятся к связыванию аминокислотных последовательностей так, чтобы это приводило к продуцированию функционального белка.

[0132] Используемый здесь термин «трансфекция» означает введение чужеродной ДНК в клетки. Трансфекция может быть осуществлена различными известными способами, включая копреципитацию фосфатом кальция-ДНК, трансфекцию, опосредуемую DEAE-декстраном, трансфекцию, опосредуемую полибренном, электропорацию, микроинъекцию, слияние липосом, трансфекцию, опосредуемую липофектином, слияние протопластов, ретровирусную инфекцию, биобаллистические методы (т.е. бомбардировку частицами) и т.п.

[0133] Используемые здесь термины «комплементарный» или «комплементарность» относятся к «полинуклеотидам» и «олигонуклеотидам» (которые являются здесь синонимами, и означают нуклеотидные последовательности), связанным в соответствии с правилами спаривания оснований. Так, например, последовательность «5'-CAGT-3'» является комплементарной последовательности «5'-ACTG-3'». Комплементарность может быть «частичной» или «полной». «Частичная» комплементарность означает ситуацию, когда одно или более оснований нуклеиновой кислоты не подчиняются правилам спаривания оснований. «Общая» или «полная» комплементарность между нуклеиновыми кислотами наблюдается в случае, когда каждое и все основания нуклеиновой кислоты спариваются с другим основанием в соответствии с правилами спаривания оснований. Степень комплементарности между цепями нуклеиновой кислоты может оказывать существенное влияние на эффективность и силу гибридизации между цепями нуклеиновой кислоты. Это может иметь особое значение в реакциях амплификации, а также в методах детектирования, которые зависят от связывания между нуклеиновыми кислотами. Для удобства, термины «полинуклеотиды» и «олигонуклеотиды» включают молекулы, которые содержат нуклеозиды.

[0134] Термины «гомология» и «гомологичный», используемые в настоящей заявке применительно к нуклеотидным последовательностям, означают степень комплементарности с другими нуклеотидными последовательностями. Возможна частичная гомология или полная гомология (то есть, идентичность). Термин «по существу гомологичный», если он используется в отношении двухцепочечной последовательности

нуклеиновой кислоты, такой как кДНК или геномный клон, относится к любой последовательности нуклеиновой кислоты (например, к зонду), которая может гибридизоваться с одной или обеими цепями двухцепочечной последовательности нуклеиновой кислоты в условиях низкой жесткости, как описано выше. Нуклеотидная последовательность, которая является частично комплементарной, то есть, «по существу гомологичной» последовательности нуклеиновой кислоты, представляет собой последовательность, которая по меньшей мере частично ингибирует гибридизацию полностью комплементарной последовательности с последовательностью нуклеиновой кислоты-мишени. Ингибирование гибридизации полностью комплементарной последовательности с последовательностью-мишенью может быть оценено с помощью анализа на гибридизацию (Саузерн или Нозерн-блот-анализа, анализа на гибридизацию в растворе и т.п.) в условиях низкой жесткости. По существу, гомологичная последовательность или зонд будут конкурировать за связывание и ингибировать связывание (то есть, гибридизацию) полностью гомологичной последовательности с последовательностью-мишенью в условиях низкой жесткости. Это не означает, что условия низкой жесткости допускают неспецифическое связывание; то есть, условия низкой жесткости требуют, чтобы связывание двух последовательностей друг с другом было специфическим (то есть селективным). Отсутствие неспецифического связывания может быть подтверждено с использованием второй последовательности-мишени, в которой отсутствует даже частичная степень комплементарности (например, менее чем приблизительно 30% идентичности); причем, в отсутствие неспецифического связывания, зонд не будет гибридизоваться со второй некомплементарной мишенью.

[0135] Условия низкой жесткости включают условия, эквивалентные связыванию или гибридизации при 68°C в растворе, состоящем из 5 × SSPE (43,8 г/л NaCl, 6,9 г/л NaH₂PO₄ • H₂O и 1,85 г/л EDTA, pH был доведен до 7,4 с использованием NaOH), 0,1% ДСН, 5 × реагент Денхардта (50 × Денхардта содержит на 500 мл: 5 г фикола (типа 400, Pharmacia), 5 г BSA (фракция V; Sigma)) и 100 мкг/мл денатурированной ДНК спермы лосося с последующей промывкой в растворе, содержащем 2,0 × SSPE, 0,1% ДСН, при комнатной температуре, если используется зонд длиной приблизительно от 100 до приблизительно 1000 нуклеотидов.

[0136] Кроме того, условия высокой жесткости, которые способствуют гибридизации (например, повышение температуры стадий гибридизации и/или промывки, использование формамида в растворе для гибридизации и т.п.) хорошо известны специалистам. Условия высокой жесткости применительно к гибридизации нуклеиновых кислот включают условия, эквивалентные условиям связыванию или гибридизации при 68°C, в растворе, состоящем из 5 × SSPE, 1% ДСН, 5 × реагента Денхардта и 100 мкг/мл денатурированной ДНК спермы лосося с последующей промывкой в растворе, содержащем 0,1 × SSPE и 0,1% ДСН при 68°C, если используется зонд длиной приблизительно от 100 до приблизительно 1000 нуклеотидов.

[0137] Специалистам в данной области хорошо известно, что для установления

условий низкой жесткости может быть использовано множество эквивалентных факторов, таких, как длина и природа (ДНК, РНК, состав оснований) зонда и природа мишени (ДНК, РНК, состав оснований, присутствующих в растворе или являющихся иммобилизованными и т.п.) и концентрация солей и других компонентов (например, присутствие или отсутствие формамида, сульфата декстрана, полиэтиленгликоля), а также компонентов раствора для гибридизации, которые могут варьироваться для создания условий гибридизации низкой жесткости, отличающихся от вышеперечисленных условий, но эквивалентных им.

[0138] Термин «эквивалент», если он используется применительно к условиям гибридизации, поскольку он относится к представляющему интерес условию гибридизации, означает, что условие гибридизации и представляющие интерес условия гибридизации позволяют осуществлять гибридизацию последовательностей нуклеиновых кислот, которые имеют одинаковые пределы процента (%) гомологии. Так, например, если представляющие интерес условия гибридизации позволяют осуществлять гибридизацию первой последовательности нуклеиновой кислоты с другими последовательностями нуклеиновой кислоты, которые на 50% - 70% гомологичны первой последовательности нуклеиновой кислоты, то такое другое условие гибридизации будет рассматриваться как эквивалентное представляющему интерес условию гибридизации, если это другое условие гибридизации также будет приводить к гибридизации первой последовательности нуклеиновой кислоты с другими последовательностями нуклеиновой кислоты, которые на 50-70% гомологичны первой последовательности нуклеиновой кислоты.

[0139] Используемый здесь термин «гибридизация» относится к спариванию комплементарных нуклеиновых кислот по любому механизму, при котором цепь нуклеиновой кислоты связывается с комплементарной цепью посредством спаривания оснований с образованием гибридного комплекса. На гибридизацию и силу гибридизации (то есть, на силу ассоциации между нуклеиновыми кислотами) влияют такие факторы, как степень комплементарности между нуклеиновыми кислотами, жесткость условий процесса гибридизации, T_m образовавшегося гибрида и отношение G:C в нуклеиновых кислотах.

[0140] Используемый здесь термин «гибридный комплекс» означает комплекс, образованный между двумя последовательностями нуклеиновых кислот в результате образования водородных связей между комплементарными основаниями G и C и между комплементарными основаниями A и T, где эти водородные связи могут быть дополнительно стабилизированы посредством стэкинг-взаимодействий между основаниями. Водородные связи двух комплементарных последовательностей нуклеиновых кислот имеют антипараллельную конфигурацию. Гибридный комплекс может быть образован в растворе (например, как определено в анализе Cot или Rot) или между одной последовательностью нуклеиновой кислоты, присутствующей в растворе, и другой последовательностью нуклеиновой кислоты, иммобилизованной на твердом носителе (например, на найлоновой мембране или на нитроцеллюлозном фильтре,

используемых в Саузерн- и Нозерн-блот-анализах и в дот-блот-анализе или на предметном стекле, используемом для гибридизация *in situ*, включая FISH (флуоресцентная гибридизация *in situ*)).

[0141] Используемый здесь термин « T_m » означает «температуру плавления». Температура плавления представляет собой температуру, при которой группа двухцепочечных молекул нуклеиновой кислоты наполовину диссоциирует на отдельные цепи. Уравнение для определения T_m нуклеиновых кислот хорошо известно специалистам в данной области. Как указано в литературе, величина T_m может быть просто вычислена по уравнению: $T_m = 81,5 + 0,41 (\%G+C)$, если нуклеиновая кислота присутствует в водном растворе при 1M NaCl (см., например, Anderson and Young, Quantitative Filter Hybridization, in Nucleic Acid Hybridization, 1985). В других документах описаны более сложные вычисления T_m , которые проводят с учетом структурных свойств, а также свойств последовательностей.

[0142] Используемый здесь термин «жесткость» относится к температурным условиям, ионной силе и присутствию других соединений, таких как органические растворители, при которых проводят гибридизацию нуклеиновых кислот. «Жесткость» обычно соответствует температуре в пределах приблизительно от $T_m - 5^\circ\text{C}$ (на 5°C ниже температуры плавления зонда) до приблизительно $20-25^\circ\text{C}$ ниже T_m . Как будет очевидно специалистам в данной области, жесткая гибридизация может быть применена для идентификации или детектирования идентичных полинуклеотидных последовательностей или для идентификации или детектирования сходных или родственных полинуклеотидных последовательностей.

[0143] Термины «специфическое связывание», «специфичность связывания» и их грамматические эквиваленты, если они относятся к связыванию первой нуклеотидной последовательности со второй нуклеотидной последовательностью, означают преимущественно взаимодействие первой нуклеотидной последовательности со второй нуклеотидной последовательностью по сравнению со взаимодействием второй нуклеотидной последовательности с третьей нуклеотидной последовательностью. Специфическое связывание является относительным термином, и не означает абсолютную специфичность связывания; другими словами, термин «специфическое связывание» не означает, что вторая нуклеотидная последовательность будет взаимодействовать с первой нуклеотидной последовательностью в отсутствие взаимодействия между второй нуклеотидной последовательностью и третьей нуклеотидной последовательностью. Скорее, достаточно, чтобы уровень взаимодействия между первой нуклеотидной последовательностью и второй нуклеотидной последовательностью был выше, чем уровень взаимодействия между второй нуклеотидной последовательностью и третьей нуклеотидной последовательностью. «Специфическое связывание» первой нуклеотидной последовательности со второй нуклеотидной последовательностью также означает, что взаимодействие между первой нуклеотидной последовательностью и второй нуклеотидной последовательностью зависит от наличия конкретной структуры на первой

нуклеотидной последовательности или внутри нее; другими словами, вторая нуклеотидная последовательность распознает и связывается со специфической структурой на первой нуклеотидной последовательности или внутри нее, но не на всех нуклеиновых кислотах или нуклеотидных последовательностях. Так, например, если вторая нуклеотидная последовательность является специфичной для структуры «А», которая находится на первой нуклеотидной последовательности или внутри нее, то присутствие третьей последовательности нуклеиновой кислоты, содержащей структуру А, будет уменьшать уровень связывания второй нуклеотидной последовательности с первой нуклеотидной последовательностью.

[0144] Используемый здесь термин «амплифицируемая нуклеиновая кислота» означает нуклеиновые кислоты, которые могут быть амплифицированы любым методом амплификации. При этом предполагается, что «амплифицируемая нуклеиновая кислота» обычно охватывает «матрицу в качестве образца».

[0145] Термины «гетерологичная последовательность нуклеиновой кислоты» или «гетерологичная ДНК» являются синонимами и употребляются для обозначения нуклеотидной последовательности, которая лигирована с последовательностью нуклеиновой кислоты, с которой она не лигирована по природе, или с которой она лигирована в природе, но в другом положении. Гетерологичная ДНК не является эндогенной для клетки, в которую она была введена, поскольку она была получена из другой клетки. Обычно, хотя и не обязательно, такая гетерологичная ДНК кодирует РНК и белки, которые обычно не продуцируются клеткой, в которой она экспрессируется. Примеры гетерологичной ДНК включают репортерные гены, последовательности регуляции транскрипции и трансляции, селективные маркерные белки (например, белки, которые сообщают резистентность к лекарственному средству) и т.п.

[0146] «Амплификацию» определяют как получение дополнительных копий последовательности нуклеиновой кислоты и обычно осуществляют с применением технологий полимеразной цепной реакции, хорошо известной специалистам в данной области (Dieffenbach, C. W. and G. S. Dveksler (1995) PCR Primer, a Laboratory Manual, Cold Spring Harbor Press, Plainview, N.Y.). Используемый здесь термин «полимеразная цепная реакция» («ПЦР») означает способ, описанный в патентах США К. Mullis. Nos. 4683195 и 4683202, которые включены в настоящее описание посредством ссылки, и в которых описан способ увеличения концентрации сегмента последовательности-мишени в смеси геномной ДНК без клонирования или очистки. Длину амплифицированного сегмента желаемой последовательности-мишени определяют по относительным положениям двух олигонуклеотидных праймеров относительно друг друга, и, следовательно, эта длина является регулируемым параметром. Благодаря повторяющемуся действию такого процесса, этот метод называется «полимеразной цепной реакцией» («ПЦР»). Поскольку желаемые амплифицированные сегменты последовательности-мишени становятся преобладающими последовательностями (с точки зрения концентрации) в смеси, то они называются «ПЦР-амплифицированными».

[0147] С помощью ПЦР можно амплифицировать одну копию определенной последовательности-мишени в геномной ДНК до уровня, определяемого несколькими различными методами (например, путем гибридизация с меченым зондом; включения биотинилированных праймеров с последующим детектированием конъюгата авидин-фермент, включения ^{32}P -меченных дезоксирибонуклеотид-трифосфатов, таких как dCTP или dATP, в амплифицированный сегмент). Помимо геномной ДНК, любая олигонуклеотидная последовательность может быть амплифицирована с использованием соответствующего набора молекул праймеров. В частности, амплифицированные сегменты, полученные посредством самой ПЦР, сами по себе являются эффективными матрицами для последующей ПЦР-амплификации.

[0148] Один из таких предпочтительных методов, особенно для коммерческих целей, основан на широко применяемой технологии ПЦР в реальном времени TaqMan® и сочетает в себе аллель-специфическую ПЦР с блокирующим реагентом (ASB-ПЦР) для подавления амплификации аллеля дикого типа. ASB-ПЦР может быть проведена для детектирования зародышевой линии или соматических мутаций в ДНК или РНК, экстрагированных из ткани любого типа, включая образцы опухолей, залитых в парафин, фиксированный формалином. Была разработан ряд правил конструирования реагентов, позволяющих проводить чувствительное и избирательное детектирование одиночных точковых замен, инсерций или делеций по сравнению с фоном аллеля дикого типа в тысячекратном избытке или более (Morlan J, Baker J, Sinicropi D Mutation Detection by Real-Time PCR: A Simple, Robust and Highly Selective Method. PLoS ONE 4(2): e4584, 2009).

[0149] Термины «полимеразная цепная реакция с обратной транскриптазой» и «ОТ-ПЦР» означают метод обратной транскрипции последовательности РНК с получением смеси последовательностей кДНК с последующим увеличением концентрации желаемого сегмента транскрибированных последовательностей кДНК в смеси без клонирования или очистки. Обычно, РНК подвергают обратной транскрипции с использованием одного праймера (например, олиго-dT-праймера) перед ПЦР-амплификацией желаемого сегмента транскрибируемой ДНК с использованием двух праймеров.

[0150] Используемый здесь термин «праймер» означает олигонуклеотид, который независимо от того, встречается ли он в природе в виде очищенного рестрикционного гидролизата или был получен путем синтеза, способен действовать в качестве точки инициации синтеза, если он находится в условиях, при которых индуцируется синтез продукта удлинения праймера, комплементарного цепи нуклеиновой кислоты (то есть, в присутствии нуклеотидов и индуцирующего агента, такого как ДНК-полимераза, и при подходящей температуре и рН). В некоторых вариантах осуществления изобретения, праймер является одноцепочечным, что дает максимально эффективную амплификацию, но альтернативно, он может быть двухцепочечным. Праймер, если он является двухцепочечным, сначала обрабатывают для разделения цепей, а затем используют для получения продуктов удлинения. В некоторых вариантах осуществления изобретения,

праймер представляет собой олигодезоксирибонуклеотид. Праймер должен иметь длину, достаточную для инициации синтеза продуктов удлинения в присутствии индуцирующего агента. Точная длина праймеров зависит от многих факторов, включая температуру, источник праймера и применяемый метод.

[0151] Используемый здесь термин «зонд» означает олигонуклеотид (то есть, последовательность нуклеотидов), который, независимо от того, встречается ли он в природе в виде очищенного рестрикционного гидролизата, или был получен путем синтеза, рекомбинантным методом или путем ПЦР-амплификации, способен гибридизоваться с другим представляющим интерес олигонуклеотидом. Зонд может быть одноцепочечным или двухцепочечным. Зонды являются подходящими для детектирования, идентификации и выделения конкретных последовательностей генов. При этом предполагается, что любой зонд, используемый в настоящем изобретении, будет помечен любой «репортерной молекулой», так, чтобы он мог быть детектирован в любой детектирующей системе, включая, но не ограничиваясь ими, фермент (например, в ELISA, а также в гистохимических анализах на основе фермента), флуоресцентные, радиоактивные и люминесцентные системы. При этом предполагается, что настоящее раскрытие не будет ограничено какой-либо конкретной системой детектирования или меткой.

[0152] Используемые здесь термины «рестриктирующие эндонуклеазы» и «рестриктирующие ферменты» означают бактериальные ферменты, каждый из которых разрезает или обеспечивает одноцепочечный разрыв в двухцепочечной или одноцепочечной ДНК в определенной нуклеотидной последовательности или рядом с ней, например, в эндонуклеазном домене рестриктирующей эндонуклеазы типа PS (например, может быть использована FokI, как описано Kim et al., 1996, Proc. Nat'l. Acad. Sci. USA, 6:1 156-60).

[0153] Используемый здесь термин «олигонуклеотид, имеющий нуклеотидную последовательность, кодирующую ген», означает последовательность нуклеиновой кислоты, содержащую кодирующую область гена, то есть последовательность нуклеиновой кислоты, которая кодирует продукт гена. Кодирующая область может присутствовать в форме кДНК, геномной ДНК или РНК. Олигонуклеотид, если он присутствует в форме ДНК, может быть одноцепочечным (то есть, иметь смысловую цепь) или двухцепочечным. Кроме того, «олигонуклеотид, имеющий нуклеотидную последовательность, кодирующую ген», может включать подходящие регуляторные элементы, такие как энхансеры, промоторы, стыки сплайсинга, сигналы полиаденилирования и т.п., если это необходимо для обеспечения надлежащей инициации транскрипции и/или правильного процессинга первичного РНК-транскрипта. Кроме того, кодирующая область согласно изобретению может содержать эндогенные энхансеры, стыки сплайсинга, промежуточные последовательности, сигналы полиаденилирования и т.п.

[0154] Сигналы регуляции транскрипции у эукариотов содержат «энхансерные»

элементы. Эхансеры состоят из коротких массивов последовательностей ДНК, которые специфически взаимодействуют с клеточными белками, участвующими в транскрипции (Maniatis, T. et al., Science 236: 1237, 1987). Эхансерные элементы были выделены из различных эукариотических источников, включая гены в клетках растений, дрожжей, насекомых и млекопитающих, и вирусы. Выбор конкретного эхансера зависит от того, какой тип клеток должен использоваться для экспрессии представляющего интерес белка.

[0155] Наличие «сигналов сплайсинга» на экспрессионном векторе часто приводит к более высоким уровням экспрессии рекомбинантного транскрипта. Сигналы сплайсинга опосредуют удаление интронов из первичного РНК-транскрипта и состоят из донорного и акцепторного сайта сплайсинга (Sambrook, J. et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York, pp. 16.7-16.8, 1989). Обычно используемый донорный и акцепторный сайт сплайсинга представляет собой стык сплайсинга из РНК 16S SV40.

[0156] Эффективная экспрессия рекомбинантных последовательностей ДНК в эукариотических клетках требует экспрессии сигналов, регулирующих эффективную терминацию и полиаденилирование полученного транскрипта. Сигналы терминации транскрипции обычно находятся ниже сигнала полиаденилирования и имеют длину в несколько сотен нуклеотидов. Используемый здесь термин «поли-А-сайт» или «поли-А-последовательность» означает последовательность ДНК, которая регулирует терминацию и полиаденилирование растущей цепи РНК-транскрипта. Эффективное полиаденилирование рекомбинантного транскрипта является желательным, поскольку транскрипты, не имеющие поли-А-хвоста, являются нестабильными и быстро разрушаются. Поли-А-сигнал, используемый в экспрессионном векторе, может быть «гетерологичным» или «эндогенным». Эндогенный поли-А-сигнал представляет собой сигнал, который по своей природе присутствует у 3'-конца кодирующей области данного гена в геноме. Гетерологичный поли-А-сигнал представляет собой сигнал, который был выделен из одного гена и помещен в 3'-конец другого гена.

[0157] Используемый здесь термин «промотор», «промоторный элемент» или «промоторная последовательность» означает последовательность ДНК, которая, если она присутствует на 5'-конце (то есть, предшествует) олигонуклеотидной последовательности, способна регулировать транскрипцию олигонуклеотидной последовательности в мРНК. Промотор обычно расположен у 5'-конца (то есть, выше) олигонуклеотидной последовательности и регулирует ее транскрипцию в мРНК, а также обеспечивает сайт специфического связывания под действием РНК-полимеразы и сайт инициации транскрипции.

[0158] Термин «промоторная активность», если он используется в отношении последовательности нуклеиновой кислоты, означает способность последовательности нуклеиновой кислоты иницировать транскрипцию олигонуклеотидной последовательности в мРНК.

[0159] Термин «тканеспецифичный», применительно к промотору, относится к

промотору, который может регулировать селективную экспрессию олигонуклеотидной последовательности в ткани конкретного типа при относительном отсутствии экспрессии того же самого олигонуклеотида в ткани другого типа. Тканеспецифичность промотора может быть оценена, например, путем функционального связывания репортерного гена с промоторной последовательностью с получением репортерной конструкции; путем введения репортерной конструкции в геном растения или животного так, чтобы эта репортерная конструкция интегрировалась в каждую ткань полученного трансгенного животного, и путем детектирования экспрессии репортерного гена (например, детектирования мРНК, белка или активности белка, кодируемого репортерным геном) в различных тканях трансгенного растения или животного. Селективность не обязательно должна быть абсолютной. Детектирование более высокого уровня экспрессии репортерного гена в одной или более тканях по сравнению с уровнем экспрессии репортерного гена в других тканях показало, что промотор является специфичным для тканей, в которых обнаружены более высокие уровни экспрессии.

[0160] Термин «специфичный для клеток конкретного типа», применительно к промотору, относится к промотору, который способен регулировать селективную экспрессию олигонуклеотидной последовательности в клетке определенного типа при относительном отсутствии экспрессии той же самой олигонуклеотидной последовательности в клетке другого типа в той же ткани. Термин «специфичный для клеток определенного типа», применительно к промотору, также означает промотор, способный стимулировать селективную экспрессию олигонуклеотида в области одной ткани. И в этом случае, селективность необязательно должна быть абсолютной. Специфичность промотора к клетке определенного типа может быть оценена хорошо известными методами, например, посредством описанного здесь иммуногистохимического окрашивания. Вкратце, срезы ткани заливают в парафин, и срезы парафина подвергают реакции взаимодействия с первым антителом, которое является специфичным для полипептидного продукта, кодируемого олигонуклеотидной последовательностью, экспрессия которой регулируется промотором. В качестве альтернативы получению парафиновых срезов, образцы могут быть получены в виде криосрезов. Так, например, срезы могут быть заморожены до и во время получения срезов, что позволяет избежать потенциального влияния остаточного парафина. Меченное (например, конъюгированное с пероксидазой) второе антитело, которое является специфичным для первого антитела, приготавливают так, чтобы оно связывалось со срезами ткани, и детектировало специфическое связывание (например, с авидином/биотином) под микроскопом.

[0161] Термины «селективная экспрессия», «селективно экспрессирует» и их грамматические эквиваленты относятся к сравнению относительных уровней экспрессии в двух или более представляющих интерес областях. Так, например, термин «селективная экспрессия», применительно к тканями, относится к значительно более высокому уровню экспрессии представляющего интерес гена в конкретной ткани или к значительно

большому числу клеток, которые экспрессируют ген в этой ткани, по сравнению, соответственно, с уровнем экспрессии и числом клеток, экспрессирующих один и тот же ген в другой ткани (то есть, селективность необязательно должна быть абсолютной). Селективная экспрессия не требует, хотя и может включать, экспрессии представляющего интерес гена в конкретной ткани и полное отсутствие экспрессии того же гена в другой ткани. Аналогичным образом, «селективная экспрессия», используемая здесь применительно к типам клеток, относится к значительно более высокому уровню экспрессии представляющего интерес гена в конкретной ткани или к значительно большему числу клеток, которые экспрессируют представляющий интерес ген в клетках конкретного типа, по сравнению, соответственно, с уровнями экспрессии гена и количеством клеток, экспрессирующих ген в клетках другого типа.

[0162] Термин «смежный», если он используется здесь в отношении двух или более нуклеотидных последовательностей, означает, что нуклеотидные последовательности лигированы тандемно либо в отсутствие промежуточных последовательностей, либо в присутствии промежуточных последовательностей, которые не содержат один или более регуляторных элементов.

[0163] Используемые здесь термины «кодирующая молекула нуклеиновой кислоты», «кодирующий нуклеотид», «кодирующая последовательность ДНК» и «кодирующая ДНК» относятся к порядку или к последовательности расположения дезоксирибонуклеотидов вдоль цепи дезоксирибонуклеиновой кислоты. Порядок этих дезоксирибонуклеотидов определяет порядок аминокислот вдоль полипептидной (белковой) цепи. Таким образом, последовательность ДНК кодирует аминокислотную последовательность.

[0164] Термин «выделенный», если он используется по отношению к нуклеиновой кислоте, как и термин «выделенный олигонуклеотид», относится к последовательности нуклеиновой кислоты, которая была отделена по меньшей мере от одной примесной нуклеиновой кислоты, с которой она обычно связана в своем природном источнике. Выделенная нуклеиновая кислота представляет собой нуклеиновую кислоту, присутствующую в форме или в окружении, отличающихся от их природной формы или окружения. В противоположность этому, невыделенные нуклеиновые кислоты представляют собой нуклеиновые кислоты, такие как ДНК и РНК, находящиеся в состоянии, в котором они существуют в природе. Так, например, данная последовательность ДНК (например, ген) присутствует на хромосоме клетки-хозяина в непосредственной близости от соседних генов; последовательности РНК, такие как специфическая последовательность мРНК, кодирующая специфический белок, обнаруживаются в клетке в виде смеси с множеством других мРНК, которые кодируют множество белков. Однако, выделенная нуклеиновая кислота, кодирующая представляющий интерес полипептид, включает, например, такую нуклеиновую кислоту в клетках, которые обычно экспрессируют представляющий интерес полипептид, где нуклеиновая кислота находится в хромосомном или во внехромосомном положении,

отличающимся от такового в природных клетках, или как либо иначе фланкированы другой последовательностью нуклеиновой кислоты, чем это наблюдается в природе. Выделенная нуклеиновая кислота или олигонуклеотид могут присутствовать в одноцепочечной или двухцепочечной форме. Выделенная нуклеиновая кислота может быть легко идентифицирована (при желании) различными способами (например, путем гибридизации, дот-блот-анализа и т.п.). Если выделенная нуклеиновая кислота или олигонуклеотид используются для экспрессии белка, то олигонуклеотид будет содержать как минимум смысловую или кодирующую цепь (то есть, олигонуклеотид может быть одноцепочечным). Альтернативно, он может содержать как смысловые, так и антисмысловые цепи (то есть, олигонуклеотид может быть двухцепочечным).

[0165] Используемый здесь термин «очищенный» или «для очистки» относится к удалению одного или более (нежелательных) компонентов из образца. Так, например, если рекомбинантные полипептиды экспрессируются в бактериальных клетках-хозяевах, то полипептиды очищают путем удаления белков клетки-хозяина, увеличивая тем самым процент рекомбинантных полипептидов в образце.

[0166] Используемый здесь термин «по существу очищенный» относится к молекулам, как к последовательностям нуклеиновой кислоты, так и к аминокислотным последовательностям, которые были удалены из их природной среды, выделены или отделены и по меньшей мере на 60%, в некоторых вариантах на 75%, а в некоторых вариантах на 90% не содержат других компонентов, с которыми они связаны в природе. Следовательно, «выделенный полинуклеотид» представляет собой, по существу, очищенный полинуклеотид.

[0167] Используемый здесь термин «кодирующая область», применительно к структурному гену, означает нуклеотидные последовательности, которые кодируют аминокислоты, обнаруженные в растущем полипептиде в результате трансляции молекулы мРНК. Кодированная область у эукариотов ограничена со стороны 5'-конца обычно нуклеотидным триплетом «АТГ», который кодирует иницирующий метионин, а со стороны 3'-конца, одним из трех триплетов, которые определяют стоп-кодона (то есть, ТАА, TAG, TGA).

[0168] Под «кодирующей последовательностью» подразумевается последовательность нуклеиновой кислоты или ее комплемент или ее часть, которые могут транскрибироваться и/или транслироваться с образованием мРНК и/или полипептида или его фрагмента. Кодированные последовательности включают экзоны в геномной ДНК или незрелые первичные РНК-транскрипты, которые связываются друг с другом по биохимическому клеточному механизму с образованием зрелой мРНК. Антисмысловая цепь является комплементом такой нуклеиновой кислоты, и из нее может быть выведена кодирующая последовательность.

[0169] Под «некодирующей последовательностью» подразумевается последовательность нуклеиновой кислоты или ее комплемент или ее часть, которые не транскрибируются в аминокислоту *in vivo*, или последовательность тРНК не

взаимодействует так, чтобы она продуцировала или могла продуцировать аминокислоту. Некодирующие последовательности включают интронные последовательности в геномной ДНК или незрелые первичные РНК-транскрипты, а также геноассоциированные последовательности, такие как промоторы, энхансеры, сайленсеры и т.п.

[0170] Используемый здесь термин «структурный ген» или «структурная нуклеотидная последовательность» относится к последовательности ДНК, кодирующей РНК или белок, которая не регулирует экспрессию других генов. В противоположность этому, «регуляторный ген» или «регуляторная последовательность» представляют собой структурный ген, который кодирует продукты (например, факторы транскрипции), регулирующие экспрессию других генов.

[0171] Используемый здесь термин «регуляторный элемент» означает генетический элемент, который регулирует некоторые аспекты экспрессии последовательностей нуклеиновых кислот. Так, например, промотор является регуляторным элементом, который облегчает инициацию транскрипции функционально связанной кодирующей области. Другие регуляторные элементы включают сигналы сплайсинга, сигналы полиаденилирования, сигналы терминации и т.п.

[0172] Используемый здесь термин «сайт связывания с пептидным фактором транскрипции» или «сайт связывания с фактором транскрипции» означает нуклеотидную последовательность, которая связывается с белковыми факторами транскрипции и, таким образом, регулирует некоторые аспекты экспрессии последовательностей нуклеиновых кислот. Так, например, сайты связывания с Sp-1 и AP1 (белком-активатором 1) являются примерами сайтов связывания с пептидным фактором транскрипции.

[0173] Используемый здесь термин «ген» означает дезоксирибонуклеотидные последовательности, содержащие кодирующую область структурного гена. «Ген» также может включать нетранслируемые последовательности, расположенные рядом с кодирующей областью как на 5'-, так и на 3'-концах, так что этот ген будет соответствовать длине полноразмерной мРНК. Последовательности, которые расположены у 5'-конца кодирующей области, и которые присутствуют в мРНК, называются 5'-нетранслируемыми последовательностями. Последовательности, которые расположены у 3'-конца или ниже кодирующей области, и которые присутствуют в мРНК, называются в 3'-нетранслируемыми последовательностями. Термин «ген» охватывает как кДНК, так и геномные формы гена. Геномная форма или клон гена содержит кодирующую область, прерываемую некодирующими последовательностями, называемыми «интронами» или «промежуточными областями» или «промежуточными последовательностями». Интроны представляют собой сегменты гена, которые транскрибируются в гетерогенную ядерную РНК (гяРНК), и эти интроны могут содержать регуляторные элементы, такие как энхансеры. Интроны удаляются или «вырезаются» из ядерного или первичного транскрипта, а поэтому интроны отсутствуют в транскрипте матричной РНК (мРНК). Во время трансляции, функции мРНК заключаются в определении последовательности или порядка расположения аминокислот в растущем

полипептиде.

[0174] Геномные формы гена, помимо интронов, могут также включать последовательности, расположенные как у 5'- и 3'-конца последовательностей, которые присутствуют в транскрипте РНК. Эти последовательности называются «фланкирующими» последовательностями или областями (эти фланкирующие последовательности расположены у 5'- и 3'-конца по отношению к нетранслируемым последовательностям, присутствующим на мРНК-транскрипте). 5'-фланкирующая область может содержать регуляторные последовательности, такие как промоторы и энхансеры, которые регулируют или влияют на транскрипцию гена. 3'-фланкирующая область может содержать последовательности, которые регулируют терминацию транскрипции, посттранскрипционное расщепление и полиаденилирование.

[0175] «Животное, не являющееся человеком» означает любое животное кроме человека и включает позвоночных, таких как грызуны, приматы, не являющиеся человеком, овцы, коровы, жвачные животные, зайцеобразные, свиньи, козы, лошади, собаки, кошки, птицы и т.п. Предпочтительные животные, не являющиеся человеком, выбраны из отряда грызунов (Rodentia). Термин «животное, не являющееся человеком» также включает амфибии (например, *Xenopus*), рептилии, насекомых (например, дрозофил) и животных других видов, не относящихся к млекопитающим.

[0176] Используемый здесь термин «трансгенный» относится к организму или к клетке, в которую встроена ДНК, происходящая от другого организма, которая интегрируется в геном соматических и/или зародышевых клеток растения или животного. «Трансген» означает последовательность ДНК, которая является частично или полностью гетерологичной (то есть, не присутствует в природе) для растения или животного, в котором она обнаружена, или которая является гомологичной эндогенной последовательности (то есть, последовательности, которая обнаруживается у животного в природе) и встраивается в геном растения или животного в положении, отличающимся от такового в природной последовательности. Трансгенные растения или животные, которые включают один или более трансгенов, входят в объем настоящего изобретения. Кроме того, используемый здесь термин «трансгенный» относится к организму, который имеет один или более генов, модифицированных и/или «нокаутированных» (лишенных функциональности или функционирующих на пониженном уровне, например, в результате «нокаут»-мутации) с применением описанных здесь способов, путем гомологичной рекомбинации, мутации ТФО или аналогичными способами. Так, например, в некоторых вариантах осуществления изобретения, трансгенный организм или клетка включает встроенную ДНК, которая содержит чужеродный промотор и/или кодирующую область.

[0177] «Трансформированная клетка» представляет собой клетку или клеточную линию, которая приобрела способность расти в клеточной культуре на протяжении нескольких поколений, способность расти в мягком агаре и/или способность не препятствовать росту клеток посредством контактирования клеток. В связи с этим,

трансформация означает введение чужеродного генетического материала в клетку или в организм. Трансформация может быть осуществлена любым известным способом, который позволяет успешно вводить нуклеиновые кислоты в клетки, и который приводит к экспрессии введенной нуклеиновой кислоты. Термин «трансформация» включает, но не ограничивается ими, такие методы, как трансфекция, микроинъекция, электропорация, нуклеофекция и липофекция (перенос генов, опосредуемый липосомами). Трансформация может быть осуществлена с использованием любого экспрессионного вектора. Так, например, рассматривается применение бакуловируса для введения чужеродной нуклеиновой кислоты в клетки насекомых. Термин «трансформация» также включает такие методы, как трансформация зародышевой линии целых насекомых, опосредуемая Р-элементом. Кроме того, трансформация относится к клеткам, которые были трансформированы естественным путем, обычно посредством генетической мутации.

[0178] Используемый здесь термин «экзогенный» означает, что ген, кодирующий белок, обычно не экспрессируется в клетке. Кроме того, термин «экзогенный» относится к гену, введенному в клетку для повышения уровня экспрессии этого гена по сравнению с нормальным уровнем (то есть, природным уровнем).

[0179] Пептидная последовательность и нуклеотидная последовательность могут быть «эндогенными» или «гетерологичными» (то есть «чужеродными»). Термин «эндогенный» относится к последовательности, которая обычно присутствует в клетке, в которую введена эта последовательность, при условии, что она не будет содержать каких-либо модификаций по сравнению с природной последовательностью. Термин «гетерологичный» относится к последовательности, которая не является эндогенной для клетки, в которую она была введена. Так, например, гетерологичная ДНК включает нуклеотидную последовательность, которая была лигирована или в результате манипуляций стала лигированной с последовательностью нуклеиновой кислоты, с которой она не была лигирована в природе, или с которой она лигирована в природе в другом участке. Гетерологичная ДНК также включает нуклеотидную последовательность, которая обычно присутствует в клетке, в которую она была введена, и которая содержит некоторую модификацию по сравнению с природной последовательностью. Обычно, хотя и необязательно, гетерологичная ДНК кодирует гетерологичную РНК и гетерологичные белки, которые обычно не продуцируются клеткой, в которую она они были введены. Примеры гетерологичной ДНК включают репортерные гены, последовательности регуляции транскрипции и трансляции, последовательности ДНК, которые кодируют селективные маркерные белки (например, белки, которые сообщают резистентность к лекарственному средству) и т.п.

[0180] В некоторых своих аспектах и вариантах, настоящее изобретение относится к способам введения мутации, опосредуемой олигонуклеосоединением для репарации генов (GRON), в последовательность дезоксирибонуклеиновой кислоты (ДНК)-мишени в клетке растения; например, в целях модификации опианного здесь гена SHP. В некоторых вариантах осуществления изобретения, способы могут включать, *inter alia*,

культивирование клетки растения в условиях, которые улучшают один или более процессов репарации клеточной ДНК до и/или во время доставки GRON в клетку растения, и/или доставку GRON длиной более 15 оснований в клетку растения, где GRON включает, но необязательно, один или более или два или более сайтов мутаций (таких как описанные здесь сайты мутаций SHP) для введения в ДНК-мишень.

[0181] Используемый здесь термин «олигонуклеотид для репарации генов» или «GRON» означает олигонуклеосооснование (например, смешанные дуплексные олигонуклеотиды; молекулы, содержащие не-нуклеотид; одноцепочечные олигодезоксинуклеотиды, двухцепочечные олигодезоксинуклеотиды и другие молекулы репарации генов), которые могут, в определенных условиях, регулировать одну, а в некоторых вариантах, множество нуклеотидных делеций, инсерций или замен в последовательности ДНК. Это опосредованное олигонуклеотидами редактирование генома посредством репарации генов может включать использование негомологичных систем репарации (например, негомологичное присоединение по концам), и гомологичных систем репарации (например, репарации, опосредуемой гомологией). GRON обычно конструируют так, чтобы он был пригоден для выравнивания с геномной мишенью, за исключением спланированного(ых) ошибочного(ых) спаривания(й). Эти ошибочные спаривания могут распознаваться и корректироваться посредством одной или более систем репарации эндогенной ДНК в клетке. В некоторых вариантах осуществления изобретения, GRON или олигонуклеотид могут быть сконструированы так, чтобы они содержали множественные различия по сравнению с последовательностью-мишенью организма. Не все эти различия могут влиять на последовательность белка, транслируемую из указанной последовательности-мишени, и в одном или более случаях, они могут быть известны как молчащие мутации. Многочисленные варианты структуры, химических свойств и функций GRON описаны в настоящей заявке. В различных вариантах осуществления изобретения, используемый здесь GRON может иметь одну или более модификаций. Так, например, используемый здесь GRON может иметь одну или более модификаций, которые иницируют механизмы репарации ДНК в целевом сайте (с ошибочными спариваниями) и/или которые предотвращают рекомбинацию части или всех GRON (кроме желаемых целевых делеций, инсерций, замен или т.п) в геномной последовательности ДНК-мишени, и/или которые повышают стабильность GRON.

[0182] В различных вариантах осуществления изобретения, GRON может содержать нуклеотиды РНК и ДНК и/или нуклеосооснования других типов. В некоторых вариантах осуществления изобретения, один или более нуклеотидов ДНК или РНК содержат модификацию.

[0183] В одном своем аспекте, настоящее изобретение относится к способу генетической модификации в клетке растения (например, генетической модификации в гене SHP), где указанный способ включает обработку клетки ДНК-разрезающим ферментом и GRON, например, GRON, который был модифицирован как рассматривается в настоящей заявке. В некоторых вариантах осуществления изобретения, GRON может

быть модифицирован группой Су3, группой ЗPS и 2'О-метильной группой, или он может иметь другие рассматриваемые здесь модификации. В другом своем аспекте, настоящее изобретение относится к клетке растения, которая включает ДНК-разрезающий фермент и GRON (такой как GRON, который связывается с геном SHP и/или модифицирует его), например, где GRON был модифицирован группой Су3, группой ЗPS и 2'О-метильной группой или другой модификацией. В некоторых вариантах осуществления изобретения, ДНК-разрезающий фермент представляет собой один и более ферментов, выбранных из CRISPR, которые включают, но не ограничиваются ими, Cas9, Cpf1 и их соответствующие гомологи, ортологи и/или паралоги, редактор оснований, TALEN, «цинковый палец», мегануклеазу и ДНК-разрезающий антибиотик. В некоторых вариантах осуществления изобретения, ДНК-разрезающим ферментом является CRISPR. В некоторых вариантах осуществления, ДНК-разрезающий фермент представляет собой TALEN. ДНК-разрезающий фермент может быть получен на основе плазмиды (ДНК), РНК и/или белка. В некоторых вариантах осуществления изобретения, длина GRON составляет от 15 до 60 нуклеоснований; или от 30 до 40 нуклеоснований; или от 35 до 45 нуклеоснований; или от 20 до 70 нуклеоснований; или от 20 до 200 нуклеоснований; или от 30 до 180 нуклеоснований; или от 50 до 160 нуклеоснований; или от 70 до 150 нуклеоснований; или от 80 до 120 нуклеоснований; или от 90 до 110 нуклеиновых оснований; или от 95 до 105 нуклеиновых оснований; или от 80 до 300 нуклеоснований; или от 90 до 250 нуклеоснований; или от 100 до 150 нуклеоснований; или от 100 до 300 нуклеоснований; или от 150 до 200 нуклеоснований; или от 200 до 300 нуклеоснований; или от 250 до 350 нуклеоснований; или от 50 до 110 нуклеиновых оснований; или от 50 до 200 нуклеоснований; или от 150 до 210 нуклеиновых оснований; или от 20 до 1000 нуклеоснований; или от 100 до 1000 нуклеоснований; или от 200 до 1000 нуклеоснований; или от 300 до 1000 нуклеоснований; или от 400 до 1000 нуклеоснований; или от 500 до 1000 нуклеоснований; или от 600 до 1000 нуклеоснований; или от 700 до 1000 нуклеоснований; или от 800 до 1000 нуклеоснований; или от 900 до 1000 нуклеоснований; или от 300 до 800 нуклеоснований; или от 400 до 600 нуклеоснований; или от 500 до 700 нуклеоснований; или от 600 до 800 нуклеоснований; или более, чем 30 нуклеоснований; или более, чем 35 нуклеоснований; или длиной более 40 нуклеоснований; или длиной более 50 нуклеоснований; или более, чем 60 нуклеоснований; или более, чем 65 нуклеоснований; или длиной более 70 нуклеоснований; или более, чем 75 нуклеоснований; или более, чем 80 нуклеоснований; или более, чем 85 нуклеоснований; или более, чем 90 нуклеоснований; или более, чем 95 нуклеоснований; или более, чем 100 нуклеоснований; или более, чем 110 нуклеоснований; или более, чем 125 нуклеоснований; или более, чем 150 нуклеоснований; или более, чем 165 нуклеотидов; или более, чем 175 нуклеоснований; или более, чем 200 нуклеоснований; или более, чем 250 нуклеоснований; или длиной более 300 нуклеоснований; или более, чем 350

нуклеосоманий; или более, чем 400 нуклеосоманий; или более, чем 450 нуклеосоманий; или более, чем 500 нуклеосоманий; или более, чем 550 нуклеосоманий; или более, чем 600 нуклеосоманий; или более, чем 700 нуклеосоманий; или более, чем 800 нуклеосоманий; или более, чем 900 нуклеосоманий.

[0184] GRON могут быть нацелены на некодирующие (NC) и на кодирующие (C) области гена-мишени.

[0185] Используемый здесь термин «CRISPR» означает элементы, то есть, ген *cas* (ассоциированный с CRISPR), транскрипт (например, мРНК) и/или белок и по меньшей мере одну спейсерную последовательность CRISPR (регулярно кластеризованные встроенные короткие палиндромные повторы, также известные как SPDR, то есть, прямые повторы, перемежающиеся спейсерами), которые, при их эффективном действии или экспрессии в клетке могут влиять на расщепление последовательности ДНК-мишени посредством клеточного механизма CRISPR/CAS, такого как механизм, описанный, например, в публикациях Cong, L. et al., *Science*, vol. 339, № 6121, pp. 819-823 (2013); Jinek et al., *Science*, vol. 337: 816-821 (2013); Wang et al., *RNA*, vol. 14, pp. 903-913 (2008); Zhang et al. *Plant Physiology*, vol. 161, pp. 20-27 (2013), Zhang et al., в заявке PCT № PCT/US2013/074743; и Charpentier et al., в заявке PCT № PCT/US2013/032589. В некоторых вариантах осуществления изобретения, таких как, например, CRISPR для использования в эукариотической клетке, CRISPR, рассматриваемый в настоящей заявке, может также содержать дополнительный элемент, который включает последовательность для одного или более функциональных сигналов локализации в ядре. Рассматриваемые здесь CRISPR могут экспрессироваться в клетке, могут быть введены в клетку любыми из множества способов и/или могут присутствовать в клетке (такой как клетка растения) в любых своих проявлениях. Так, например, рассматриваемый здесь CRISPR, может содержать или включать один или более CRISPR на плазмиде, никазу CRISPR на плазмиде, CRISPRa на плазмиде или CRISPRi на плазмиде, как описано ниже:

[0186] **CRISPR на плазмиде:** Рекомбинантный экспрессионный вектор, содержащий:

(i) нуклеотидную последовательность, кодирующую РНК, нацеленную на ДНК (например, руководящую РНК), где РНК, нацеленная на ДНК, содержит:

а. первый сегмент, содержащий нуклеотидную последовательность, которая комплементарна последовательности в ДНК-мишени (например, протоспейсер, спейсер или cr-РНК); и

б. второй сегмент, который взаимодействует с сайт-направленным модифицирующим полипептидом (например, транс-активирующей cr-РНК или tracr-РНК); и

(ii) нуклеотидную последовательность, кодирующую сайт-направленный модифицирующий полипептид (например, ген *cas*), где сайт-направленный полипептид содержит:

а. РНК-связывающую часть, которая взаимодействует с РНК, нацеленной на ДНК (например, на участке REC); и

б. участок активности, который вызывает двухцепочечные разрывы в ДНК-мишени (например, на участке NUC), где сайт двухцепочечных разрывов в ДНК-мишени определяют по РНК, нацеленной на ДНК.

[0187] **Никаза CRISPR на плазмиде.** Рекомбинантный экспрессионный вектор, содержащий:

(i) нуклеотидную последовательность, кодирующую РНК, нацеленную на ДНК (например, руководящую РНК), где РНК, нацеленная на ДНК, содержит:

а. первый сегмент, содержащий нуклеотидную последовательность, которая комплементарна последовательности в ДНК-мишени (например, протоспейсер, спейсер или cr-РНК); и

б. второй сегмент, который взаимодействует с сайт-направленным модифицирующим полипептидом (например, транс-активирующей cr-РНК или tracr-РНК); и

(ii) нуклеотидную последовательность, кодирующую сайт-направленный модифицирующий полипептид (например, ген cas), где сайт-направленный полипептид содержит:

а. РНК-связывающую часть, которая взаимодействует с РНК, нацеленной на ДНК (например, на участке REC); и

б. участок активности, который вызывает двухцепочечные разрывы в ДНК-мишени (например, на участке NUC), где сайт одноцепочечных разрывов в ДНК-мишени определяют по РНК, нацеленной на ДНК.

[0188] **CRISPRa на плазмиде.** Рекомбинантный экспрессионный вектор, содержащий:

(i) нуклеотидную последовательность, кодирующую РНК, нацеленную на ДНК (например, руководящую РНК), где РНК, нацеленная на ДНК, содержит:

а. первый сегмент, содержащий нуклеотидную последовательность, которая комплементарна последовательности в ДНК-мишени (например, протоспейсер, спейсер или cr-РНК); и

б. второй сегмент, который взаимодействует с сайт-направленным модифицирующим полипептидом (например, транс-активирующей cr-РНК или tracr-РНК); и

(ii) нуклеотидную последовательность, кодирующую сайт-направленный модифицирующий полипептид (например, ген cas), где сайт-направленный полипептид содержит:

а. РНК-связывающую часть, которая взаимодействует с РНК, нацеленной на ДНК (например, на участке REC); и

б. участок активности, который модулирует транскрипцию (например, на участке NUC; в некоторых вариантах осуществления изобретения, повышает транскрипцию) в

ДНК-мишени, где сайт модуляции транскрипции в ДНК-мишени определяют по РНК, нацеленной на ДНК.

[0189] **CRISPRi на плазмиде.** Рекомбинантный экспрессионный вектор, содержащий:

(i) нуклеотидную последовательность, кодирующую РНК, нацеленную на ДНК (например, руководящую РНК), где РНК, нацеленная на ДНК, содержит:

a. первый сегмент, содержащий нуклеотидную последовательность, которая комплементарна последовательности в ДНК-мишени (например, протоспейсер, спейсер или cr-РНК); и

b. второй сегмент, который взаимодействует с сайт-направленным модифицирующим полипептидом (например, транс-активирующей cr-РНК или tracr-РНК); и

(ii) нуклеотидную последовательность, кодирующую сайт-направленный модифицирующий полипептид (например, ген cas), где сайт-направленный полипептид содержит:

a. РНК-связывающую часть, которая взаимодействует с РНК, нацеленной на ДНК (например, на участке REC); и

b. участок активности, который модулирует транскрипцию/трансляцию (например, на участке NUC; в некоторых вариантах осуществления изобретения, повышает транскрипцию/трансляцию) в ДНК-мишени, где сайт модуляции транскрипции/трансляции в ДНК-мишени определяют по РНК, нацеленной на ДНК.

[0190] Каждый из CRISPR на плазмиде, никазы CRISPR на плазмиде, CRISPRa на плазмиде и CRISPRi на плазмиде могут, в некоторых вариантах осуществления изобретения, альтернативно иметь один или более соответствующих элементов, которые могут быть введены, могут экспрессироваться или могут присутствовать в клетке в качестве РНК (например, мРНК) или белка, но не на плазмиде. Доставка защищенной мРНК может быть такой, как это описано Kariko et al., в патенте США № 8278036.

[0191] В некоторых вариантах осуществления изобретения, каждый из CRISPRi и CRISPRa может включать дезактивированный cas9 (dCas9). Дезактивированный cas9 все еще связывается с ДНК-мишенью, но не обладает разрезающей активностью. Дефицитный по нуклеазе Cas9 может быть результатом точковых мутаций D10A и H840A, которые инактивируют два его каталитических домена.

[0192] В некоторых вариантах осуществления изобретения, CRISPRi ингибирует инициацию транскрипции или элонгацию посредством стерического затруднения РНК-полимеразы II. CRISPRi может быть, но необязательно, активирован (CRISPRei) посредством присоединения домена сильного репрессора к С-концу белка dCas9. В некоторых вариантах осуществления изобретения, домен-репрессор осуществляет рекрутиг модификаторов хроматина и использует их. В некоторых вариантах осуществления изобретения, домен-репрессор может включать, но не ограничивается ими, домены, описанные в публикации Kagale, S. et al., Epigenetics, vol. 6 № 2, pp. 141-146

(2011).

1. LDLNRPPPVEN - домен-репрессор OsERF3 (мотив LxLxPP) (SEQ ID NO: 17)
2. LRLFGVNM - домен-репрессор AtBRD (мотив R/KLFGV) (SEQ ID NO: 18)
3. LKLFGVWL - домен-репрессор AtHsfB1 (мотив R/KLFGV) (SEQ ID NO: 19)
4. LDLELRLGFA - домен-репрессор AtSUP (мотив EAR) (SEQ ID NO: 20)

5. ERSNSIELRNSFYGRARTSPWSYGDYDNCQQDHDYLLGFSWPPRSYTCSFCKR
EFRSAQALGGHMNVHRRDRARLRLQQSPSSSSTPSPYPNPNYSYSTMANSPPPHHSPLT
LFPTLSPSSPRYRAGLIRSLSPKSKHTPENACKTKKSSLLVEAGEATRFTSKDACKILRN
DEISLELEIGLINESEQDLLELRLGFA* - полноразмерный ген AtSUP, содержащий
домен-репрессор (мотив EAR) (SEQ ID NO: 21).

[0193] В некоторых вариантах осуществления изобретения, активация транскрипции CRISPRa достигается под действием белка dCas9, содержащего активатор транскрипции, присоединенный к С-концу. В некоторых вариантах осуществления изобретения, активация может включать, но не ограничивается ими, VP64 (4X VP16), домен активации AtERF98 или конкатемеры AtERF98×4, описанные в публикациях Cheng, AW et al., Cell Research, pp1-9 (2013); Perez-Pinera, P. et al., Nature Methods, vol. 10 pp. 913-976 (2013); Maeder, ML. et al., Nature Methods, vol. 10 pp. 977-979 (2013) и Mali, P., et al., Nature Biotech., Vol. 31 pp. 833-838 (2013).

[0194] В некоторых вариантах осуществления изобретения, CRISPR включает никазу. В некоторых вариантах осуществления изобретения используются две или более никаз CRISPR. В некоторых вариантах осуществления изобретения, две или более никаз осуществляют разрезание на противоположных цепях нуклеиновой кислоты-мишени. В других вариантах осуществления изобретения, две или более никаз осуществляют разрезание на одной и той же цепи нуклеиновой кислоты-мишени.

[0195] Используемый здесь термин «белок-репрессор» или «репрессор» означает белок, который связывается с оператором ДНК или РНК и предотвращает транскрипцию или трансляцию, соответственно.

[0196] Используемый здесь термин «репрессия» относится к ингибированию транскрипции или трансляции посредством связывания белка-репрессора со специфическим сайтом в ДНК или мРНК. В некоторых вариантах осуществления изобретения, репрессия включает значительное изменение уровня транскрипции или трансляции по меньшей мере в 1,5 раза, в других вариантах, по меньшей мере в два раза, а в других вариантах по меньшей мере в пять раз.

[0197] Используемый здесь термин «белок-активатор» или «активатор», если он относится к транскрипции и/или трансляции гена, означает белок, который связывается с оператором ДНК или РНК и усиливает или повышает транскрипцию или трансляцию, соответственно.

[0198] Используемый здесь термин «активация», если он относится к транскрипции и/или трансляции гена, означает усиление или повышение уровня транскрипции или трансляции посредством связывания белка-активатора со специфическим сайтом в ДНК

или мРНК. В некоторых вариантах осуществления изобретения, активация включает значительное изменение уровня транскрипции или трансляции по меньшей мере в 1,5 раза, в некоторых вариантах, по меньшей мере в два раза, а в других вариантах, по меньшей мере в пять раз.

[0199] В некоторых вариантах осуществления изобретения, условия, которые улучшают один или более процессов репарации клеточной ДНК, могут включать одну или более из таких стадий, как введение одного или более сайтов в GRON или в ДНК клетки растения, которые являются мишенями для репарации посредством вырезания основания; введение одного или более сайтов в GRON или в ДНК клетки растения, которые являются мишенями для негомологичного присоединения по концам; введение одного или более сайтов в GRON или в ДНК клетки растения, которые являются мишенями для присоединения по концам, опосредуемого микрогомологией; введение одного или более сайтов в GRON или в ДНК клетки растения, которые являются мишенями для гомологичной рекомбинации; и введение одного или нескольких сайтов в GRON или в ДНК клетки растения, которые являются мишенями для осуществления репарации (например, репарации посредством вырезания оснований (BER); репарации посредством гомологичной рекомбинации (HR); репарации посредством ошибочных спариваний (MMR); репарации посредством негомологичного присоединения по концам (NHEJ), которая включает классическую и альтернативную NHEJ; и репарацию посредством вырезания нуклеотидов (NER)).

[0200] Как описано в настоящей заявке, GRON, используемые в настоящем изобретении, могут включать одну или более из нижеследующих (неограничивающих) модификаций по сравнению со стандартными нуклеотидами РНК и ДНК:

один или более нуклеотидов, не содержащих нуклеотидных оснований;

один или более 8'-оксо-dA- и/или 8'-оксо-dG-нуклеотидов;

обратное основание у 3'-конца;

один или более 2'-О-метил-нуклеотидов;

один или более РНК-нуклеотидов;

один или более РНК-нуклеотидов у 5'-конца, а в некоторых вариантах осуществления изобретения, 2, 3, 4, 5, 6, 7, 8, 9, 10 или более нуклеотидов, где один или более РНК-нуклеотидов могут быть дополнительно модифицированы; один или более РНК-нуклеотидов у 3'-конца, а в некоторых вариантах осуществления изобретения, 2, 3, 4, 5, 6, 7, 8, 9, 10 или более нуклеотидов, где один или более РНК-нуклеотидов могут быть дополнительно модифицированы;

один или более 2'-О-метил-РНК-нуклеотидов у 5'-конца, а в некоторых вариантах осуществления изобретения, 2, 3, 4, 5, 6, 7, 8, 9, 10 или более нуклеотидов;

интеркалирующий краситель;

5'-концевой кэп;

модификацию остова, выбранную из группы, состоящей из модификации фосфотиоатом, модификации метилфосфонатом, модификации блокированной

нуклеиновой кислотой (LNA), модификации O-(2-метоксиэтил)(MOE), модификации ди-PS и модификации пептид-содержащей нуклеиновой кислотой (PNA);

одну или более внутрицепевых перекрестных связей;

один или более флуоресцентных красителей, конъюгированных с ними, а в некоторых вариантах осуществления изобретения, у 5'- или 3'-конца GRON; и

одно или более оснований, которые повышают энергию гибридизации.

[0201] Используемый здесь термин «неоднозначно спаренное основание» относится к замене одного или более нуклеотидных оснований эталонной нуклеотидной последовательности, где такая замена не приводит к модификации аминокислотной последовательности, кодируемой нуклеотидом, по сравнению с эталонной последовательностью.

[0202] Используемый здесь термин «не-нуклеотид» или «нуклеотид, не содержащий нуклеотидных оснований» означает любую группу или соединение, которые могут быть включены в цепь нуклеиновой кислоты вместо одного или более нуклеотидных звеньев, включая замену сахаром и/или фосфатом, и которые позволяют оставшимся основаниям проявлять свою ферментативную активность. Группа или соединение, не содержащие главных оснований, представляют собой группу или соединение, которые не содержат общеизвестных нуклеотидных оснований, таких как аденозин, гуанин, цитозин, урацил или тимин. Они могут иметь замены на 2'- или 3'-Н или ОН, как описано в настоящей заявке и в литературе.

[0203] Как описано в настоящей заявке, в некоторых вариантах осуществления изобретения, качество GRON и эффективность конверсии могут быть улучшены путем синтеза всего или части GRON с использованием нуклеотидных мультимеров, таких как димеры, тримеры, тетрамеры и т.п., повышающие его чистоту.

[0204] В некоторых вариантах осуществления изобретения, последовательность дезоксирибонуклеиновой кислоты-мишени (ДНК) присутствует в клетке растения, например, последовательность ДНК-мишени присутствует в геноме клетки растения. Клетка растения может быть нетрансгенной или трансгенной, и последовательность ДНК-мишени может представлять собой трансген или эндогенный ген клетки растения.

[0205] В некоторых вариантах осуществления изобретения, условия, которые способствуют улучшению одного или более процессов репарации клеточной ДНК, включают введение одного или более соединений, которые индуцируют разрывы одноцепочечной или двухцепочечной ДНК в клетке растения до, во время или после доставки GRON в клетку растения. Типичные соединения описаны в настоящей заявке.

[0206] В некоторых вариантах осуществления изобретения, способы дополнительно включают регенерацию растения, имеющего мутацию, внесенную посредством GRON из клетки растения и могут включать сбор семян этого растения.

[0207] В своих родственных аспектах, настоящее изобретение относится к клеткам растений, содержащим геномную модификацию, введенную посредством GRON в соответствии с описанными здесь способами; к растению, содержащему геномную

модификацию, введенную посредством GRON в соответствии с описанными здесь способами; или к семенам, содержащим геномную модификацию, введенную посредством GRON в соответствии с описанными здесь способами; или к потомству семян, содержащему геномную модификацию, введенную посредством GRON в соответствии с описанными здесь способами.

Конструкции

[0208] Описанные здесь молекулы нуклеиновой кислоты (например, сайт-специфические нуклеазы или руководящая РНК для CRISPR) могут быть использованы для получения рекомбинантных конструкций нуклеиновых кислот. В одном варианте осуществления изобретения, молекулы нуклеиновой кислоты согласно изобретению могут быть использованы при получении конструкций нуклеиновой кислоты, например, экспрессионных кластеров для экспрессии в представляющих интерес растениях, микроорганизмах или животных, таких как конструкции для экспрессии SHP, которые имеют, но необязательно, одну или более описанных здесь мутаций. Такая экспрессия может быть временной, например, если конструкция не интегрирована в геном хозяина или поддерживается под контролем промотора и присутствует в положении генома-хозяина, если она становится интегрированной.

[0209] Экспрессионные кластеры могут включать регуляторные последовательности, функционально присоединенные к описанной здесь сайт-специфической нуклеазе или к последовательностям руководящей РНК. Кластер может дополнительно содержать по меньшей мере один дополнительный ген для совместной трансформации в организм. Альтернативно, дополнительный ген или гены могут быть представлены на множестве экспрессионных кластеров.

[0210] Конструкции нуклеиновых кислот могут иметь множество рестрикционных сайтов для инсерции последовательности, кодирующей сайт-специфическую нуклеазу, которая находится под транскрипционным контролем регуляторных областей. Конструкции нуклеиновой кислоты могут дополнительно содержать молекулы нуклеиновой кислоты, кодирующие гены селективного маркера.

[0211] Для получения конструкций нуклеиновой кислоты может быть использован любой промотор. Промотор может быть нативным или аналогичным или чужеродным или гетерологичным по отношению к описанным здесь последовательностям нуклеиновой кислоты растений, микроорганизмов или животных-хозяев. Кроме того, промотор может представлять собой природную последовательность или, альтернативно, синтетическую последовательность. Если промотор является «чужеродным» или «гетерологичным» по отношению к растению, микроорганизму или животному-хозяину, то предполагается, что промотор не присутствует в нативном растении, микроорганизме или животном, в которые его вводят. Используемый здесь химерный ген содержит кодирующую последовательность, функционально присоединенную к области инициации транскрипции, которая является гетерологичной по отношению к кодирующей последовательности.

[0212] Описанные здесь сайт-направленные нуклеазные последовательности могут быть экспрессированы с использованием гетерологичных промоторов.

[0213] При получении конструкций для регуляции экспрессии сайт-направленных нуклеазных последовательностей могут быть использованы любые промоторы, такие как конститутивные, тканеспецифические, индуцибельные или другие промоторы для экспрессии в растениях, микроорганизмах или животных. Конститутивными промоторами являются, например, коровый промотор Rsyn7 и другие конститутивные промоторы, раскрытые в WO 99/43 838 и в патенте США № 6075050; коровый промотор 35S CaMV (Odell et al., *Nature* 313: 810-812; 1985); актина риса (McElroy et al., *Plant Cell* 2: 163-171, 1990); убихитина (Christensen et al., *Plant Mol. Biol.* 12: 619-632, 1989 и Christensen et al., *Plant Mol. Biol.* 18: 675-689, 1992); pEMU (Last et al., *Theor. Appl. Genet.* 81: 581-588, 1991); MAS (Velten et al., *EMBO J.* 3: 2723-2730, 1984); промотор ALS (патент США № 5659026) и т.п. Другими конститутивными промоторами являются, например, промоторы, описанные в патентах США №№ 5608149; 5608144; 5604121; 5569597; 5466785; 5399680; 5268463; 5608142; и 6177611.

[0214] Для сайт-направленной экспрессии нуклеазы в конкретной ткани растения могут быть использованы тканеспецифические промоторы. Такими тканеспецифическими промоторами являются, но не ограничиваются ими, промоторы, специфичные для листьев, корней, семян и стеблей. Тканеспецифические промоторы описаны в публикациях Yamamoto et al., *Plant J.* 12(2):255-265, 1997; Kawamata et al., *Plant Cell Physiol.* 38(7):792-803, 1997; Hansen et al., *Mol. Gen Genet.* 254(3):337-343, 1997; Russell et al., *Transgenic Res.* 6(2):157-168, 1997; Rinehart et al., *Plant Physiol.* 112(3):1331-1341, 1996; Van Camp et al., *Plant Physiol.* 112(2):525-535, 1996; Canevascini et al., *Plant Physiol.* 112(2):513-524, 1996; Yamamoto et al., *Plant Cell Physiol.* 35(5):773-778, 1994; Lam, *Results Probl. Cell Differ.* 20:181-196, 1994; Orozco et al. *Plant Mol Biol.* 23(6):1129-1138, 1993; Matsuoka et al., *Proc Nat'l. Acad. Sci. USA* 90(20):9586- 9590, 1993; и Guevara-Garcia et al., *Plant J.* 4(3):495-505, 1993.

[0215] Конструкции нуклеиновых кислот могут также включать области терминации транскрипции. Если используются области терминации транскрипции, то для получения конструкций нуклеиновой кислоты может быть использована любая область терминации. Так, например, область терминации может происходить от другого источника (то есть, чужеродного или гетерологичного для промотора). Примерами областей терминации, которые доступны для их использования в конструкциях согласно изобретению, являются области, происходящие от Ti-плазмиды *A. tumefaciens*, такие как области терминации октопинсинтазы и нопалинсинтазы. См. также Guerineau et al., *Mol. Gen. Genet.* 262: 141-144, 1991; Proudfoot, *Cell* 64: 671-674, 1991; Sanfacon et al., *Genes Dev.* 5: 141-149, 1991; Mogen et al., *Plant Cell* 2: 1261-1272, 1990; Munroe et al., *Gene* 91: 151-158, 1990; Ballas et al., *Nucleic Acids Res.* 17: 7891-7903, 1989; и Joshi et al., *Nucleic Acid Res.* 15: 9627-9639, 1987.

[0216] В комбинации с любым раскрытыми здесь аспектами, вариантами,

способами и/или композициями, нуклеиновые кислоты могут быть оптимизированы для повышения уровня экспрессии в трансформированном растении. То есть, нуклеиновые кислоты, кодирующие сайт-направленные нуклеазные белки, могут быть синтезированы с использованием предпочтительных для растений кодонов, повышающих уровень экспрессии. См., например, Campbell and Gowri (Plant Physiol. 92: 1-11, 1990), где обсуждается использование кодонов, предпочтительных для хозяина. Для синтеза предпочтительных для растения генов могут быть применены методы, описанные в литературе. См., например, патенты США №№ 5380831 и 5436391 и Murray et al., Nucleic Acids Res. 17: 477-498, 1989. См. также, например, Lanza et al., BMC Systems Biology 8: 33-43, 2014; Burgess-Brown et al., Protein Expr. Purif. 59: 94-102, 2008; Gustafsson et al., Trends Biotechnol 22: 346-353, 2004.

[0217] Кроме того, в описанные здесь последовательности нуклеиновой кислоты могут быть внесены и другие модификации. Так, например, известно, что дополнительные модификации последовательности повышают уровень экспрессии генов в клетке-хозяине. Эти модификации включают удаление последовательностей, кодирующих ложные сигналы полиаденилирования, сигналы сайтов сплайсинга экзонов/интронов, транспозон-подобные повторы и другие хорошо охарактеризованные последовательности, которые могут оказывать негативное влияние на экспрессию генов. Содержание G-C в последовательности также может быть скорректировано до средних уровней для целевой клетки-хозяина, и такое содержание может быть вычислено по отношению к известным генам, экспрессируемым в клетке-хозяине. Кроме того, последовательность может быть модифицирована для предотвращения образования предсказанных шпилечных вторичных структур мРНК.

[0218] Для получения конструкций согласно изобретению могут быть также использованы и другие последовательности нуклеиновой кислоты, например, для повышения уровня экспрессии сайт-направленной последовательности, кодирующей нуклеазу. Такие последовательности нуклеиновых кислот включают интроны AdhI кукурузы, ген интрона 1 (Callis et al., Genes and Development 1: 1183-1200, 1987) и лидерные последовательности (W-последовательность) вируса мозаики табака (TMV), вируса хлоротической пятнистости кукурузы и вируса мозаики люцерны (Gallie et al., Nucleic Acid Res. 15: 8693-8711, 1987; и Skuzeski et al., Plant Mol. Biol. 15: 65-79, 1990). Было показано, что первый интрон из локуса 1 сморщивания семян кукурузы повышает уровень экспрессии генов в химерных генных конструкциях. В патентах США №№ 5424412 и 5593874 раскрывается использование специфических интронов в экспрессионных конструкциях генов, и Gallie et al. (Plant Physiol. 106: 929-939, 1994) также показали, что эти интроны являются подходящими для регуляции экспрессии генов на тканеспецифической основе. Для дальнейшего повышения или оптимизации сайт-направленной экспрессии гена нуклеазы, в описанные здесь экспрессионные векторы растений могут быть также включены последовательности ДНК, содержащие области присоединения матрицы (MAR). Клетки растений, трансформированные такими

модифицированными системами экспрессии, могут затем демонстрировать сверхэкспрессию или конститутивную экспрессию нуклеотидной последовательности согласно изобретению.

[0219] Описанные здесь экспрессионные конструкции могут также включать последовательности нуклеиновых кислот, способные регулировать экспрессию сайт-направленной нуклеазной последовательности в хлоропластах или в других органеллах и в структурах как у прокариот, так и у эукариот. Такие последовательности нуклеиновой кислоты включают последовательности, нацеленные на хлоропласты, которые кодируют пептид переноса в хлоропласты для встраивания представляющего интерес генного продукта в хлоропласты клетки растения. Такие пептиды для переноса известны специалистам. Термин «функционально присоединенный» если он относится к последовательностям для доставки в хлоропласты, означает, что последовательность нуклеиновой кислоты, кодирующая транспортный пептид (то есть последовательность для доставки в хлоропласты), присоединена к сайт-направленным молекулам нуклеиновой кислоты нуклеазы, так, что две последовательности являются смежными и находятся в одной и той же рамке считывания. См., например, Von Heijne et al., *Plant Mol. Biol. Rep.* 9:104-126, 1991; Clark et al., *J. Biol. Chem.* 264:17544-17550, 1989; Della-Cioppa et al., *Plant Physiol.* 84:965-968, 1987; Romer et al., *Biochem. Biophys. Res. Commun.* 196:1414-1421, 1993; and Shah et al., *Science* 233:478-481, 1986.

[0220] Нацеленные на хлоропласты последовательности известны специалистам в данной области и включают небольшую субъединицу рибулозо-1,5-бисфосфаткарбоксилазы хлоропласта (Rubisco) (de Castro Silva Filho et al., *Plant Mol. Biol.* 30:769-780, 1996; Schnell et al., *J. Biol. Chem.* 266(5):3335-3342, 1991); 5-(энолпирувил)шикимат-3-фосфат-синтазы (EPSPS) (Archer et al., *J. Bioenerg. Biomemb.* 22(6):789-810, 1990); триптофан-синтазы (Zhao et al., *J. Biol. Chem.* 270(11):6081-6087, 1995); пластоцианина (Lawrence et al., *J. Biol. Chem.* 272(33):20357-20363, 1997); хоризмат-синтазы (Schmidt et al., *J. Biol. Chem.* 268(36):27447-27457, 1993); и светособирающий белок, связывающийся с хлорофиллом a/b (LHBP) (Lampira et al., *J. Biol. Chem.* 263:14996-14999, 1988). См. также Von Heijne et al., *Plant Mol. Biol. Rep.* 9:104-126, 1991; Clark et al., *J. Biol. Chem.* 264:17544-17550, 1989; Della-Cioppa et al., *Plant Physiol.* 84:965-968, 1987; Romer et al., *Biochem. Biophys. Res. Commun.* 196:1414-1421, 1993; и Shah et al., *Science* 233: 478-481, 1986.

[0221] В комбинации с любыми раскрытыми здесь аспектами, вариантами, способами и/или композициями, конструкции нуклеиновой кислоты могут быть получены для регуляции экспрессии последовательности, кодирующей мутантную сайт-направленную нуклеазу из хлоропласта клетки растения. Способы трансформации хлоропластов известны в данной области. См., например, Svab et al., *Proc. Nat'l. Acad. Sci. USA* 87:8526-8530, 1990; Svab and Maliga, *Proc. Nat'l. Acad. Sci. USA* 90:913-917, 1993; Svab and Maliga, *EMBO J.* 12:601-606, 1993. Этот метод основан на доставке ДНК, содержащей селективный маркер, путем выстреливания частиц, и нацеливании ДНК на

пластидный геном посредством гомологичной рекомбинации. Кроме того, пластидная трансформация может быть осуществлена путем трансактивации молчащего пластидного трансгена посредством тканеспецифической экспрессии РНК-полимеразы, кодируемой в ядре и доставляемой в пластид. Такая система описана в McBride et al. Proc. Nat'l. Acad. Sci. США 91: 7301-7305, 1994.

[0222] Представляющие интерес нуклеиновые кислоты, нацеленные на хлоропласт, могут быть оптимизированы для экспрессии в хлоропласте для учета различий в использовании кодонов между ядром растения и этой органеллой. Таким образом, представляющие интерес нуклеиновые кислоты могут быть синтезированы с использованием кодонов, предпочтительных для хлоропластов. См., например, патент США № 5380831, который включен в настоящее описание посредством ссылки.

[0223] Конструкции нуклеиновых кислот могут быть использованы для трансформации клеток растений и регенерации трансгенных растений, содержащих сайт-направленные последовательности, кодирующие нуклеазу. При этом, для трансформации растений имеется множество векторов и способов, которые являются доступными. См., например, патент США № 6753458, An, G. et al., Plant Physiol., 81:301-305, 1986; Fry, J. et al., Plant Cell Rep. 6:321-325, 1987; Block, M., Theor. Appl Genet. 76:767-774, 1988; Hinchee et al., Stadler. Genet. Symp. 203:212, 1990; Cousins et al., Aust. J. Plant Physiol. 18:481-494, 1991; Chee, P. P. and Slightom, J. L., Gene. 118:255-260, 1992; Christou et al., Trends. Biotechnol. 10:239-246, 1992; D'Halluin et al., Bio/Technol. 10:309-314, 1992; Dhir et al., Plant Physiol. 99:81-88, 1992; Casas et al., Proc. Nat'l. Acad. Sci. USA 90:11212-11216, 1993; Christou, P., In Vitro Cell. Dev. Biol.-Plant 29P:119-124, 1993; Davies et al., Plant Cell Rep. 12:180-183, 1993; Dong, J. A. and Mc Hughen, A., Plant Sci. 91:139-148, 1993; Franklin, C.I., Trieu, T.N., Cassidy, B.G., Dixon, R.A., Nelson, R.S. 1993, Plant Cell Report 12, 74-79; Golovkin et al., Plant Sci. 90:41-52, 1993; Guo Chin Sci. Bull. 38:2072-2078; Asano et al., Plant Cell Rep. 13, 1994; Ayeres, N. M. and Park, W. D., Crit. Rev. Plant. Sci. 13:219-239, 1994; Barcelo et al., Plant. J. 5:583-592, 1994; Becker et al., Plant. J. 5:299-307, 1994; Borkowska et al., Acta. Physiol Plant. 16:225-230, 1994; Christou, P., Agro. Food. Ind. Hi Tech. 5:17-27, 1994; Eapen et al., Plant Cell Rep. 13:582-586, 1994; Hartman et al., Bio-Technology 12:919-923, 1994; Ritala et al., Plant. Mol. Biol. 24:317-325, 1994; and Wan, Y. C. and Lemaux, P. G., Plant Physiol. 104:3748, 1994. Конструкции могут быть также перенесены в клетки растений посредством гомологичной рекомбинации.

[0224] Термин «дикого типа», если он относится к пептидной последовательности и нуклеотидной последовательности, означает, что пептидная последовательность и нуклеотидная последовательность (локус/ген/аллель), соответственно, обладают свойствами пептидной последовательности и нуклеотидной последовательности, происходящих от природного источника. Пептидная последовательность и нуклеотидная последовательность дикого типа представляют собой последовательности, которые наиболее часто встречаются в конкретной популяции и, таким образом, иногда называются «нормальной пептидной и нуклеотидной последовательностью» или

«последовательностью дикого типа», соответственно. «Последовательность дикого типа» может также означать последовательность, в которых конкретные нуклеотиды или конкретные кодоны или конкретные аминокислоты находятся в определенном положении или положениях.

[0225] «Консенсусную последовательность» определяют как последовательность аминокислот или нуклеотидов, которые содержат по меньшей мере на 25% идентичные аминокислоты или нуклеотиды или функционально эквивалентные аминокислоты или нуклеотиды. Идентичные или функционально эквивалентные аминокислоты или нуклеотиды необязательно должны быть смежными.

[0226] Нуклеоснование представляет собой основание, которое, в некоторых предпочтительных вариантах осуществления изобретения, представляет собой пурин, пиримидин или его производное или аналог. Нуклеозиды представляют собой нуклеоснования, которые содержат пентозофуранозильную группу, например, необязательно замещенный рибозид или 2'-деоксирибозид. Нуклеозиды могут быть связаны одной из нескольких связывающих групп, которые могут содержать, а могут и не содержать, фосфор. Нуклеозиды, которые связаны незамещенными фосфодиэфирными связями, называются нуклеотидами. Используемый здесь термин «нуклеоснование» включает пептидные нуклеоснования, субъединицы пептид-содержащих нуклеиновых кислот и морфолиновые нуклеоснования, а также нуклеозиды и нуклеотиды.

[0227] Олигонуклеоснование представляет собой полимер, содержащий нуклеоснования, а в некоторых вариантах осуществления изобретения, по меньшей мере часть, которая может гибридизоваться посредством спаривания оснований Уотсона-Крика с ДНК, имеющей комплементарную последовательность. Цепь олигонуклеоснований может содержать один 5'- и 3'-конец, которые представляют собой концевые нуклеоснования полимера. Конкретная цепь олигонуклеоснований может содержать нуклеоснования всех типов. Соединение олигонуклеоснований представляет собой соединение, содержащее одну или более цепей олигонуклеоснований, которые могут быть комплементарными и гибридизуются посредством спаривания оснований Уотсона-Крика. Нуклеоснования рибо-типа включают пентозофуранозил-содержащие нуклеоснования, где 2'-углерод представляет собой метилен, замещенный гидроксилом, алкилокси или галогеном. Нуклеоснования дезоксирибо-типа представляют собой нуклеоснования, отличающиеся от нуклеоснований рибо-типа, и включают все нуклеоснования, которые не содержат пентозофуранозильную группу.

[0228] В некоторых вариантах осуществления изобретения, цепь олигонуклеоснований может включать цепи олигонуклеоснований и сегменты или области цепей олигонуклеоснований. Цепь олигонуклеоснований может иметь 3'-конец и 5'-конец, а если цепь олигонуклеоснований является также и нитью, то 3'- и 5'-концы нити также являются 3'- и 5'-концами цепи.

[0229] Используемый здесь термин «кодон» означает последовательность из трех смежных нуклеотидов (РНК или ДНК), составляющих генетический код, который

определяет inserцию конкретной аминокислоты в полипептидную цепь во время синтеза белка или передачи сигнала остановки синтеза белка. Термин «кодон» также используется для обозначения соответствующих (и комплементарных) последовательностей из трех нуклеотидов в матричной РНК, в которую транскрибируется исходная ДНК.

[0230] Используемый здесь термин «гомология» относится к сходству последовательностей белков и ДНК. Термин «гомология» или «гомологичный» относится к степени идентичности. Гомология может быть частичной или полной. Частично гомологичная последовательность представляет собой последовательность, которая менее, чем на 100% идентична другой последовательности.

[0231] Термин «гетерозиготный» относится к наличию различных аллелей в одном или более генетических локусах в гомологичных сегментах хромосомы. Используемый здесь термин «гетерозиготный» может также относиться к образцу, клетке, клеточной популяции или к организму, в которых могут быть детектированы различные аллели в одном или более генетических локусах. Гетерозиготные образцы могут быть также определены известными методами, такими как, например, секвенирование нуклеиновых кислот. Так, например, если секвенирующая электрофореграмма показывает два пика в одном локусе, и оба пика имеют приблизительно одинаковый размер, то образец можно охарактеризовать как гетерозиготный. Или, если один пик меньше другого, но составляет по меньшей мере приблизительно 25% от размера большего пика, то этот образец можно охарактеризовать как гетерозиготный. В некоторых вариантах осуществления изобретения, меньший пик составляет по меньшей мере приблизительно 15% от большего пика. В других вариантах осуществления изобретения, меньший пик составляет по меньшей мере приблизительно 10% от большего пика. В других вариантах осуществления изобретения, меньший пик составляет по меньшей мере приблизительно 5% от большего пика. В других вариантах осуществления изобретения детектируется минимальное количество меньшего пика.

[0232] Используемый здесь термин «гомозиготный» относится к наличию идентичных аллелей в одном или более генетических локусах в сегментах гомологичной хромосомы. Термин «гомозиготный» может также относиться к образцу, клетке, клеточной популяции или организму, в которых могут быть обнаружены одни и те же аллели в одном или более генетических локусах. Гомозиготные образцы могут быть определены известными методами, такими как, например, секвенирование нуклеиновых кислот. Так, например, если секвенирующая электрофореграмма показывает один пик в одном конкретном локусе, то образец можно охарактеризовать как «гомозиготный» по отношению к этому локусу.

[0233] Термин «гемизиготный» относится к гену или к сегменту гена, присутствующим только один раз в генотипе клетки или организма, поскольку второй аллель был делетирован или отсутствует в гомологичном сегменте хромосомы. Используемый здесь термин «гемизиготный» может также относиться к образцу, к клетке, к клеточной популяции или к организму, в которых аллель в одном или более

генетических локусах может быть обнаружен в этом генотипе только один раз.

[0234] Используемый здесь термин «статус зиготности» относится к образцу, к клеточной популяции или к организму, которые являются гетерозиготными, гомозиготными или гемизиготными, как было определено методами тестирования, известными специалистам и описанными в настоящей заявке. Термин «статус зиготности нуклеиновой кислоты» означает ее зиготность после того, как было определено, является ли источник нуклеиновой кислоты гетерозиготным, гомозиготным или гемизиготным. «Статус зиготности» может относиться к различиям в положениях одного нуклеотида в последовательности. В некоторых способах, статус зиготности образца в отношении одной мутации может быть классифицирован как гомозиготный образец дикого типа, гетерозиготный образец (то есть, с одним аллелем дикого типа и одним мутантным аллелем), гомозиготный мутант или гемизиготный образец (то есть, с одной копией аллеля дикого типа или мутантного аллеля).

[0235] Используемый здесь термин «**RTDS**» относится к Системе Быстрой Идентификации Признаков™ (**RTDS™**), разработанной Cibus. **RTDS** представляет собой сайт-специфическую систему модификации генов, которая является эффективной для внесения точных изменений в последовательность генов без включения чужеродных генов или регуляторных последовательностей.

[0236] Термин «приблизительно», используемый здесь в количественном выражении, означает плюс или минус 10%. Так, например, «приблизительно 3%» будет охватывать 2,7-3,3%, а «приблизительно 10%» будет охватывать 9-11%. Кроме того, если термин «приблизительно» используется здесь в сочетании с количественным выражением, то подразумевается, что в дополнение к значению плюс или минус 10%, также рассматривается и описывается точная величина этого количественного значения. Так, например, выражение «приблизительно 3%» явно означает, описывает и включает точно 3%.

RTDS и олигонуклеотиды для репарации (GRON)

[0237] Различные аспекты и варианты рассматриваемых здесь способов и композиций включают способы повышения эффективности нацеливания модификаций на конкретные локусы в геномных или других нуклеотидных последовательностях (например, модификации гена *SHP*, такие как рассматриваемые здесь модификации).

[0238] В некоторых вариантах осуществления изобретения, **RTDS** разработаны на основе изменения целевого гена с использованием собственной системы генной репарации клетки для специфической модификации последовательности гена *in situ* без встраивания чужеродной ДНК и последовательности регуляции экспрессии гена. Эта процедура может привести к точному изменению генетической последовательности, но с сохранением остальной части генома в неизменном виде. В некоторых вариантах осуществления изобретения, в данном случае, в отличие от традиционных трансгенных ГМО, в растении не происходит интеграции чужеродного генетического материала, а также не присутствует какого-либо чужеродного генетического материала. Модификации

в генетической последовательности, введенные с помощью **RTDS**, не встраиваются случайным образом. Поскольку модифицированные гены остаются в своем природном положении, то не наблюдается какого-либо случайного, неконтролируемого или неблагоприятного паттерна экспрессии.

[0239] Молекула, которая влияет на такое изменение, представляет собой описанный здесь химически синтезированный олигонуклеотид (GRON), который может состоять из оснований ДНК, и модифицированных оснований РНК, а также из других химических молекул, и предназначен для гибридизации в положении целевого гена в целях создания пары оснований с ошибочным спариванием. Эта пара оснований с ошибочным спариванием действует как сигнал для доставки собственной природной системы репарации генов клетки в этот сайт и коррекции (замены, инсерции или делеции) сконструированного нуклеотида или нуклеотидов в гене. После завершения процессов коррекции, молекула GRON разлагается, и только что модифицированный или репарированный ген экспрессируется по механизмам нормальной эндогенной регуляции генов.

[0240] Описанные здесь способы и композиции могут быть осуществлены на практике или разработаны с использованием «олигонуклеооснований для репарации генов» (GRON), имеющих конформации и химические структуры, подробно описанные здесь и ниже. Рассматриваемые здесь термины «олигонуклеооснования для репарации генов» были также описаны в опубликованной научной и патентной литературе под другими названиями, включая «рекомбинагенные олигонуклеооснования»; «химерные олигонуклеотиды РНК/ДНК»; «химерные олигонуклеотиды»; «смешанные дуплексные олигонуклеотиды» (MDON); «РНК-ДНК-олигонуклеотиды (RDO)»; «олигонуклеотиды для нацеливания на гены»; «генопласты»; «одноцепочечные модифицированные олигонуклеотиды»; «одноцепочечные олигодезоксинуклеотидные мутантные векторы» (SSOMV); «дуплексные мутантные векторы»; и «гетеродуплексные мутантные векторы». Олигонуклеооснование для репарации генов может быть введено в клетку растения любым известным методом, включая, но не ограничиваясь ими, использование микроносителей (биобаллистическая доставка), использование микроволокон, поглощение, опосредованное полиэтиленгликолем (ПЭГ), электропорацию и микроинъекцию.

[0241] В одном варианте осуществления изобретения, олигонуклеооснование для репарации гена представляет собой смешанный дуплексный олигонуклеотид (MDON), в котором нуклеотиды РНК-типа смешанного дуплексного олигонуклеотида становятся резистентными к РНКазе в результате замены 2'-гидроксила на функциональные группы фтора, хлора или брома, или в результате введения заместителя в положение 2'-О. Подходящими заместителями являются заместители, описанные в Kmiec II. Альтернативными заместителями являются заместители, описанные в патенте США No. 5334711 (Sproat) и заместители, описанные в патентных публикациях EP 629387 и EP 679657 (совместно поданные заявки Martin Applications), которые включены в настоящее

описание посредством ссылки. Используемое здесь 2'-фтор-, хлор- или бром-содержащее производное рибонуклеотида или рибонуклеотида, имеющего Т-ОН, замещенный заместителем, описанным в заявках Martin Applications или Sproat, называется «Т-замещенным рибонуклеотидом». Используемый здесь термин «нуклеотид РНК-типа» означает Т-гидроксил или 2'-замещенный нуклеотид, который связан с другими нуклеотидами смешанного дуплексного олигонуклеотида посредством незамещенной фосфодиэфирной связи или любой из не-природных связей, описанных в Kmiec I или Kmiec II. Используемый здесь термин «нуклеотид дезоксирибо-типа» означает нуклеотид, имеющий Т-Н, который может быть связан с другими нуклеотидами олигонуклеоснования для репарации гена посредством незамещенной фосфодиэфирной связи или любых не-природных связей, описанных в Kmiec I или Kmiec II.

[0242] В конкретных вариантах осуществления изобретения, олигонуклеоснование для репарации гена может представлять собой смешанный дуплексный олигонуклеотид (MDON), который связан исключительно незамещенными фосфодиэфирными связями. В альтернативных вариантах осуществления изобретения, связь представляет собой замещенные фосфодиэфиры, производные фосфодиэфира и нефосфорные связи, как описано в Kmiec II. В еще одном варианте осуществления изобретения, каждый нуклеотид РНК-типа в смешанном дуплексном олигонуклеотиде представляет собой 2'-замещенный нуклеотид. Особенно предпочтительными вариантами 2'-замещенных рибонуклеотидов являются 2'-фтор-, Т-метокси-, 2'-пропилокси-, 2'-аллилокси-, 2'-гидроксилэтилокси-, 2'-метоксиэтилокси-, 2'-фторпропилокси- и 2'-трифторпропилокси-замещенные рибонуклеотиды. Более предпочтительными вариантами 2'-замещенных рибонуклеотидов являются 2'-фтор-, 2'-метокси-, 2'-метоксиэтилокси- и 2'-аллилокси-замещенные нуклеотиды. В другом варианте осуществления изобретения, смешанный дуплексный олигонуклеотид связан незамещенными фосфодиэфирными связями.

[0243] Хотя смешанные дуплексные олигонуклеотиды (MDON), имеющие 2'-замещенный нуклеотид РНК-типа только одного типа, является более удобным для синтеза, однако, способы согласно изобретению могут быть осуществлены с использованием смешанных дуплексных олигонуклеотидов, имеющих нуклеотиды РНК-типа двух или более типов. На функцию сегмента РНК может не влиять разрыв, вызванный введением дезоксинуклеотида между двумя тринуклеотидами РНК-типа, и, соответственно, термин «сегмент РНК» охватывает такой «сегмент РНК с разрывом». Сегмент РНК без разрывов называется непрерывным сегментом РНК. В альтернативном варианте осуществления изобретения, сегмент РНК может содержать чередующиеся резистентные к РНКазе и незамещенные 2'-ОН-нуклеотиды. В некоторых вариантах осуществления изобретения, смешанные дуплексные олигонуклеотиды имеют менее 100 нуклеотидов, а в других вариантах, менее, чем 85 нуклеотидов, но более, чем 50 нуклеотидов. Первая и вторая цепи представляют собой пары оснований Уотсона-Крика. В одном варианте осуществления изобретения, цепи смешанного дуплексного

олигонуклеотиды ковалентно связаны посредством линкера, такого как одноцепочечный гекса-, пента- или тетрануклеотид, а поэтому первая и вторая цепи представляют собой сегменты одной олигонуклеотидной цепи, имеющие один 3'-конец и один 5'-конец. 3'- и 5'-концы могут быть защищены путем введения «шпилечного кэпа», а поэтому 3'- и 5'-концевые нуклеотиды являются нуклеотидами Уотсона-Крика, спаренными со смежными нуклеотидами. Второй шпилечный кэп может быть дополнительно введен на стыке между первой и второй цепями, удаленными от 3'- и 5'-концов, что позволяет стабилизировать пары Уотсона-Крика между первой и второй цепями.

[0244] Первая и вторая цепи содержат две области, которые гомологичны двум фрагментам гена/аллеля-мишени, то есть, они имеют такую же последовательность, как и ген/аллель-мишень. Гомологичная область содержит нуклеотиды сегмента РНК и может содержать один или более нуклеотидов ДНК-типа, соединяющих сегмент ДНК, а также может содержать нуклеотиды ДНК-типа, которые не находятся внутри промежуточного сегмента ДНК. Две области гомологии разделены областью, и каждая из них является смежной с областью, имеющей последовательность, которая отличается от последовательности гена-мишени, и эти области называются «гетерологичными областями». Гетерологичная область может содержать один, два или три ошибочно спаренных нуклеотида. Ошибочно спаренные нуклеотиды могут быть смежными или, альтернативно, они могут быть разделены одним или двумя нуклеотидами, которые гомологичны гену/аллелю-мишени. Альтернативно, гетерологичная область может также содержать инсерцию одного, двух, трех или пяти или нескольких нуклеотидов. Альтернативно, последовательность смешанного дуплексного олигонуклеотида может отличаться от последовательности гена/аллеля-мишени только делецией одного, двух, трех или пяти или нескольких нуклеотидов из смешанного дуплексного олигонуклеотида. В этом случае, длина и положение гетерологичной области, предположительно считается длиной делеции, даже если в гетерологичной области отсутствуют нуклеотиды смешанного дуплексного олигонуклеотида. Расстояние между фрагментами гена-мишени, которые комплементарны двум гомологичным областям, идентично длине гетерологичной области, если предполагается введение замены или замен. Если гетерологичная область содержит инсерцию, то гомологичные области разделяются в смешанном дуплексном олигонуклеотиде дальше, чем их комплементарные гомологичные фрагменты в гене/аллеле, и наоборот, когда гетерологичная область кодирует делецию.

[0245] Каждый из сегментов РНК смешанных дуплексных олигонуклеотидов является частью гомологичной области, то есть, области, которая по своей последовательности идентична фрагменту гена-мишени, причем, в некоторых вариантах осуществления изобретения, эти сегменты вместе содержат по меньшей мере 13 нуклеотидов РНК-типа, а в некоторых вариантах, от 16 до 25 нуклеотидов РНК-типа или в других вариантах, 18-22 нуклеотидов РНК-типа или в некоторых вариантах, 20 нуклеотидов. В одном варианте осуществления изобретения, сегменты РНК областей гомологии разделены промежуточным сегментом ДНК, и являются смежными с этим

сегментом, то есть «сединены» с ним. В одном варианте осуществления изобретения, каждый нуклеотид гетерологичной области представляет собой нуклеотид промежуточного сегмента ДНК. Промежуточный сегмент ДНК, который содержит гетерологичную область смешанного дуплексного олигонуклеотида, называется «сегментом-мутатором».

[0246] В другом варианте способов и композиций согласно изобретению, олигонуклеосодержащее основание для репарации генов (GRON) представляет собой одноцепочечный олигодезоксинуклеотидный мутантный вектор (SSOMV), такой как вектор, описанный в Международной патентной заявке PCT/US00/23457, в патентах США № 6271360, 6479292 и 7060500, которые включены в настоящее описание в полном объеме посредством ссылки. Последовательность SSOMV основана на тех же принципах, что и мутантные векторы, описанные в патентах США №. 5756325; 5871984; 5760012; 5888983; 5795972; 5780296; 5945339; 6004804; и 6010907 и в публикациях Международных заявок WO 98/49350; WO 99/07865; WO 99/58723; WO 99/58702; и WO 99/40789. Последовательность SSOMV содержит две области, которые гомологичны последовательности-мишени, разделенной областью, которая содержит желаемую генетическую модификацию, называемую областью-мутатором. Область-мутатор может иметь последовательность, которая имеет такую же длину, как и последовательность, которая разделяет гомологичные области в последовательности-мишени, но имеет другую последовательность. Такая область-мутатор может приводить к замене. Альтернативно, гомологичные области в SSOMV могут быть смежными друг с другом, а области в гене-мишени, имеющие одинаковую последовательность, разделены одним, двумя или более нуклеотидами. Такой SSOMV вызывает делецию нуклеотидов в гене-мишени, которые отсутствуют в SSOMV. И наконец, последовательность гена-мишени, которая идентична гомологичным областям, может быть смежной в гене-мишени, но разделенной одним, двумя или более нуклеотидами в последовательности SSOMV. Такой SSOMV вызывает инсерцию в последовательности гена-мишени. В некоторых вариантах осуществления изобретения, SSOMV не гибридизуется сам с собой.

[0247] Нуклеотиды SSOMV представляют собой дезоксирибонуклеотиды, которые связаны немодифицированными фосфодиэфирными связями, за исключением того, что 3'-концевая и/или 5'-концевая межнуклеотидная связь или, альтернативно, две 3'-концевые и/или 5'-концевые межнуклеотидные связи могут представлять собой фосфотиоат или фосфоамидат. Используемый здесь термин «межнуклеотидная связь» означает связь между нуклеотидами SSOMV и не включает связь между 3'-концевым нуклеотидом или 5'-концевым нуклеотидом и блокирующим заместителем. В конкретном варианте осуществления изобретения, длина SSOMV составляет от 21 до 55 дезоксинуклеотидов; области гомологии имеют, соответственно, общую длину по меньшей мере 20 дезоксинуклеотидов; а по меньшей мере каждая из двух областей гомологии должна иметь длину по меньшей мере 8 дезоксинуклеотидов.

[0248] SSOMV может быть сконструирован так, чтобы он был комплементарен

либо кодирующей, либо не кодирующей цепи гена-мишени. Если желаемая мутация представляет собой замену одного основания, то предпочтительно, чтобы оба нуклеотида-мутатора представляли собой пиримидин. Для достижения желаемого функционального результата, предпочтительно, чтобы нуклеотид-мутатор и целевой нуклеотид в комплементарной цепи представляли собой пиримидины. Особенно предпочтительными являются SSOMV, которые кодируют мутации трансверсии, то есть, нуклеотид-мутатор С или Т, который ошибочно спаривается, соответственно, с нуклеотидом С или Т в комплементарной цепи.

[0249] Конструирование 2'-OME-GRON. В различных вариантах осуществления изобретения, GRON может иметь нуклеотиды РНК и ДНК и/или нуклеос основания других типов. В некоторых вариантах осуществления изобретения, один или более нуклеотидов ДНК или РНК содержат модификацию. В некоторых вариантах осуществления изобретения, первый 5'-нуклеотид представляет собой нуклеотид РНК, а остальные нуклеотиды представляют собой ДНК. В других вариантах осуществления изобретения, первый 5'-нуклеотид РНК модифицирован 2'-О-Ме. В других вариантах осуществления изобретения, первые два, три, четыре, пять, шесть, семь, восемь, девять, десять или более 5'-нуклеотидов представляют собой нуклеотиды РНК, а остальные нуклеотиды представляют собой ДНК. В других вариантах осуществления изобретения, один или более из первых двух, трех, четырех, пяти, шести, семи, восьми, девяти, десяти или более 5'-нуклеотидов РНК модифицированы 2'-О-Ме. В клетках растения, двухцепочечные разрывы в ДНК обычно репарируют по пути репарации ДНК NHEJ. Этот путь не требует присутствия матрицы для репарации ДНК, а поэтому существует вероятность ошибки. Преимущество этого пути репарации ДНК в клетке растения состоит в том, что он является быстрым, универсальным и, что самое важное, может быть осуществлен в тех случаях, когда клетка не подвергается репликации ДНК. Другой путь репарации ДНК, который функционирует при репарации двухцепочечных разрывов за пределами репликативной вилки в клетках растения, называется матричной репарацией, однако, в отличие от пути NHEJ, этот тип репарации является точным и требует использования ДНК-матрицы (GRON).

Повышение эффективности

[0250] Настоящее изобретение может включать любой из ряда подходов для повышения эффективности конверсии гена-мишени с использованием репаративных олигонуклеотидов, и эти подходы могут быть использованы отдельно или в комбинации друг с другом. К ним относятся, например:

1. Введение модификаций в репаративные олигонуклеотиды, которые запускают механизмы репарации ДНК в нужном сайте (при ошибочном спаривании).

А. Введение одного или более не-нуклеотидных сайтов в олигонуклеотид (например, в пределах 10 оснований, а в некоторых вариантах, в пределах 5 оснований желаемого сайта с ошибочным спариванием), которое приводит к образованию повреждения, представляющего собой промежуточное звено в репарации посредством

вырезания основания (BER), и которое запускает механизм BER в непосредственной близости от сайта, нацеленного на конверсию под действием репаративного олигонуклеотида. Олигонуклеотиды, модифицированные dSpacer (не-нуклеотидным фураном), могут быть получены, как описано, например, Takeshita et al., *J. Biol. Chem.*, 262: 10171-79, 1987.

В. Включение соединений, которые вызывают одно- или двухцепочечные разрывы, либо в олигонуклеотид, либо в комбинацию олигонуклеотидов, что приводит к образованию повреждения, которое репарируется посредством NHEJ, присоединения по концам, опосредуемого микрогомologies (MMEJ), и гомологичной рекомбинации. Так, например, антибиотики семейства блеомицинов, цинковые пальцы, FokI (или рестриктирующий фермент класса IIS любого типа) и другие нуклеазы могут быть ковалентно связаны с 3' или 5'-концом олигонуклеотидов для репарации в целях введения двухцепочечных разрывов в непосредственной близости от сайта, нацеленного на конверсию под действием олигонуклеотида для репарации. Антибиотики семейства блеомицинов представляют собой ДНК-расщепляющие гликопептиды, в состав которых входят блеомицин, зеоцин, флеомицин, таллисомидин, пепломицин и другие.

С. Введение одного или более 8'-оксо-dA или dG, включенных в олигонуклеотид (например, в пределах 10 оснований, а в некоторых вариантах, в пределах 5 оснований желаемого сайта с ошибочным спариванием), которое приводит к образованию повреждения, аналогичного повреждениям, создаваемым молекулами активного кислорода. Эти повреждения индуцируют так называемую систему «принудительной репарации». См., например, Kim et al., *J. Biochem. Mol. Biol.* 37: 657-62, 2004.

2. Повышение стабильности олигонуклеотидов для репарации:

- Введение обратного основания (idC) у 3'-конца олигонуклеотида для создания 3'-блокированного конца на олигонуклеотиде для репарации.

- Введение одного или более 2'-О-метил-нуклеотидов или оснований, которые повышают энергию гибридизации (см., например, WO2007/073149) у 5'- и/или 3'-конца олигонуклеотида для репарации.

- Введение одного или множества 2'-О-метил-РНК-нуклеотидов у 5'-конца олигонуклеотида для репарации, в основания ДНК, которые обеспечивают желаемый сайт с ошибочным спариванием и с получением структуры нуклеиновой кислоты, подобной фрагменту Оказаки.

- Добавление конъюгированных (5' или 3') интеркалирующих красителей, таких как акридин, псорален, этидийбромид и красители Syber.

- Введение 5'-концевого кэпа, такого как молекула для захвата Т/А, холестеринная группа, SIMA (HEX), giboC и амидит.

- Модификации остова, такие как модификации фосфотиоатом, модификации 2'-О-метилом, модификации метилфосфонатом, модификации заблокированной нуклеиновой кислотой (LNA), модификации MOE (метоксиэтилом), модификации ди-PS и модификации пептид-содержащей нуклеиновой кислотой (PNA).

- Перекрестное связывание репаративного олигонуклеотида, например, с внутрицепевыми перекрестно связывающими реагентами, такими как цисплатин и митомицин С.

- Конъюгирование с флуоресцентными красителями, такими как Cy3, DY547, Cy3.5, Cy3B, Cy5 и DY647.

3. Повышение энергии гибридизации репаративного олигонуклеотида путем включения оснований, которые повышают энергию гибридизации (см., например, WO2007/073149).

4. Повышение качества синтеза репаративных олигонуклеотидов с использованием нуклеотидных мультимеров (димеров, тримеров, тетрамеров и т.п.) в качестве структурных блоков для синтеза. Это позволяет уменьшить число стадий связывания и облегчить отделение полноразмерных продуктов от структурных блоков.

5. Использование длинных олигонуклеотидов для репарации (то есть, длиной более, чем 55 нуклеотидов, например, длиной, указанной в настоящей заявке), например, имеющих одну или более мутаций или две или более мутаций, нацеленных на олигонуклеотид для репарации.

[0251] Примеры вышеуказанных подходов приведены в Таблице А.

Таблица А: Репрезентативные химические структуры GRON

	Тип олигонуклеотида	Модификации
5'-модификации	Молекула для захвата Т/А	Молекула для захвата Т/А
Модификации остова	Фосфориоат	PS
Интеркалирующие красители	5'-акридин-3'	idC-акридин, idC
2'-О-метил		ДНК/РНК
Замены Cy3		DY547
Агенты, ускоряющие реакцию	2'-О-Ме-олигонуклеотиды, сконструированные у 5'- и 3'-концов конвертирующего олигонуклеотида	2'-О-Ме
Не-нуклеотид	Не-нуклеотидный сайт, расположенный в различных положениях у 5'- и 3'-концов по отношению к конвертирующему основанию. 44-мер	Не-нуклеотид 2
Вспомогательный метод	Вспомогательный метод Перекрывание:	Cy3, idC на одном, но не на другом:

	2 олигонуклеотида: 1 с Су3/idC, 1 немодифицированный олигонуклеотид для репарации	
Вспомогательный метод	Вспомогательный метод Без перекрываения: 2 олигонуклеотида: 1 с Су3/idC, 1 немодифицированный олигонуклеотид для репарации	Только для получения немодифицированного олигонуклеотидао
Не-нуклеотид	ТГФ-сайт, расположенный в различных положениях у 5'- и 3'-концов по отношению к конвертирующему основанию. 44-мер	Тетрагидрофуран (d-спейсер)
Модификации остова	9	2'-О-Ме
Тримеры		Тримерные амидиты, Су3. idC
Принудительная репарация		8'-оксо-dA, 5'-Су3, idC
Принудительная репарация		8'-оксо-dA, 5'-Су3, idC
Двухцепочечный разрыв		Блеомицин
Сшивающий линкер		Цисплатин
Сшивающий линкер		Митомицин С
Агенты, ускоряющие реакцию	Супероснования у 5'- и 3'- концов конвертирующего олигонуклеотида	2-амино-dA и 2-тио-Т
Суперолигонуклеотиды		2'-амино-d, 5'-Су3, idC
Суперолигонуклеотиды		2-тио-Т, 5'-Су3, idC
Суперолигонуклеотиды		7-деаза-А, 5'-Су3, idC
Суперолигонуклеотиды		7-деаза-G, 5'-Су3, idC

Суперолигонуклеотиды		Пропанил-dC, 5'-Cy3, idC
Интеркалирующие красители	5'-псорален/3'-idC	Псорален, idC
Интеркалирующие красители	5'-этидийбромид	Этидийбромид
Интеркалирующие красители	5'-красители Syber	Красители Syber
5'-модификации	5'-Chol/3'-idC	Холестерин
Двойная мутация	Длинный олигонуклеотид (55+ оснований) с 2 мутациями	Любая модификация
5'-модификации	5'-SIMA-HEX/3'-idC	SIMA-HEX, idC
Модификации остова	9	Метилфосфонаты
Модификации остова		LNA
Модификации остова	1	МОЕ (метоксиэтил)
Замены Cy3		Cy3.5
Замены Cy3		Cy5
Модификации остова		Ди-PS
5'-модификации		riboC для ветви NM
Модификации остова		PNA
Замены Cy3		DY647
5'-модификации	5'-ветвь	Амидит с симметричной ветвью/idC

[0252] Вышеуказанные модификации могут также включать известные модификации нуклеотидов, такие как метилирование, введение 5'-интеркалирующих красителей, модификации у 5'- и 3'-концов, модификации остова, введение сшивающих линкеров, циклизация и «кэпирование» и замена одного или более природных нуклеотидов таким аналогом, как инозин. Модификации нуклеотидов включают добавление акридина, амина, биотина, каскадного синего, холестерина, Cy3@, Cy5@, Cy5.5@, дабоила, дигоксигенина, динитрофенила, Эданса, 6-FAM, флуоресцеина, 3'-глицерила, HEX, IRD-700, IRD-800, JOE, фосфата псоралена, родамина, ROX, тиола (SH), спейсеров, TAMRA, TET, AMCA-S'', SE, BODIPY°, флотского синего@, бледно-синего@, Орегона зеленого@, родаминового зеленого@, родаминового красного@, родолового зеленого@ и тexasского красного@. Модификации полинуклеотидного остова включают метилфосфонат, 2'-ОМе-метилфосфонат-РНК, фосфортиоат, РНК, 2'-ОМе-РНК. Модификации оснований включают 2-амино-dA, 2-аминопурин, 3'-(ddA), 3'-dA

(кордицепин), 7-деаза-dA, 8-Br-dA, 8-оксо-dA, N6-Me-dA, не-нуклеотидный сайт (d-спейсер), биотин-dT, 2'-OMe-5Me-C, 2'-OMe-пропинил-C, 3'-(5-Me-dC), 3'-(ddC), 5-Br-dC, 5-I-dC, 5-Me-dC, 5-F-dC, карбокси-dT, конвертируемый dA, конвертируемый dC, конвертируемый dG, конвертируемый dT, конвертируемый dU, 7-деаза-dG, 8-Br-dG, 8-оксо-dG, O6-Me-dG, S6-DNP-dG, 4-метилендол, 5-нитроиндол, 2'-OMe-инозин, 2'-dl, об-фенил-dl, 4-метил-индол, 2'-дезоксинебуларин, 5-нитроиндол, 2-аминопурин, dP (аналог пурина), dK (аналог пиримидина), 3-нитропиррол, 2-тио-dT, 4-тио-dT, биотин-dT, карбокси-dT, O4-Me-dT, O4-триазол-dT, 2'-OMe-пропинил-U, 5-Br-dU, 2'-dU, 5-F-dU, 5-I-dU, O4-триазол-dU. Указанные термины также охватывают пептид-содержащие нуклеиновые кислоты (PNA), то есть, аналог ДНК, в котором остов представляет собой псевдопептид, состоящий из N-(2-аминоэтил)глициновых звеньев, а не сахара. PNA имитируют поведение ДНК и связывают комплементарные цепи нуклеиновых кислот. Нейтральный остов PNA обеспечивает более сильное связывание и более высокую специфичность, чем это обычно достигается. Кроме того, уникальные химические, физические и биологические свойства PNA были использованы для продуцирования эффективных биомолекулярных средств, антисмысловых и антигенных агентов, молекулярных зондов и биосенсоров.

[0253] Олигонуклеос основания могут иметь ник(и), разрыв(ы), модифицированные нуклеотиды, такие как модифицированные олигонуклеотидные остовы, нуклеотиды, не имеющие нуклеотидных оснований, или другие химические молекулы. В другом варианте осуществления изобретения, по меньшей мере одна цепь олигонуклеос основания включает по меньшей мере один дополнительный модифицированный нуклеотид, например, 2'-O-метил-модифицированный нуклеотид, такой как МОЕ (метоксиэтил), нуклеотид, имеющий 5'-фосфортиоатную группу, концевой нуклеотид, связанный с холестерилловым производным; 2'-дезоксид-2'-фтор-модифицированный нуклеотид; 2'-дезоксид-модифицированный нуклеотид; блокированный нуклеотид; нуклеотид, не имеющий нуклеотидных оснований (в котором нуклеос основание отсутствует или вместо него присутствует гидроксильная группа) (см., например, Glen Research, <http://www.glenresearch.com/GlenReports/GR21-14.html>), 2'-амино-модифицированный нуклеотид, 2'-алкил-модифицированный нуклеотид, морфолино-нуклеотид, фосфорамидит и нуклеотид, содержащий не-природное основание. Также включены различные соли, смешанные соли и формы свободных кислот.

[0254] Предпочтительные модифицированные олигонуклеотидные каркасы включают, например, фосфортиоаты, хиральные фосфортиоаты, фосфодитиоаты, фосфотриэфиры, аминоалкилфосфотриэфиры, метил- и другие алкилфосфонаты, включая 3'-алкиленфосфонаты, 5'-алкиленфосфонаты и хиральные фосфонаты, фосфофинаты, фосфорамидаты, включая 3'-аминофосфорамидат и аминоалкилфосфорамидаты, тионофосфорамидаты, тионоалкилфосфонаты, тионоалкилфосфотриэфиры, селенофосфаты и боранофосфаты, имеющие нормальные 3'-5'-связи, их 2'-5'-связанные аналоги, и аналоги, имеющие обращенную полярность, где одна или более

межнуклеотидных связей представляют собой 3'-3'-, 5'-5' или 2'-2'-связь. Предпочтительные олигонуклеотиды, имеющие обращенную полярность, содержат одну 3'-3'-связь у крайней 3'- межнуклеотидной связи, то есть, один инвертированный нуклеозидный остаток, который может быть не-нуклеотидным (в котором нуклеооснование отсутствует или вместо него присутствует гидроксильная группа). Наиболее распространенным применением инвертированной связи является присоединение 3'-3'-связи к концу антисмыслового олигонуклеотида с фосфортиоатным остовом. 3'-3'-связь также предотвращает разложение антисмыслового олигонуклеотида экзонуклеазой посредством образования олигонуклеотида с двумя 5'-ОН-концами и без 3'-ОН-конца. Инвертированная связь может быть введена в специфические положения во время синтеза олигонуклеотидов посредством использования «обратных фосфорамидитов». Эти реагенты имеют фосфорамидитные группы в 5'-ОН-положении и диметокситритильную (DMT) защитную группу в 3'-ОН-положении. Обычно, защитная группа DMT находится на 5'-ОН, а фосфорамидит находится на 3'-ОН.

[0255] Примерами модифицированных оснований являются, но не ограничиваются ими, 2-аминопурин, 2'-амино-бутирилпиренуридин, 2'-аминоуридин, 2'-дезоксинуридин, 2'-фторцитидин, 2'-фторуридин, 2,6-диаминопурин, 4-тиоуридин, 5-бромурин, 5-фторцитидин, 5-фторуридин, 5-индоуридин, 5-метилцитидин, инозин, N3-метилуридин, 7-дезагуанин, 8-аминогексиламиноаденин, 6-тиогуанин, 4-тиотимин, 2-тиотимин, 5-йодуридин, 5-йодцитидин, 8-бромгуанин, 8-бромаденин, 7-дезааденин, 7-диазагуанин, 8-оксо-гуанин, 5,6-дигидроуридин и 5-гидроксиметилуридин. Эти синтетические молекулы являются коммерчески доступными (например, могут быть закуплены у Glen Research Company) и могут быть включены в ДНК посредством химического синтеза.

[0256] Примерами модификации сахарной молекулы являются 3'-дезоксилирование, 2'-фторирование и арабаноцидирование, однако они не рассматриваются как ограничивающие. Включение этих модификаций в ДНК может быть также осуществлено путем химического синтеза.

[0257] Примерами модификации 5'-конца являются 5'-аминирование, 5'-биотинилирование, 5'-флуоресцеинилирование, 5'-тетрафторфлуоресцеинилирование, 5'-тионирование и 5'-дабсилирование, однако эти модификации не рассматриваются как ограничивающие.

[0258] Примерами модификации 3'-конца являются 3'-аминирование, 3'-биотинилирование, 2,3-дидезоксилирование, 3'-тионирование, 3'-дабсилирование, 3'-карбоксилирование и 3'-холестерилирование, однако, эти модификации не рассматриваются как ограничивающие.

[0259] В одном предпочтительном варианте осуществления изобретения, олигонуклеооснование может содержать 5'-блокирующий заместитель, который присоединен к 5'-концевым атомам углерода посредством линкера. Химическая природа линкера не имеет существенного значения, за исключением того, что, в некоторых вариантах, его длина должна составлять по меньшей мере 6 атомов, и этот линкер должен

быть гибким. При этом могут быть использованы различные нетоксичные заместители, такие как биотин, холестерин или другие стероиды или неинтеркалирующий катионный флуоресцентный краситель. Особенно предпочтительными реагентами для получения олигонуклеоснований являются реагенты, поставляемые как $Cy3^{TM}$ и $Cy5^{TM}$ компанией Glen Research, Sterling VA (ныне называемой GE Healthcare), которые представляют собой блокированные фосфоамидиты, которые при включении в олигонуклеотид дают 3,3,3',3'-тетраметил-N, N'-изопропил-замещенные индомонокарбоцианиновые и индодикарбоцианиновые красители, соответственно. $Cy3$ является наиболее предпочтительным. Если индокарбоцианин замещен N-оксиалкилом, то он может легко связываться с 5'-концом олигодезоксинуклеотида посредством фосфодиэфира, имеющего 5'-концевой фосфат. Если используется коммерчески доступный $Cy3$ -фосфоамидит, то полученная 5'-модификация состоит из блокирующего заместителя и линкера, которые вместе представляют собой N-гидроксипропил, N'-фосфатидилпропил-3,3,3',3'-тетраметилиндомонокарбоцианин. Другие рассматриваемые красители включают родамин 6G, тетраметилродамин, сульфородамин 101, мероцианин 540, Atto565, Atto550 26, $Cy3.5$, Dy547, Dy548, Dy549, Dy554, Dy555, Dy556, Dy560, mStrawberry и mCherry.

[0260] В предпочтительном варианте осуществления изобретения, индокарбоцианиновый краситель является тетразамещенным в 3- и 3'-положениях индольных колец. Не ограничиваясь какой-либо конкретной теорией, авторы лишь отмечают, что эти заместители не позволяют красителю быть интеркалирующим красителем. Идентичность заместителей в этих положениях не играет важной роли.

[0261] Описанные здесь олигонуклеотидные конструкции могут быть также использованы в качестве более эффективных донорных матриц в комбинации с другими методами редактирования ДНК или рекомбинации, включая, но не ограничиваясь ими, таргетинг гена посредством сайт-специфической гомологичной рекомбинации под действием нуклеаз «цинковый палец», мегануклеаз, эффекторных нуклеаз, подобных активатору транскрипции (TALEN), или регулярно кластеризованных перемежающихся коротких палиндромных повторов (CRISPR).

[0262] В некоторых своих аспектах и вариантах, настоящее изобретение может включать способы и композиции, а именно, способы эффективной модификации геномной клеточной ДНК и/или рекомбинации ДНК в геномную ДНК клетки. Хотя некоторые описанные здесь способы не ограничиваются каким-либо конкретным применением, однако, в некоторых вариантах осуществления изобретения, они могут быть применены, например, для введения модификации в геном клетки в целях определения влияния модификации на клетку. Так, например, модификация может быть введена в нуклеотидную последовательность, кодирующую фермент, для того, чтобы определить, изменяет ли такая модификация ферментативную активность фермента, и/или определить местоположение каталитической области фермента. Альтернативно, модификация может быть введена в последовательность, кодирующую ДНК-связывающий белок для того, чтобы определить, изменяется ли ДНК-связывающая

активность белка, и, таким образом, определить конкретную ДНК-связывающую область внутри белка. Еще одной альтернативой является введение модификации в некодирующую регуляторную последовательность (например, в промоторную, энхансерную, регуляторную последовательность РНК (миРНК) и т.п. для определения влияния модификации на уровень экспрессии второй последовательности, которая функционально присоединена к некодирующей регуляторной последовательности. Это может быть желательным, например, для определения конкретной последовательности, которая обладает регуляторной активностью.

ДНК-разрезающие ферменты

[0263] Одной из стратегий получения нацеленной дизрупции гена является продуцирование одноцепочечных или двухцепочечных разрывов ДНК с использованием ДНК-разрезающего фермента, такого как сайт-специфическая эндонуклеаза. Эндонуклеазы чаще всего используются для нацеленной дизрупции гена в организмах, которые традиционно являются чувствительными к более традиционным механизмам таргетинга генов, таких как водоросли, растения и крупные животные-модели, включая человека. Так, например, в настоящее время проводятся клинические испытания с участием человека с использованием нуклеаз «цинковые пальцы» для лечения и профилактики ВИЧ-инфекции. Кроме того, в настоящее время проводится конструирование эндонуклеаз в попытке дизрупции генов, которые дают нежелательные фенотипы в сельскохозяйственных культурах.

[0264] В некоторых своих аспектах, настоящее изобретение относится к введению одной или более мутаций в нуклеиновую кислоту-мишень с использованием ДНК-эндонуклеазы. В некоторых вариантах осуществления изобретения, ДНК-эндонуклеаза представляет собой ДНК-эндонуклеазу, зависимую от РНК. Репрезентативными ДНК-эндонуклеазами, зависимыми от РНК, являются Cas9, Cpf1 и т.п. ДНК-эндонуклеазы, зависимые от РНК, являются подходящими для их применения в описанных здесь способах и композициях, как очевидно для специалиста в данной области. Дополнительные ДНК-эндонуклеазы для их использования в способах и композициях согласно изобретению описаны в настоящей заявке.

Цинковые пальцы

[0265] Одним из классов искусственных эндонуклеаз является эндонуклеаза «цинковые пальцы». Эндонуклеазы «цинковые пальцы» объединяют неспецифический домен расщепления, обычно домен эндонуклеазы FokI, и домены белка «цинковый палец», которые были сконструированы для связывания со специфическими последовательностями ДНК. Модульная структура эндонуклеаз «цинковые пальцы» делает их универсальной платформой для доставки сайт-специфических двухцепочечных разрывов в геном. Поскольку эндонуклеаза FokI расщепляется в виде димера, то одной из стратегий предотвращения событий расщепления вне мишени является создание доменов «цинковые пальцы», которые связываются с соседними 9 сайтами пар оснований. См. также патенты США №№ 7285416; 7521241; 7361635; 7273923; 7262054; 7220719;

7070934; 7013219; 6979539; 6933113; 6824978; каждый из которых включен в настоящее описание в полном объеме посредством ссылки.

TALEN

[0266] TALEN представляют собой нацеливаемые нуклеазы, которые используются для индуцирования одно- и двухцепочечных разрывов в специфических сайтах ДНК, которые затем репарируются по механизмам, которые можно использовать для введения модификаций в последовательности в сайте расщепления.

[0267] Фундаментальный структурный блок, который используется для конструирования ДНК-связывающей области TALEN, представляет собой высококонсервативный повторяющийся домен, происходящий от встречающихся в природе TALE, кодируемых протеобактериями *Xanthomonas* spp. Связывание ДНК посредством TALEN опосредуется массивами их высококонсервативных 33-35 аминокислотных повторов, которые фланкируются дополнительными происходящими от TALE доменами у амино- и карбокси-концов повторов.

[0268] Эти повторы TALE специфически связываются с одним основанием ДНК, идентичность которого определяется двумя гипервариабельными остатками, обычно находящимися в положениях 12 и 13 повтора, причем, число повторов в массиве соответствует длине желаемой нуклеиновой кислоте-мишени, а идентичность повтора выбирают так, чтобы она соответствовала последовательности нуклеиновой кислоты-мишени. В некоторых вариантах осуществления изобретения, нуклеиновая кислота-мишень имеет от 15 до 20 пар оснований, что позволяет максимизировать селективность сайта-мишени. Расщепление нуклеиновой кислоты-мишени обычно происходит в пределах 50 пар оснований сайта связывания с TALEN. Компьютерные программы для конструирования сайта распознавания TALEN описаны в литературе. См., например, Cermak et al., *Nucleic Acids Res.* 2011 July; 39 (12): c82.

[0269] TALEN, после их конструирования так, чтобы они соответствовали нужной последовательности-мишени, могут быть экспрессированы рекомбинантно и введены в протопласты в виде экзогенных белков, либо они могут быть экспрессированы из плазмиды внутри протопласта или введены в виде мРНК или белка.

Мегануклеазы

[0270] Хоминг-эндонуклеазы, также известные как мегануклеазы, представляют собой последовательность-специфические эндонуклеазы, которые генерируют двухцепочечные разрывы в геномной ДНК с высокой степенью специфичности благодаря их крупным (например >14 п.о.) сайтам расщепления. Хотя специфичность хоминг-эндонуклеаз к своим сайтам-мишеням позволяет осуществлять точную доставку в индуцированные ДНК-разрывы, однако, сайты расщепления хоминг-эндонуклеазы являются редкими, и вероятность обнаружения встречающегося в природе сайта расщепления в целевом гене является низкой.

[0271] Другим классом искусственных эндонуклеаз являются сконструированные мегануклеазы. Сконструированные хоминг-эндонуклеазы получают путем модификации

специфичности уже существующих хоминг-эндонуклеаз. В одном подходе, в аминокислотную последовательность природных хоминг-эндонуклеаз вводят модификации, а затем полученные сконструированные хоминг-эндонуклеазы скринируют для отбора функциональных белков, которые расщепляют целевой сайт связывания. В соответствии с другим подходом, химерные хоминг-эндонуклеазы конструируют путем объединения сайтов распознавания двух различных хоминг-эндонуклеаз для создания нового сайта распознавания, состоящего из половины сайта каждой хоминг-эндонуклеазы. См., например, патенты США №№ 8338157 и 8445251.

Системы CRISPR или CRISPR/cas

[0272] Система CRISPR-Cas содержит три основных компонента конструкции: 1) ген Cas, транскрипт (например, мРНК) или белок; 2) руководящую РНК (gРНК); и 3) cr-РНК (РНК CRISPR), которые представляют собой сегменты РНК, процессированные из РНК-транскриптов, кодирующих массивы повторов CRISPR, которые содержат «протоспейсерную область», комплементарную чужеродному сайту ДНК (например, эндогенной области-мишени ДНК) и части повтора CRISPR. См., например, заявки РСТ № WO/2014/093661 и WO/2013/176772.

Ген Cas (ассоциированный с CRISPR), транскрипт (например, мРНК) или белок

[0273] Временная экспрессия Cas из плазмидного вектора непосредственно регулирует доставку белка Cas и/или мРНК Cas в клетки растения. Гены Cas оптимизированы по кодонам для экспрессии в высших растениях, водорослях или в дрожжах и индуцируются конститутивным, индуцибельным, тканеспецифическим или видоспецифическим промотором, когда это возможно. Сигналы терминации транскрипта Cas и полиаденилирования представляют собой либо NosT, RBCT, HSP18.2T, либо другие геноспецифические или видоспецифические терминаторы. Генные кластеры Cas могут содержать интроны, либо нативные, либо в комбинации с геноспецифическими промоторами и/или синтетическими промоторами. Белок Cas может содержать одну или более последовательностей сигнала локализации в ядре (NLS), мутации, делеции, модификации или усечения. Во временных экспрессионных системах, генные кластеры Cas могут быть объединены с другими компонентами системы CRISPR-Cas, такими как кластеры gРНК, на одном и том же временном экспрессионном векторе. Альтернативно, генные кластеры Cas могут быть локализованы и экспрессированы из конструкций, не зависящих от кластеров gРНК, или из других компонентов системы CRISPR-Cas. CRISPR-ассоциированный (Cas) ген кодирует белки с различными предсказанными активностями, модифицирующими нуклеиновую кислоту, такими как нуклеазы, геликазы и полимеразы. Гены Cas включают Cas9. Cas9 представляет собой ген, кодирующий крупный белок, содержащий предсказанные RuvC-подобные домены и домены эндонуклеазы HNH, и ассоциируется с адаптивной иммунной системой CRISPR, которая присутствует у большинства археобактерий и многих других бактерий. Белок Cas9 состоит из двух частей:

- 1) сайта распознавания (REC), состоящего из трех доменов:
 - а) ВН (мостиковой спирали)

b) REC1, облегчающего таргетинг ДНК, зависимой от РНК

c) REC2, облегчающего таргетинг ДНК, зависимой от РНК

2) Сайта нуклеазы (NUC), состоящего из трех доменов:

a) RuvC, облегчающего таргетинг ДНК, зависимой от РНК, и обладающего эндонуклеазной активностью

b) Домена HNH, обладающего эндонуклеазной активностью

c) Домена PI, взаимодействующего с РАМ

[0274] В других вариантах осуществления изобретения, ген Cas может представлять собой гомолог Cas9, в котором домены RuvC, HNH, REC и NH являются в высокой степени консервативными. В некоторых вариантах осуществления изобретения, гены Cas представляют собой гены, происходящие от нижеследующих видов, перечисленных в Таблице В.

Таблица В: Репрезентативные гены Cas

Локус ID/ GI	Виды	Профиль Cas ID	Ген Cas
352684361	Acidaminococcus_intestini_RyC_MR95_uid74445	mkCas019 3	cas9
117929158	Acidothermus_cellulolyticus_11B_uid58501	cd09643	cas9
326315085	Acidovorax_avenae_ATCC_19860_uid42497	cd09643	cas9
222109285	Acidovorax_ebreus_TPSY_uid59233	COG3513	cas9
152978060	Actinobacillus_succinogenes_130Z_uid58247	COG3513	cas9
407692091	Actinobacillus_suis_H91_0380_uid176363	COG3513	cas9
187736489	Akkermansia_muciniphila_ATCC_BAA_835_uid58985	cd09643	cas9
319760940	Alicyclophilus_denitrificans_BC_uid49953	cd09643	cas9
330822845	Alicyclophilus_denitrificans_K601_uid66307	cd09643	cas9
288957741	Azospirillum_B510_uid46085	cd09643	cas9
549484339	Bacteroides_CF50_uid222805	cd09643, COG3513	cas9
375360193	Bacteroides_fragilis_638R_uid84217	COG3513, COG3513	cas9
60683389	Bacteroides_fragilis_NCTC_9343_uid57639	COG3513, COG3513	cas9
471261880	Bdellovibrio_exovorus_JSS_uid194119	COG3513	cas9
390944707	Belliella_baltica_DSM_15883_uid168182	cd09643, COG3513	cas9

470166767	<i>Bibersteinia_trehalosi_192_uid193709</i>	COG3513	cas9
310286728	<i>Bifidobacterium_bifidum_S17_uid59545</i>	mkCas019 3	cas9
283456135	<i>Bifidobacterium_dentium_Bd1_uid43091</i>	cd09643	cas9
189440764	<i>Bifidobacterium_longum_DJO10A_uid58833</i>	cd09643	cas9
384200944	<i>Bifidobacterium_longum_KACC_91563_uid158861</i>	cd09643	cas9
479188345	<i>Butyrivibrio_fibrisolvens_uid197155</i>	cd09643	cas9
544063172	<i>Campylobacter_jejuni_00_2425_uid219359</i>	COG3513	cas9
543948719	<i>Campylobacter_jejuni_00_2426_uid219324</i>	COG3513	cas9
543946932	<i>Campylobacter_jejuni_00_2538_uid219325</i>	COG3513	cas9
543950499	<i>Campylobacter_jejuni_00_2544_uid219326</i>	COG3513	cas9
549693479	<i>Campylobacter_jejuni_4031_uid222817</i>	COG3513	cas9
157415744	<i>Campylobacter_jejuni_81116_uid58771</i>	COG3513	cas9
384448746	<i>Campylobacter_jejuni_IA3902_uid159531</i>	COG3513	cas9
384442102	<i>Campylobacter_jejuni_M1_uid159535</i>	COG3513	cas9
384442103	<i>Campylobacter_jejuni_M1_uid159535</i>	COG3513	cas9
403056243	<i>Campylobacter_jejuni_NCTC_11168_BN148_uid174152</i>	COG3513	cas9
218563121	<i>Campylobacter_jejuni_NCTC_11168_ATCC_700819_uid57587</i>	COG3513	cas9
407942868	<i>Campylobacter_jejuni_PT14_uid176499</i>	COG3513	cas9
153952471	<i>Campylobacter_jejuni_doylei_269_97_uid58671</i>	COG3513	cas9
294086111	<i>Candidatus_Puniceispirillum_marinum_IMCC1322_uid47081</i>	cd09643	cas9
340622236	<i>Capnocytophaga_canimorsus_Cc5_uid70727</i>	COG3513, cd09643	cas9
220930482	<i>Clostridium_cellulolyticum_H10_uid58709</i>	COG3513	cas9
479136975	<i>Coprococcus_catus_GD_7_uid197174</i>	mkCas019 3	cas9
328956315	<i>Coriobacterium_glomerans_PW2_uid65787</i>	mkCas019 3	cas9
375289763	<i>Corynebacterium_diphtheriae_241_uid83607</i>	cd09643	cas9
376283539	<i>Corynebacterium_diphtheriae_31A_uid84309</i>	cd09643	cas9
376286566	<i>Corynebacterium_diphtheriae_BH8_uid84311</i>	cd09643	cas9

376289243	<i>Corynebacterium diphtheriae</i> _C7__beta__uid84313	cd09643	cas9
376244596	<i>Corynebacterium diphtheriae</i> _HC01_uid84297	cd09643	cas9
376292154	<i>Corynebacterium diphtheriae</i> _HC02_uid84317	cd09643	cas9
38232678	<i>Corynebacterium diphtheriae</i> _NCTC_13129_uid57691	cd09643	cas9
376256051	<i>Corynebacterium diphtheriae</i> _VA01_uid84305	cd09643	cas9
159042956	<i>Dinoroseobacter shibae</i> _DFL_12_uid58707	cd09643	cas9
339445983	<i>Eggerthella</i> _YY7918_uid68707	mkCas019 3	cas9
187250660	<i>Elusimicrobium minutum</i> _Pei191_uid58949	cd09643	cas9
479180325	<i>Enterococcus</i> _7L76_uid197170	cd09643	cas9
397699066	<i>Enterococcus faecalis</i> _D32_uid171261	mkCas019 3	cas9
384512368	<i>Enterococcus faecalis</i> _OG1RF_uid54927	mkCas019 3	cas9
392988474	<i>Enterococcus hirae</i> _ATCC_9790_uid70619	mkCas019 3	cas9
558685081	<i>Enterococcus mundtii</i> _QU_25_uid229420	mkCas019 3	cas9
238924075	<i>Eubacterium rectale</i> _ATCC_33656_uid59169	cd09643	cas9
385789535	<i>Fibrobacter succinogenes</i> _S85_uid161919	cd09643, cd09643	cas9
261414553	<i>Fibrobacter succinogenes</i> _S85_uid41169	cd09643, cd09643	cas9
374307738	<i>Filifactor alocis</i> _ATCC_35896_uid46625	mkCas019 3	cas9
169823755	<i>Fingoldia magna</i> _ATCC_29328_uid58867	mkCas019 3	cas9
150025575	<i>Flavobacterium psychrophilum</i> _JIP02_86_uid61627	cd09643, cd09643	cas9
327405121	<i>Fluviicola taffensis</i> _DSM_16823_uid65271	cd09643, cd09643	cas9
387824704	<i>Francisella</i> _cf__novicida_3523_uid162107	cd09704	cas9
118497352	<i>Francisella novicida</i> _U112_uid58499	cd09704	cas9

134302318	<i>Francisella_tularensis_WY96_3418_uid58811</i>	cd09704	cas9
89256630	<i>Francisella_tularensis_holarctica_LVS_uid58595</i>	cd09704	cas9
89256631	<i>Francisella_tularensis_holarctica_LVS_uid58595</i>	cd09704	cas9
534508854	<i>Fusobacterium_3_1_36A2_uid55995</i>	mkCas019 3	cas9
530600688	<i>Geobacillus_JF8_uid215234</i>	COG3513	cas9
209542524	<i>Gluconacetobacter_diazotrophicus_PAI_5_uid59075</i>	COG3513	cas9
162147907	<i>Gluconacetobacter_diazotrophicus_PAI_5_uid61587</i>	COG3513	cas9
479173968	<i>Gordonibacter_pamelaee_7_10_1_b_uid197167</i>	mkCas019 3	cas9
345430422	<i>Haemophilus_parainfluenzae_T3T1_uid72801</i>	COG3513	cas9
471315929	<i>Helicobacter_cinaedi_ATCC_BAA_847_uid193765</i>	COG3513	cas9
386762035	<i>Helicobacter_cinaedi_PAGU611_uid162219</i>	COG3513	cas9
291276265	<i>Helicobacter_mustelae_12198_uid46647</i>	COG3513	cas9
385811609	<i>Ignavibacterium_album_JCM_16511_uid162097</i>	cd09643, COG3513	cas9
310780384	<i>Ilyobacter_polytropus_DSM_2926_uid59769</i>	COG3513	cas9
331702228	<i>Lactobacillus_buchneri_NRRL_B_30929_uid66205</i>	mkCas019 3	cas9
406027703	<i>Lactobacillus_buchneri_uid73657</i>	mkCas019 3	cas9
385824065	<i>Lactobacillus_casei_BD_II_uid162119</i>	mkCas019 3	cas9
191639137	<i>Lactobacillus_casei_BL23_uid59237</i>	mkCas019 3	cas9
385820880	<i>Lactobacillus_casei_LC2W_uid162121</i>	mkCas019 3	cas9
523514789	<i>Lactobacillus_casei_LOCK919_uid210959</i>	mkCas019 3	cas9
409997999	<i>Lactobacillus_casei_W56_uid178736</i>	mkCas019 3	cas9
301067199	<i>Lactobacillus_casei_Zhang_uid50673</i>	mkCas019 3	cas9
385815562	<i>Lactobacillus_delbrueckii_bulgaricus_2038_uid1619</i>	mkCas019	cas9

	29	3	
385815563	Lactobacillus_delbrueckii_bulgaricus_2038_uid1619 29	mkCas019 3	cas9
385815564	Lactobacillus_delbrueckii_bulgaricus_2038_uid1619 29	mkCas019 3	cas9
385826041	Lactobacillus_johnsonii_DPC_6026_uid162057	mkCas019 3	cas9
532357525	Lactobacillus_paracasei_8700_2_uid55295	mkCas019 3	cas9
448819853	Lactobacillus_plantarum_ZJ316_uid188689	mkCas019 3	cas9
385828839	Lactobacillus_rhamnosus_GG_uid161983	mkCas019 3	cas9
258509199	Lactobacillus_rhamnosus_GG_uid59313	mkCas019 3	cas9
523517690	Lactobacillus_rhamnosus_LOCK900_uid210957	mkCas019 3	cas9
385839898	Lactobacillus_salivarius_CECT_5713_uid162005	mkCas019 3	cas9
385839899	Lactobacillus_salivarius_CECT_5713_uid162005	mkCas019 3	cas9
385839900	Lactobacillus_salivarius_CECT_5713_uid162005	mkCas019 3	cas9
90961083	Lactobacillus_salivarius_UCC118_uid58233	mkCas019 3	cas9
90961084	Lactobacillus_salivarius_UCC118_uid58233	mkCas019 3	cas9
347534532	Lactobacillus_sanfranciscensis_TMW_1_1304_uid7 2937	mkCas019 3	cas9
54296138	Legionella_pneumophila_Paris_uid58211	cd09704	cas9
406600271	Leuconostoc_gelidum_JB7_uid175682	mkCas019 3	cas9
16801805	Listeria_innocua_Clip11262_uid61567	cd09643, COG3513	cas9

386044902	Listeria_monocytogenes_10403S_uid54461	COG3513, COG3513	cas9
550898770	Listeria_monocytogenes_EGD_uid223288	COG3513, COG3513	cas9
386048324	Listeria_monocytogenes_J0161_uid54459	COG3513, COG3513	cas9
405756714	Listeria_monocytogenes_SLCC2540_uid175106	COG3513, COG3513	cas9
404411844	Listeria_monocytogenes_SLCC5850_uid175110	COG3513, COG3513	cas9
404282159	Listeria_monocytogenes_serotype_1_2b_SLCC2755_uid52455	COG3513, COG3513	cas9
404287973	Listeria_monocytogenes_serotype_7_SLCC2482_uid174871	COG3513, COG3513	cas9
433625054	Mycoplasma_cynos_C142_uid184824	cd09643	cas9
401771107	Mycoplasma_gallisepticum_CA06_2006_052_5_2P_uid172630	cd09643	cas9
385326554	Mycoplasma_gallisepticum_F_uid162001	cd09643	cas9
401767318	Mycoplasma_gallisepticum_NC95_13295_2_2P_uid172625	cd09643	cas9
401768090	Mycoplasma_gallisepticum_NC96_1596_4_2P_uid172626	cd09643	cas9
401768851	Mycoplasma_gallisepticum_NY01_2001_047_5_1P_uid172627	cd09643	cas9
385325798	Mycoplasma_gallisepticum_R_high_uid161999	cd09643	cas9
294660600	Mycoplasma_gallisepticum_R_low_uid57993	cd09643	cas9
565627373	Mycoplasma_gallisepticum_S6_uid200523	cd09643	cas9
401769598	Mycoplasma_gallisepticum_WI01_2001_043_13_2P_uid172628	cd09643	cas9
47458868	Mycoplasma_mobile_163K_uid58077	cd09643	cas9
71894592	Mycoplasma_synoviae_53_uid58061	cd09643	cas9
313669044	Neisseria_lactamica_020_06_uid60851	COG3513	cas9
161869390	Neisseria_meningitidis_053442_uid58587	COG3513	cas9
385324780	Neisseria_meningitidis_8013_uid161967	COG3513	cas9

385337435	<i>Neisseria meningitidis</i> _WUE_2594_uid162093	COG3513	cas9
218767588	<i>Neisseria meningitidis</i> _Z2491_uid57819	COG3513	cas9
254804356	<i>Neisseria meningitidis</i> _alpha14_uid61649	COG3513	cas9
319957206	<i>Nitratifactor salsuginis</i> _DSM_16511_uid62183	cd09643	cas9
325983496	<i>Nitrosomonas</i> _AL212_uid55727	COG3513	cas9
302336020	<i>Olsenella uli</i> _DSM_7084_uid51367	mkCas019 3	cas9
392391493	<i>Ornithobacterium rhinotracheale</i> _DSM_15997_uid168256	cd09643	cas9
154250555	<i>Parvibaculum lavamentivorans</i> _DS_1_uid58739	cd09643	cas9
15602992	<i>Pasteurella multocida</i> _Pm70_uid57627	COG3513	cas9
557607382	<i>Pediococcus pentosaceus</i> _SL4_uid227215	mkCas019 3	cas9
294674019	<i>Prevotella ruminicola</i> _23_uid47507	COG3513	cas9
408489713	<i>Psychroflexus torquis</i> _ATCC_700755_uid54205	cd09643, cd09643	cas9
90425961	<i>Rhodopseudomonas palustris</i> _BisB18_uid58443	COG3513	cas9
91975509	<i>Rhodopseudomonas palustris</i> _BisB5_uid58441	COG3513	cas9
83591793	<i>Rhodospirillum rubrum</i> _ATCC_11170_uid57655	cd09643	cas9
386348484	<i>Rhodospirillum rubrum</i> _F11_uid162149	cd09643	cas9
383485594	<i>Riemerella anatipestifer</i> _ATCC_11845__DSM_15868_uid159857	COG3513, cd09643	cas9
407451859	<i>Riemerella anatipestifer</i> _RA_CH_1_uid175469	COG3513, cd09643	cas9
442314523	<i>Riemerella anatipestifer</i> _RA_CH_2_uid186548	COG3513, cd09643	cas9
386321727	<i>Riemerella anatipestifer</i> _RA_GD_uid162013	COG3513, cd09643	cas9
479204792	<i>Roseburia intestinalis</i> _uid197164	COG3513	cas9
470213512	<i>Sphingomonas</i> _MM_1_uid193771	COG3513	cas9
325972003	<i>Spirochaeta Buddy</i> _uid63633	cd09643	cas9
563693590	<i>Spiroplasma apis</i> _B31_uid230613	cd09643	cas9
507384108	<i>Spiroplasma syrphidicola</i> _EA_1_uid205054	cd09643	cas9
556591142	<i>Staphylococcus pasteurii</i> _SP1_uid226267	cd09643	cas9

386318630	Staphylococcus_pseudintermedius_ED99_uid162109	mkCas0193	cas9
269123826	Streptobacillus_moniliformis_DSM_12112_uid41863	COG3513	cas9
552737657	Streptococcus_I_G2_uid224251	cd09643	cas9
512539130	Streptococcus_agalactiae_09mas018883_uid208674	mkCas0193	cas9
22537057	Streptococcus_agalactiae_2603V_R_uid57943	mkCas0193	cas9
494703075	Streptococcus_agalactiae_2_22_uid202215	mkCas0193	cas9
76788458	Streptococcus_agalactiae_A909_uid57935	mkCas0193	cas9
406709383	Streptococcus_agalactiae_GD201008_001_uid175780	mkCas0193	cas9
512544670	Streptococcus_agalactiae_ILRI005_uid208676	mkCas0193	cas9
512698372	Streptococcus_agalactiae_ILRI112_uid208675	mkCas0193	cas9
25010965	Streptococcus_agalactiae_NEM316_uid61585	mkCas0193	cas9
410594450	Streptococcus_agalactiae_SA20_06_uid178722	mkCas0193	cas9
538370328	Streptococcus_anginosus_C1051_uid218003	cd09643	cas9
410494913	Streptococcus_dysgalactiae_equisimilis_AC_2713_uid178644	COG3513	cas9
386317166	Streptococcus_dysgalactiae_equisimilis_ATCC_12394_uid161979	COG3513	cas9
251782637	Streptococcus_dysgalactiae_equisimilis_GGS_124_uid59103	COG3513	cas9
408401787	Streptococcus_dysgalactiae_equisimilis_RE378_uid176684	COG3513	cas9
195978435	Streptococcus_equi_zooepidemicus_MGCS10565_uid59263	COG3513	cas9

386338081	Streptococcus_gallolyticus_ATCC_43143_uid16210 3	cd09643	cas9
386338091	Streptococcus_gallolyticus_ATCC_43143_uid16210 3	mkCas019 3	cas9
325978669	Streptococcus_gallolyticus_ATCC_BAA_2069_uid6 3617	mkCas019 3	cas9
288905632	Streptococcus_gallolyticus_UCN34_uid46061	cd09643	cas9
288905639	Streptococcus_gallolyticus_UCN34_uid46061	mkCas019 3	cas9
157150687	Streptococcus_gordonii_Challis_substr_CH1_uid5 7667	cd09643	cas9
379705580	Streptococcus_infantarius_CJ18_uid87033	mkCas019 3	cas9
508127396	Streptococcus_iniae_SF1_uid206041	mkCas019 3	cas9
508127399	Streptococcus_iniae_SF1_uid206041	COG3513	cas9
538379999	Streptococcus_intermedius_B196_uid218000	cd09643	cas9
527330434	Streptococcus_lutetiensis_033_uid213397	mkCas019 3	cas9
374338350	Streptococcus_macedonicus_ACA_DC_198_uid816 31	cd09643	cas9
397650022	Streptococcus_mutans_GS_5_uid169223	mkCas019 3	cas9
387785882	Streptococcus_mutans_LJ23_uid162197	mkCas019 3	cas9
290580220	Streptococcus_mutans_NN2025_uid46353	mkCas019 3	cas9
24379809	Streptococcus_mutans_UA159_uid57947	mkCas019 3	cas9
336064611	Streptococcus_pasteurianus_ATCC_43144_uid6801 9	cd09643	cas9
410680443	Streptococcus_pyogenes_A20_uid178106	COG3513	cas9
470200927	Streptococcus_pyogenes_M1_476_uid193766	COG3513	cas9
15675041	Streptococcus_pyogenes_M1_GAS_uid57845	COG3513	cas9

94990395	<i>Streptococcus pyogenes</i> _MGAS10270_uid58571	COG3513	cas9
94994317	<i>Streptococcus pyogenes</i> _MGAS10750_uid58575	COG3513	cas9
383479946	<i>Streptococcus pyogenes</i> _MGAS15252_uid158037	COG3513	cas9
383493861	<i>Streptococcus pyogenes</i> _MGAS1882_uid158061	COG3513	cas9
94992340	<i>Streptococcus pyogenes</i> _MGAS2096_uid58573	COG3513	cas9
21910213	<i>Streptococcus pyogenes</i> _MGAS315_uid57911	COG3513	cas9
71910582	<i>Streptococcus pyogenes</i> _MGAS5005_uid58337	COG3513	cas9
71903413	<i>Streptococcus pyogenes</i> _MGAS6180_uid58335	COG3513	cas9
94988516	<i>Streptococcus pyogenes</i> _MGAS9429_uid58569	COG3513	cas9
209559356	<i>Streptococcus pyogenes</i> _NZ131_uid59035	COG3513	cas9
28896088	<i>Streptococcus pyogenes</i> _SSI_1_uid57895	COG3513	cas9
387783792	<i>Streptococcus salivarius</i> _JIM8777_uid162145	cd09643	cas9
386584496	<i>Streptococcus suis</i> _D9_uid162125	cd09643	cas9
389856936	<i>Streptococcus suis</i> _ST1_uid167482	mkCas019 3	cas9
330833104	<i>Streptococcus suis</i> _ST3_uid66327	cd09643	cas9
55822627	<i>Streptococcus thermophilus</i> _CNRZ1066_uid58221	cd09643	cas9
386344353	<i>Streptococcus thermophilus</i> _JIM_8232_uid162157	cd09643	cas9
116627542	<i>Streptococcus thermophilus</i> _LMD_9_uid58327	cd09643	cas9
116628213	<i>Streptococcus thermophilus</i> _LMD_9_uid58327	mkCas019 3	cas9
55820735	<i>Streptococcus thermophilus</i> _LMG_18311_uid5821 9	cd09643	cas9
387909441	<i>Streptococcus thermophilus</i> _MN_ZLW_002_uid16 6827	cd09643	cas9
387910220	<i>Streptococcus thermophilus</i> _MN_ZLW_002_uid16 6827	mkCas019 3	cas9
386086348	<i>Streptococcus thermophilus</i> _ND03_uid162015	cd09643	cas9
386087120	<i>Streptococcus thermophilus</i> _ND03_uid162015	mkCas019 3	cas9
389874754	<i>Tistrella mobilis</i> _KA081020_065_uid167486	COG3513	cas9
42525843	<i>Treponema denticola</i> _ATCC_35405_uid57583	mkCas019 3	cas9
530892607	<i>Treponema pedis</i> _T_A4_uid215715	COG3513,	cas9

		COG3513	
121608211	<i>Verminephrobacter_eiseniae</i> _EF01_2_uid58675	cd09643	cas9
525888882	<i>Vibrio_paraahaemolyticus</i> _O1_K33_CDC_K4557_uid212977	COG3513, COG3513	cas9
525913263	<i>Vibrio_paraahaemolyticus</i> _O1_K33_CDC_K4557_uid212977	COG3513	cas9
525919586	<i>Vibrio_paraahaemolyticus</i> _O1_K33_CDC_K4557_uid212977	COG3513, COG3513	cas9
525927253	<i>Vibrio_paraahaemolyticus</i> _O1_K33_CDC_K4557_uid212977	COG3513, COG3513	cas9
325955459	<i>Weeksella_virosa</i> _DSM_16922_uid63627	cd09643, cd09643	cas9
34557790	<i>Wolinella_succinogenes</i> _DSM_1740_uid61591	cd09643	cas9
34557932	<i>Wolinella_succinogenes</i> _DSM_1740_uid61591	cd09704	cas9
295136244	<i>Zunongwangia_profunda</i> _SM_A87_uid48073	COG3513, cd09643	cas9
304313029	<i>gamma_proteobacterium</i> _HdN1_uid51635	cd09643	cas9
189485058	uncultured_Termite_group_1_bacterium_phylotype_Rs_D17_uid59059	cd09643	cas9
189485059	uncultured_Termite_group_1_bacterium_phylotype_Rs_D17_uid59059	cd09643	cas9
189485225	uncultured_Termite_group_1_bacterium_phylotype_Rs_D17_uid59059	COG3513	cas9
347536497	<i>Flavobacterium_branchiophilum</i> _FL_15_uid73421	COG3513, cd09643, COG3513	cas9,cas9
365959402	<i>Flavobacterium_columnare</i> _ATCC_49512_uid80731	COG3513, cd09643, COG3513	cas9,cas9
387132277	<i>Prevotella_intermedia</i> _17_uid163151	cd09643, COG3513, COG0188	cas9, топо- изомераз а типа IIA

Руководящая РНК (рРНК)

[0275] рРНК или оцрРНК (одноцепочечная руководящая РНК) сконструирована как гибрид между cr-РНК и частью транскрибирующей последовательности РНК CRISPR (tracr-РНК), которая направляет Cas9 на специфическую последовательность ДНК-мишени, комплементарную протоспейсерной области. Руководящая РНК может включать экспрессионный кластер, содержащий химерную конструкцию РНК с длинным гибридом tracr-РНК, коротким гибридом tracr-РНК или нативным массивом CRISPR с конформацией tracr-РНК. Химерная рРНК обладает специфичностью нацеливания на cr-РНК и свойствами образовывать каркас tracr-РНК в одном транскрипте. Временная экспрессия рРНК регулируется видоспецифическими промоторами РНК-полимеразы III высших растений, такими как промоторы семейства генов мяРНК U6 или U3 (Wang et al., 2008). Терминация транскрипции рРНК регулируется фрагментом из 6-20 нуклеотидов poly-dT как описано Wang et al., 2008. Экспрессионные кластеры рРНК расположены на одних и тех же или различных временных экспрессионных кластерах, происходящих от других компонентов системы CRISPR-Cas. Транскрипты рРНК могут быть синтезированы *in vitro* и доставлены непосредственно в клетки растений отдельно от временных экспрессионных векторов рРНК или в комбинации с ними.

[0276] В некоторых вариантах осуществления изобретения, нативная система CRISPR-Cas типа II *S. pyogenes* состоит из нуклеазы, ассоциированной с Crispr AS (Cas9), и двух различных некодирующих РНК, транскрибирующей РНК (tracr-РНК) и РНК CRISPR (cr-РНК). РНК-компоненты этой системы направляют нуклеазу Cas9 в последовательность-специфическую мишень в геноме. Все три компонента могут быть экспрессированы по отдельности в виде tracr-РНК и cr-РНК и белка Cas9. cr-РНК обеспечивает специфичность к мишени и состоит из спейсерной последовательности из 20 оснований, которые комплементарны ДНК-мишени (протоспейсерной последовательности), которая расщепляется Cas9 (Le Cong et al., 2013). tracr-РНК действует как РНК-каркас, если она ассоциирована с cr-РНК посредством спаривания оснований РНК:РНК, и именно этот комплекс ассоциируется с Cas9. tracr-РНК может быть сконструирована так, чтобы она была короче на 89 оснований, как и в случае системы Alt-RTM, разработанной Integrated DNA Technologies (IDT). В этой системе, tracr-РНК, составляющая всего 67 оснований, обладает повышенной эффективностью по отношению к мишени по сравнению с нативной tracr-РНК. Если cr-РНК и tracr-РНК искусственно объединены в одну гибридную функциональную РНК или в одну руководящую РНК (оцрРНК), то таргетинг белка Cas9 может быть значительно упрощен по сравнению с нативной системой. Подобно нативному комплексу tracr-РНК:cr-РНК, сконструированная оцрРНК направляет Cas9 к специфической последовательности ДНК-мишени.

Область-мишень

[0277] Руководящие РНК содержат два компонента, которые определяют специфичность к области-мишени ДНК, протоспейсер и мотив, смежный с

протоспейсером (PAM). Последовательность протоспейсера обычно имеет длину 20 нуклеотидов, но может варьироваться в зависимости от ДНК-мишени, что обеспечивает специфичность последовательности ДНК к комплексу CRISPR-Cas. ДНК-мишени также содержат тринуклеотидную последовательность NNG или NAG (PAM), где N означает любой нуклеотид, расположенный непосредственно у 3'-конца или ниже протоспейсера.

Однокомпонентный подход

[0278] Как описано у Le Cong et al. (2013) и др., представлен упрощенный «однокомпонентный подход» к редактированию генов CRISPR-Cas, где одна временная экспрессионная конструкция содержит все компоненты комплекса CRISPR-Cas, то есть, экспрессионные кластеры рРНК и Cas. Это обеспечивает получение простой модульной конструкции для нацеливания на один или множество локусов в любом конкретном растении или культуре. Нацеливание на множество локусов может быть достигнуто путем простого обмена мишень-специфическими кластерами рРНК. Кроме того, видоспецифические промоторы, терминаторы или другие экспрессирующие энхансерные элементы могут быть легко перемещены в транзientные векторы для «однокомпонентного подхода», и удалены из них, что будет обеспечивать оптимальную экспрессию рРНК и белка Cas видоспецифическим образом.

Двухкомпонентный подход

[0279] В двухкомпонентном подходе, экспрессионные кластеры Cas и рРНК расположены на различных транзientных экспрессионных векторах. Это позволяет доставлять компоненты редактирования CRISPR-Cas по отдельности, обеспечивая тем самым различные отношения рРНК к Cas в одной и той же клетке. Подобно однокомпонентному подходу, двухкомпонентный подход также позволяет легко модифицировать промоторы, терминаторы или другие элементы, влияющие на экспрессию компонентов CRISPR-Cas, и позволяет нацеливать ДНК видоспецифическим образом.

Антибиотики

[0280] Другим классом эндонуклеаз являются антибиотики, которые представляют собой ДНК-расщепляющие гликопептиды, такие как антибиотики семейства блеомицинов, которые представляют собой ДНК-расщепляющие гликопептиды, включая блеомицин, зеоцин, флеомицин, таллисомицин, пеплеомицин и другие, которые также описаны в настоящей заявке.

Другие ДНК-модифицирующие молекулы

[0281] Другие ДНК-модифицирующие молекулы могут быть использованы для направленной рекомбинации генов. Так, например, пептид-содержащие нуклеиновые кислоты могут быть использованы для индуцирования модификаций в геноме клетки-мишени или клеток (см., например, Ecker, патент США № 5986053, который включен в настоящее описание посредством ссылки). Вкратце, синтетические нуклеотиды, содержащие по меньшей мере частичный пептидный остов, используют для нацеливания на гомологичную геномную нуклеотидную последовательность. После связывания с

двухспиральной ДНК или посредством мутагена, лигированного с пептид-содержащей нуклеиновой кислотой, будет индуцироваться модификация и/или рекомбинация последовательности ДНК-мишени. Специфичность нацеливания определяется степенью гомологии последовательности между нацеливающей последовательностью и геномной последовательностью.

[0282] В некоторых вариантах способов и композиций согласно изобретению, гены (такие как ген SHP) могут быть нацелены с использованием олигонуклеотидов, образующих тройную спираль (TFO). TFO могут быть получены путем синтеза, например, с помощью ПЦР или с использованием устройства для синтеза генов. Кроме того, TFO могут быть выделены из геномной ДНК, если присутствуют подходящие природные последовательности. TFO могут быть использованы в различных способах, включая, например, связывание с мутагеном, таким как, но не ограничиваясь им, псорален или хлорамбуцил (см., например, Havre et al., Proc Nat'l Acad Sci, USA 90). : 7879-7883, 1993; Havre et al., J Virol 67: 7323-7331, 1993; Wang et al., Mol Cell Biol 15: 1759-1768, 1995; Takasugi et al., Proc Nat'l Acad Sci, USA 88: 5602-5606, 1991; Belousov et al., Nucleic Acids Res 25: 3440-3444, 1997). Кроме того, например, TFO могут быть связаны с донорной дуплексной ДНК (см., например, Chan et al., J Biol Chem 272: 11541-11548, 1999). TFO также могут действовать посредством связывания с аффинностью, достаточной для индуцирования репарации с вероятностью ошибки (Wang et al., Science 271: 802-805, 1996).

[0283] Раскрытые здесь способы необязательно ограничены природой или типом используемого ДНК-модифицирующего реагента. Так, например, такие ДНК-модифицирующие реагенты высвобождают радикалы, что приводит к разрыву цепи ДНК. Альтернативно, реагенты алкилируют ДНК с образованием аддуктов, которые блокируют репликацию и транскрипцию. В другом альтернативном варианте, реагенты генерируют перекрестные связи или молекулы, которые ингибируют клеточные ферменты, приводящие к разрывам цепей. Примеры ДНК-модифицирующих реагентов, которые были связаны с олигонуклеотидами с образованием TFO, включают, но не ограничиваются ими, индолкарбазолы, нафталиндиимид (NDI), трансплатин, блеомицин, аналоги циклопропапирролоиндола и фенантодигидродоксины. В частности, индолкарбазолы являются ингибиторами топоизомеразы I. Ингибирование этих ферментов приводит к разрыву цепи и образованию продукта присоединения белка к ДНК (Arimondo et al., Bioorganic and Medicinal Chem. 8, 777-784, 2000). NDI является фотооксидантом, который может окислять гуанины, которые могут вызывать мутации в сайтах остатков гуанина (Nunez et al., Biochemistry, 39, 6190-6199, 2000). Было показано, что трансплатин реагирует с ДНК в триплексе-мишени, если TFO связан с реагентом. Эта реакция приводит к образованию продуктов присоединения ДНК, которые могут быть мутагенными (Columbier et al., Nucleic Acids Research, 24: 4519-4524, 1996). Блеомицин является агентом, разрушающим ДНК, и широко используется в качестве миметика облучения. Этот агент был связан с олигонуклеотидами, и было показано, что он обладает

разрушающей активностью в этой форме (Sergeyev, *Nucleic Acids Research* 23, 4400-4406, 1995; Kane et al., *Biochemistry*, 34, 16715-16724, 1995). Аналоги циклопропапиirroлиндола были связаны с TFO, и было показано, что они алкилируют ДНК в последовательности триплекс-мишени. Затем алкилированная ДНК будет содержать химические аддукты, которые могут быть мутагенными (Lukhtanov et al., *Nucleic Acids Research*, 25, 5077-5084, 1997). Фенантодигидродоксины представляют собой маскированные хиноны, которые высвобождают радикалы при фотоактивации. Они были связаны с TFO и, как было показано, вносят разрывы в дуплексную ДНК при фотоактивации (Bendinskas et al., *Bioconjugate Chem.* 9, 555-563, 1998).

[0284] В настоящем изобретении рассматриваются и другие способы индуцирования модификаций и/или рекомбинации. Так, например, другой вариант осуществления изобретения включает индуцирование гомологичной рекомбинации между экзогенным фрагментом ДНК и целевым геном (см., например, Caracci et al., *Science* 244: 1288-1292, 1989) или ее индуцирование под действием пептид-содержащих нуклеиновых кислот (PNA), обладающих аффинностью к целевому сайту. Другие способы включают последовательность-специфическое распознавание ДНК и таргетинг под действием полиамидов (см., например, Dervan et al., *Curr Opin Chem Biol* 3: 688-693, 1999; *Biochemistry* 38: 2143-2151, 1999) и использование нуклеаз с сайт-специфической активностью (например, белков «цинковые пальцы», TALEN, мегануклеазы и/или CRISPR).

[0285] Настоящее изобретение не ограничивается какой-либо конкретной частотой модификации и/или рекомбинации. В некоторых вариантах осуществления изобретения, описанные здесь способы дают частоту модификации в нуклеотидной последовательности-мишени от 0,01% до 3%. Тем не менее, любая частота (то есть от 0% до 100%) модификации и/или рекомбинации рассматривается как входящая в объем настоящего раскрытия. Частота модификации и/или рекомбинации зависит от метода, применяемого для индуцирования модификации и/или рекомбинации; используемого типа клеток; конкретного гена-мишени; и используемого реагента для мутации ДНК, если таковой имеется. Кроме того, способ, применяемый для детектирования модификации и/или рекомбинации, из-за ограничений способа детектирования, может не обнаруживать все случаи модификации и/или рекомбинации. Кроме того, некоторые события модификации и/или рекомбинации могут быть молчашими, что не позволяет точно детектировать события модификации и/или рекомбинации. Невозможность детектировать молчащие модификации и/или рекомбинации дает ложную низкую оценку модификации и/или рекомбинации. По этим и другим причинам, настоящее изобретение необязательно ограничивается какой-либо конкретной частотой модификации и/или рекомбинации. В одном варианте осуществления изобретения, частота модификации и/или рекомбинации составляет от 0,01 до 100%. В другом варианте осуществления изобретения, частота модификации и/или рекомбинации составляет от 0,01 до 50%. В еще одном варианте осуществления изобретения, частота модификации и/или рекомбинации составляет от

0,1% до 10%. В еще одном варианте осуществления изобретения, частота модификации и/или рекомбинации составляет от 0,1% до 5%.

[0286] Термин «частота мутаций», используемый здесь в отношении популяции клеток, которые были обработаны ДНК-модифицирующей молекулой, способной вводить мутацию в сайт-мишень в геноме клеток, относится к числу клеток в обработанной популяции, которые содержат мутацию в сайте-мишени по сравнению с общим числом клеток, обработанных ДНК-модифицирующей молекулой. Так, например, что касается популяции клеток, которую обрабатывают ДНК-модифицирующей молекулой ТФО, связанной с псораленом, которая предназначена для введения мутации в сайт-мишень в геноме клетки, то частота мутации 5% означает, что из всех 100 клеток, которые были обработаны ТФО-псораленом, 5 клеток содержат мутацию в сайте-мишени.

[0287] Хотя настоящее раскрытие необязательно ограничивается какой-либо степенью точности модификации и/или рекомбинации ДНК в клетке, однако, предполагается, что в некоторых вариантах осуществления изобретения, требуется более высокая степень точности, в зависимости от желаемого результата. Так, например, конкретные изменения последовательности, необходимые для репарации генов (например, конкретные изменения оснований), требуют более высокой степени точности по сравнению с введением нокаута гена, когда необходима только дизрупция гена. При применении способов согласно изобретению, достижение более высоких уровней точности в способах модификации и/или гомологичной рекомбинации является более вероятным, чем в способах известного уровня техники.

Доставка олигонуклеоснований для репарации генов в клетки растений

[0288] Любой общеизвестный метод, применяемый для трансформации клетки растения, может быть применен для доставки олигонуклеоснований для репарации генов. Иллюстративные методы перечислены ниже. Описанные здесь способы и композиции могут включать любой из многих способов трансфекции клеток ДНК-модифицирующим реагентом или реагентами. Способы введения ДНК-модифицирующих реагентов в клетку или в клетки хорошо известны специалистам в данной области и включают, но не ограничиваются ими, микроинъекцию, электропорацию, пассивную адсорбцию, копреципитацию фосфатом кальция-ДНК, трансфекцию, опосредуемую DEAE-декстраном, трансфекцию, опосредуемую полибренном, слияние липосом, использование липофектина, нуклеофекцию, слияние протопластов, ретровирусную инфекцию, биобаллистические методы (то есть, бомбардировку частицами) и т.п.

[0289] Использование металлических микроносителей (микросфер) для введения крупных фрагментов ДНК в клетки растений, имеющие целлюлозные клеточные стенки, путем бомбардировки частицами (далее называемой биобаллистической доставкой), хорошо известно специалистам в данной области. В патентах США № 4945050; 5100792 и 5204253 описаны общие методы выбора микроносителей и устройств для их доставки.

[0290] Конкретные условия использования микроносителей в способах согласно изобретению могут включать условия, описанные в публикации Международной заявки

WO 99/07865. В репрезентативном методе, охлажденные льдом микроносители (60 мг/мл), смешанный дуплексный олигонуклеотид (60 мг/мл), 2,5М CaCl₂ и 0,1 М спермидина добавляют в указанном порядке; а затем смесь осторожно перемешивают, например, путем встряхивания в течение 10 минут, и оставляют на 10 минут при комнатной температуре, после чего микроносители разводят в 5 объемах этанола, центрифугируют и ресуспендируют в 100% этаноле. Хорошие результаты могут быть получены в адгезивном растворе при концентрациях 8-10 мкг/мкл микроносителей, 14-17 мкг/мл смешанного дуплексного олигонуклеотида, 1,1-1,4М CaCl₂ и 18-22 мМ спермидина. Оптимальные результаты наблюдались при концентрациях 8 мкг/мкл микроносителей, 16,5 мкг/мл смешанного дуплексного олигонуклеотида, 1,3 М CaCl₂ и 21 мМ спермидина.

[0291] Олигонуклеоснования для репарации генов могут быть также введены в клетки растений с использованием микроволокон для проникновения через клеточную стенку и клеточную мембрану. В патенте США № 5302523, выданном Coffee R. и Dunwell J.M. (1994), описано использование волокон карбида кремния для облегчения трансформации суспензионных культур кукурузы сорта Black Mexican Sweet. Любая механическая методика, которая может быть применена для введения ДНК для трансформации клетки растения с использованием микроволокон, может быть применена для доставки олигонуклеоснований для репарации генов в целях трансмутации.

[0292] Иллюстративный способ доставки олигонуклеоснования для репарации генов с использованием микроволокон заключается в следующем: стерильные микроволокна (2 мкг) суспендируют в 150 мкл среды для культивирования растений, содержащей приблизительно 10 мкг смешанного дуплексного олигонуклеотида. Суспензионную культуру оставляют для осаждения, и равные объемы упакованных клеток и стерильной суспензии волокна/нуклеотида встряхивают в течение 10 минут и высевают. Селективные среды наносят немедленно или с задержкой приблизительно до 120 часов, в зависимости от конкретного признака.

[0293] В альтернативном варианте осуществления изобретения, олигонуклеоснования для репарации генов могут быть доставлены в клетку растения путем электропорации протопласта, происходящего от части растения. Протопласты получают путем ферментативной обработки части растения, в частности листьев, в соответствии с методиками, хорошо известными специалистам в данной области. См., например, Gallois et al., 1996, в *Methods in Molecular Biology* 55: 89-107, Humana Press, Totowa, N.J.; Kipp et al., 1999, *Methods in Molecular Biology* 133: 213-221, Humana Press, Totowa, N.J. Протопласты необязательно должны быть культивированы в культуральных средах перед электропорацией. Иллюстративными условиями для электропорации является наличие 300000 протопластов в общем объеме 0,3 мл с концентрацией олигонуклеоснований для репарации генов, составляющей 0,6-4 мкг/мл.

[0294] В альтернативном варианте осуществления изобретения, нуклеиновые кислоты поглощаются протопластами растений в присутствии агента, модифицирующего мембрану, а именно, полиэтиленгликоля в соответствии с методами, хорошо известными

специалистам в данной области. В другом альтернативном варианте осуществления изобретения, олигонуклеос основания для репарации генов могут быть доставлены путем инъекции микрокапиллярным методом в клетки растения или в протопласты.

[0295] В альтернативном варианте осуществления изобретения, нуклеиновые кислоты погружают в микросферы, состоящие из альгината кальция, и эти нуклеиновые кислоты поглощаются растительными протопластами в присутствии агента, модифицирующего мембрану, а именно, полиэтиленгликоля (см., например, Sone et al., *Journal of Bioscience and Bioengineering*, 94 (1): 87-91, 2002; Liu et al., 2004).

[0296] В альтернативном варианте осуществления изобретения, нуклеиновые кислоты замораживают в воде и вводят в клетки растения путем бомбардировки микрочастицами (см., например, Gilmore, 1991, патент США 5219746; Brinegar et al.).

[0297] В альтернативном варианте осуществления изобретения, нуклеиновые кислоты, присоединенные к наночастицам, вводят в интактные клетки растений путем инкубирования клеток в суспензии, содержащей наночастицы (см., например, Pasupathy et al., *Biotechnology Journal: Healthcare Nutrition Technology*, 3 (8), 1078-1082, 2008) или путем их доставки в интактные клетки посредством бомбардировки частицами или в протопласты посредством совместного инкубирования (см., например, Torney et al., *Nature nanotechnology*, 2 (5), 295, 2007).

[0298] В альтернативном варианте осуществления изобретения, нуклеиновые кислоты подвергают реакции образования комплекса с проникающими пептидами и доставляют в клетки путем совместного инкубирования (см., например, Chugh and Eudes, *Journal of peptide science: an official publication of the European Peptide Society*, 14(4), 477-481, 2008; WO 2008148223 A1).

[0299] В альтернативном варианте осуществления изобретения, нуклеиновые кислоты вводят в интактные клетки посредством электропорации (см., например, He et al., 1998, US 2003/0115641 A1, Dobres et al.).

[0300] В альтернативном варианте осуществления изобретения, нуклеиновые кислоты доставляют в клетки сухих эмбрионов путем их смачивания в растворе с нуклеиновыми кислотами (см., например, Töpfer et al., 1989, Senaratna et al., 1991), или в других вариантах осуществления изобретения, эти нуклеиновые кислоты вводят с помощью Cellsqueeze (SQZ Biotech).

Методы снижения полипептидной активности и другие методы мутагенеза

[0301] В некоторых своих аспектах, настоящее изобретение относится к снижению уровней и/или активности полипептида (например, полипептида SHP). Способы модификации, уменьшающей количество/уровень или активность одного или более полипептидов согласно изобретению хорошо известны специалистам и описаны в настоящей заявке.

[0302] Клетки (например, клетки растений) согласно изобретению могут содержать один или более полипептидов с пониженной активностью по сравнению с соответствующей контрольной клеткой, такой как клетка дикого типа. В некоторых

вариантах осуществления изобретения, один или более белков SHP обладают пониженной активностью в клетке-хозяине по сравнению с соответствующей контрольной клеткой. Способы снижения уровня экспрессии, количества и/или активности полипептида хорошо известны специалистам и описаны в настоящей заявке.

[0303] В некоторых вариантах осуществления изобретения, снижение активности полипептида, такого как, например, один или более белков SHP, включает снижение уровня экспрессии нуклеиновой кислоты, кодирующей полипептид.

[0304] Снижение уровня экспрессии нуклеиновой кислоты может быть достигнуто путем введения генетической мутации в нуклеиновую кислоту-мишень. Методы мутагенеза могут быть применены для дизрупции или «нокаута» экспрессии гена-мишени путем введения мутаций. В некоторых вариантах осуществления изобретения, мутагенез приводит к частичной делеции гена-мишени. В других вариантах осуществления изобретения, мутагенез приводит к полной делеции гена-мишени. Способы мутагенеза микроорганизмов хорошо известны специалистам в данной области и включают, например, случайный мутагенез и сайт-направленный мутагенез для индуцирования мутаций. Примерами способов случайного мутагенеза являются, например, химический мутагенез (например, с использованием этанметилсульфоната), инсерционный мутагенез и облучение. В некоторых вариантах осуществления изобретения, методы мутагенеза могут быть применены для введения преждевременного стоп-кодона в нуклеиновую кислоту согласно изобретению (например, ген SHP). Это может быть достигнуто, например, путем целевого изменения одного нуклеотида в нуклеиновой кислоте-мишени в сайте, который создает преждевременный стоп-кодон.

[0305] В некоторых вариантах осуществления изобретения, нуклеиновые кислоты согласно изобретению (например, гены SHP) могут быть отредактированы способом, который не приводит к сдвигу открытой рамки считывания, или который, по существу, не устраняет экспрессию нуклеиновой кислоты и/или полипептида, который она кодирует, например, путем введения одной нуклеотидной замены в нуклеиновую кислоту. Для осуществления такого редактирования могут быть применены различные методы, такие как, например, целевое введение точковой мутации в нуклеиновую кислоту. Такое редактирование может, например, снижать уровень экспрессии нуклеиновой кислоты и/или уровень экспрессии и/или активности полипептида, который она кодирует. Так, например, точковая мутация может быть введена в ген SHP, что будет приводить к аминокислотной модификации в кодируемой полипептидной последовательности. В некоторых вариантах осуществления изобретения, такая аминокислотная модификация может быть введена в область, играющую важную роль в функционировании гена SHP, так, чтобы кодируемый мутантный полипептид SHP обладал пониженной активностью и/или измененной функцией.

[0306] Одним из способов снижения или ингибирования экспрессии гена-мишени является генетическая модификация гена-мишени и введение его в геном клетки-хозяина для замены варианта гена дикого типа путем гомологичной рекомбинации (например, как

описано в патенте США № 6924146).

[0307] Другим способом снижения или ингибирования экспрессии гена-мишени является инсерционный мутагенез с использованием Т-ДНК *Agrobacterium tumefaciens* или транспозонов (см. Winkler et al., *Methods Mol. Biol.* 82: 129-136, 1989, и Martienssen *Proc. Natl. Acad. Sci.* 95: 2021-2026, 1998). После получения инсерционных мутантов, эти мутанты могут быть скринированы для идентификации мутантов, которые содержат инсерцию в гене-мишени. Способы дизрупции гена-мишени путем инсерционного мутагенеза описаны, например, в патенте США No. 5792633. Способы дизрупции гена-мишени с помощью транспозонного мутагенеза описаны, например, в патенте США No.6207384.

[0308] Еще одним способом дизрупции гена-мишени является использование системы cre-lox (например, как описано в патенте США № 4959317).

[0309] Другим методом дизрупции гена-мишени является ПЦР-мутагенез (например, как описано в патенте США № 7501275).

[0310] Экспрессия эндогенного гена может быть также снижена или ингибирована с помощью РНК-интерференции (РНКи), в которой используется двухцепочечная РНК, имеющая последовательность, идентичную или аналогичную последовательности гена-мишени. РНКи может включать использование микроРНК, такой как искусственная миРНК, для подавления экспрессии гена.

[0311] РНКи представляет собой феномен, при котором, в случае, когда в клетку вводят двухцепочечную РНК, имеющую последовательность, идентичную или аналогичную последовательности гена-мишени, экспрессия встроенного экзогенного гена и эндогенного гена-мишени подавляется. Двухцепочечная РНК может быть образована из двух отдельных комплементарных РНК или может представлять собой одну РНК с внутренними комплементарными последовательностями, которые образуют двухцепочечную РНК.

[0312] Таким образом, в некоторых вариантах осуществления изобретения, снижение или ингибирование экспрессии генов достигается с применением методов РНКи. Так, например, для снижения или ингибирования экспрессии ДНК, кодирующей белок, с использованием РНКи, в представляющую интерес клетку-хозяина вводят двухцепочечную РНК, имеющую последовательность ДНК, кодирующую белок, или, по существу, аналогичную ей последовательность (включая последовательность, сконструированную так, чтобы она не транслировалась в белок), или ее фрагмент. Используемые здесь РНКи и дцРНК означают геноспецифический сайленсинг, который индуцируется введением двухцепочечной молекулы РНК, см., например, патенты США No. 6506559 и 6573099, и включают молекулу, которая имеет двухцепочечную область, например, молекулу короткой шпилечной РНК. Затем полученные клетки могут быть подвергнуты скринингу на фенотип, ассоциированный с пониженной экспрессией гена-мишени, например, пониженной экспрессией целлюлазы, и/или мониторингу стационарных уровней РНК на наличие транскриптов гена-мишени. Хотя

последовательности, используемые для РНКи, необязательно должны быть полностью идентичны гену-мишень, однако, они могут быть по меньшей мере на 70%, 80%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или более идентичны последовательности гена-мишени. См., например, публикацию патентной заявки США № 2004/0029283. Конструкции, кодирующие молекулу РНК со структурой стебель-петля, которая не является родственной гену-мишени, и которая расположена на значительном расстоянии от последовательности, специфичной к представляющему интерес гену, могут быть также использованы для ингибирования экспрессии гена-мишени. См., например, публикацию заявки на патент США № 2003/0221211.

[0313] Нуклеиновые кислоты РНКи могут включать полноразмерную РНК-мишень или могут соответствовать фрагменту РНК-мишени. В некоторых случаях, фрагмент будет иметь менее, чем 100, 200, 300, 400 или 500 нуклеотидов, соответствующих последовательности-мишени. Кроме того, в некоторых аспектах изобретения, эти фрагменты имеют длину по меньшей мере, например, 50, 100, 150, 200 или более нуклеотидов. Интерферирующие РНК могут быть сконструированы на основе коротких дуплексов (то есть, коротких областей двухцепочечных последовательностей). Обычно, короткий дуплекс имеет длину по меньшей мере приблизительно 15, 20 или 25-50 нуклеотидов (например, каждая комплементарная последовательность двухцепочечной РНК имеет длину 15-50 нуклеотидов), а чаще всего приблизительно 20-30 нуклеотидов, например, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 или 30 нуклеотидов. В некоторых случаях, фрагменты для использования в РНКи будут соответствовать областям белка-мишени, которые не встречаются в других белках в организме, или которые имеют небольшое сходство с другими транскриптами в организме, например, отобранными путем сравнения с последовательностями в анализе с использованием общедоступных баз данных последовательностей. Аналогичным образом, фрагменты РНКи могут быть выбраны по сходству или идентичности с консервативной последовательностью представляющего интерес семейства генов, такого как семейство, описанное в настоящей заявке, так, чтобы РНКи нацеливалась на множество различных транскриптов гена, содержащих консервативную последовательность.

[0314] РНКи могут быть введены в клетку-хозяина как часть более крупной конструкции ДНК. В большинстве случаев, такие конструкции обеспечивают стабильную экспрессию РНКи в клетках после введения, например, путем интеграции конструкции в геном хозяина. Таким образом, экспрессионные векторы, которые непрерывно экспрессируют РНКи в клетках, трансфицированных этими векторами, могут быть использованы в настоящем изобретении. Так, например, могут быть использованы векторы, которые экспрессируют небольшие шпилечные РНК или РНК со структурой «стебель-петля» или предшественники микроРНК, которые процессируются *in vivo* с образованием небольших молекул РНКи, способных осуществлять геноспецифический сайленсинг (Brummelkamp et al., *Science* 296: 550-553 2002) и Paddison et al., *Genes & Dev.* 16: 948-958 (2002)). Посттранскрипционный сайленсинг генов двухцепочечной РНК более

подробно обсуждается в публикациях Hammond et al., *Nature Rev Gen* 2: 110-119, (2001); Fire et al., *Nature* 391: 806-811, (1998); и Timmons and Fire, *Nature* 395: 854 (1998).

[0315] Способы отбора и конструирования последовательностей, которые генерируют РНКи, хорошо известны специалистам в данной области (см., например, патенты США №№ 6506559; 6511824; и 6489127).

[0316] Снижение или ингибирование экспрессии гена-мишени в в клетке-хозяине может быть также достигнуто путем введения в клетки-хозяева антисмысловых конструкций на основе последовательности нуклеиновой кислоты гена-мишени. Для ингибирования антисмысловой последовательности, последовательность-мишень располагают в обратной ориентации относительно промоторной последовательности в экспрессионном векторе. Введенная последовательность необязательно должна представлять собой полноразмерную кДНК или полноразмерный ген и необязательно должна быть идентична кДНК-мишени или гену, присутствующему в трансформируемой клетке. Однако, обычно, если введенная последовательность имеет меньшую длину, то для эффективного подавления антисмысловой последовательности используется более высокая степень гомологии по отношению к нативной последовательности-мишени. В некоторых аспектах изобретения, введенная антисмысловая последовательность в векторе будет иметь длину по меньшей мере 30 нуклеотидов, а повышенная антисмысловая супрессия обычно будет наблюдаться по мере увеличения длины антисмысловой последовательности. В некоторых аспектах изобретения, длина антисмысловой последовательности в векторе будет составлять более, чем 100 нуклеотидов. Транскрипция описанной антисмысловой конструкции приводит к образованию молекул РНК, которые представляют собой обратный комплемент молекул мРНК, транскрибируемых из эндогенного гена-мишени. Подавление экспрессии гена-мишени также может быть достигнуто с использованием рибозима. Получение и использование рибозимов раскрывается в патентах США No. №№ 4987071 и 5543508.

[0317] Экспрессионные кластеры, содержащие нуклеиновые кислоты, которые кодируют ингибиторы экспрессии гена-мишени, например, антисмысловую РНК или киРНК, могут быть сконструированы способами, хорошо известными специалистам. Конструкции содержат регуляторные элементы, включая промоторы и другие последовательности для экспрессии и отбора клеток, которые экспрессируют эту конструкцию. Обычно, векторы для трансформации грибов и/или бактерий включают одну или более клонированных кодирующих последовательностей (геномных или кДНК) под транскрипционным контролем 5'- и 3'-регуляторных последовательностей и доминантный селективный маркер. Такие векторы для трансформации обычно также содержат промотор (например, регуляторную область, регулируемую индуцибельную или конститутивную экспрессию, а также экспрессию, регулируемую окружающей средой или факторами развития), сайт инициации транскрипции, сигнал процессинга РНК (такой как сайты сплайсинга интронов), сайт терминации транскрипции и/или сигнал полиаденилирования.

[0318] В некоторых вариантах осуществления изобретения, часть нуклеиновой кислоты-мишени может быть модифицирована, например, с получением области, кодирующей каталитический домен, кодирующей области или регуляторной последовательности, необходимой для экспрессии кодирующей области. Такая регуляторная последовательность гена может представлять собой промоторную последовательность или ее функциональную часть, то есть, часть, достаточную для влияния на экспрессию гена. Так, например, промоторная последовательность может быть инактивирована, так, чтобы она не приводила к экспрессии, либо нативная промоторная последовательность может быть заменена более слабым промотором для снижения уровня экспрессии кодирующей последовательности. Другие регуляторные последовательности для возможной модификации могут включать, например, лидерную последовательность, пропептидную последовательность, сигнальную последовательность, терминатор транскрипции и активатор транскрипции.

Растения согласно изобретению

[0319] Описанные здесь способы и композиции могут, в определенных аспектах и вариантах осуществления изобретения, быть применимы к растениям в целом. Так, например, в некоторых аспектах и/или вариантах осуществления изобретения, растения могут быть выбраны из семейства Brassicaceae, включая ряд ценных сельскохозяйственных культур, таких как *Brassica napus* (канола, масличный рапс), *Brassica rapa* (например, турнепс, капуста китайская), *Brassica oleracea* (брокколи, капуста белокачанная, цветная капуста и т.п.), *Brassica juncea* (горчица) или *Raphanus sativus* (редька обыкновенная), а также многие ценные бобовые культуры, такие как горох, фасоль, чечевица и соя.

[0320] В соответствии с описанием настоящего изобретения, по существу, нормальный рост растения, органа растения, ткани растения или клетки растения определяется как скорость роста или скорость деления клеток растения, органа растения, ткани растения или клетки растения, где указанная скорость составляет по меньшей мере 35%, по меньшей мере 50%, по меньшей мере 60% или по меньшей мере 75% от скорости роста или скорости деления клеток в соответствующем растении, органе растения, ткани растения или в клетке растения, экспрессирующих белок SHP дикого типа.

[0321] В соответствии с настоящим описанием, по существу, нормальное развитие растения, органа растения, ткани растения или клетки растения определяется как одно или более событий развития растения, органа растения, ткани растения или клетки растения, которые, по существу, являются такими же, как в соответствующем растении, органе растения, ткани растения или клетке растения, экспрессирующих белок SHP дикого типа.

[0322] В соответствии с описанием настоящего изобретения, органами растения являются, но не ограничиваются ими, листья, стебли, корни, вегетативные почки, цветочные почки, меристемы, эмбрионы, семядоли, эндосперм, чашелистики, лепестки, пестики, плодолистики, тычинки, пыльники, микроспоры, пыльца, пыльцевые трубки, завязи, семяпочки и плоды или взятые из них сегменты, срезы или диски. Тканями

растений являются, но не ограничиваются ими, ткани каллуса, наземные ткани, сосудистые ткани, запасные ткани, меристемные ткани, ткани листьев, ткани побегов, ткани корней, ткани галлов, ткани опухолей растений и ткани репродуктивных органов. Клетки растения включают, но не ограничиваются ими, выделенные клетки с клеточными стенками, их агрегаты различных размеров и протопласты.

[0323] Растения согласно изобретению включают растения, у которых может наблюдаться осыпание стручков. Так, например, настоящее изобретение включает растения *Brassica spp.*, у которых наблюдается осыпание стручков.

[0324] В различных вариантах осуществления изобретения, описание растений в основном, сосредоточено на однодольных растениях, включая любые древесные породы растений, которые представляют собой деревья или кустарники, любые травянистые растения или любые растения, которые дают съедобные плоды, семена или овощи, или любые виды, которые имеют красивые или ароматные цветы. Так, например, растение может быть выбрано из видов растений, выбранных из группы, состоящей из канолы, подсолнечника, кукурузы, табака, сахарной свеклы, хлопчатника, маиса, пшеницы, ячменя, риса, люцерны, сорго, томатов, манго, персика, яблони, груши, клубники, бананов, дынь, маниоки, картофеля, моркови, салата-латука, лука, сои, сои различных видов, сахарного тростника, гороха, нута, полевого гороха, конских бобов, чечевицы, репы, брюквы, брюссельской капусты, люпина, цветной капусты, капусты кормовой, кормовой фасоли, тополя, сосны, эвкалипта, винограда, цитрусовых, тритикале, люцерны, ржи, овса, травы для газонов и фуражных трав, льна, масличного рапса, горчицы, огурца, ипомеи, бальзамина, перца, баклажана, календулы, лотоса, капусты, маргаритки, гвоздики, тюльпана, ириса, лилии и орехоносных растений, причем, это список является далеко неисчерпывающим.

[0325] Растения и клетки растения могут быть протестированы на их резистентность к растрескиванию плодов до сбора урожая с применением общеизвестных методов.

[0326] В некоторых вариантах осуществления изобретения, растения согласно изобретению имеют одну или более мутаций в одном или более генах SHP, и обладают повышенной резистентностью/пониженной чувствительностью к растрескиванию плодов до сбора урожая по сравнению с соответствующим контрольным растением (например, растением того же вида, которое не имеет каких-либо мутаций в любых генах SHP, таким как растение дикого типа). Частота осыпания стручков у растений, имеющих повышенную резистентность/пониженную чувствительность к растрескиванию плодов до сбора урожая, может быть, например, по меньшей мере приблизительно на 5%, по меньшей мере приблизительно на 10%, по меньшей мере приблизительно на 15%, по меньшей мере приблизительно на 20%, по меньшей мере приблизительно на 25%, по меньшей мере приблизительно на 30%, по меньшей мере приблизительно на 40%, по меньшей мере приблизительно на 50%, по меньшей мере приблизительно на 55%, по меньшей мере приблизительно на 60%, по меньшей мере приблизительно на 65%, по

меньшей мере приблизительно на 70%, по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 91%, по меньшей мере приблизительно на 92%, по меньшей мере приблизительно на 93%, по меньшей мере приблизительно на 94%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 96%, по меньшей мере приблизительно на 97%, по меньшей мере приблизительно на 98%, по меньшей мере приблизительно на 99% или, по меньшей мере приблизительно на 100% ниже частоты осыпания стручков у соответствующего контроля, или такая частота является пониженной.

[0327] Используемый здесь термин «по существу, нормальный рост растения, органа растения, ткани растения или клетки растения» определяют как скорость роста или скорость деления клеток растения, органа растения, ткани растения или клетки растения, которая составляет по меньшей мере 35%, по меньшей мере 50%, по меньшей мере 60% или по меньшей мере 75% от скорости роста или скорости деления клеток в соответствующем растении, органе растения, ткани растения или клетке растения, экспрессирующих представляющий интерес белок дикого типа.

[0328] Используемый здесь термин «по существу, нормальное развитие растения, органа растения, ткани растения или клетки растения» определяют как одно или более событий развития растения, органа растения, ткани растения или клетки растения, которые, по существу, являются такими же, как в соответствующем растении, органе растения, ткани растения или клетке растения, экспрессирующих белок дикого типа.

[0329] В некоторых вариантах осуществления изобретения, описанными здесь органами растения являются, но не ограничиваются ими, листья, стебли, корни, вегетативные почки, цветочные почки, меристемы, эмбрионы, семядоли, эндосперм, чашелистики, лепестки, пестики, плодолистики, тычинки, пыльники, микроспоры, пыльца, пыльцевые трубки, завязи, семяпочки и плоды или взятые из них сегменты, срезы или диски. Тканями растений являются, но не ограничиваются ими, ткани каллуса, наземные ткани, сосудистые ткани, запасные ткани, меристемные ткани, ткани листьев, ткани побегов, ткани корней, ткани галлов, ткани опухолей растений и ткани репродуктивных органов. Клетки растения включают, но не ограничиваются ими, выделенные клетки с клеточными стенками, их агрегаты различных размеров и протопласты.

Поколения растений

[0330] Культуры тканей растений различных видов и регенерация растений из этих тканей являются известными. Так, например, размножение сорта канолы из культуры ткани описано в любой из нижеследующих неограничивающих публикаций: Li et al., “Somatic embryogenesis in quite a direct way in cultures of mesophyll protoplasts of *Brassica napus* L.”, *Plant Cell Reports* 1: 209-211, 1982; Chuong et al., “A Simple Culture Method for *Brassica hypocotyls* Protoplasts,” *Plant Cell Reports* 4:4-6, 1985; Barsby et al., “A Rapid and Efficient Alternative Procedure for the Regeneration of Plants from *Hypocotyl* Protoplasts of

Brassica napus,” *Plant Cell Reports* (Spring, 1996); Kartha et al., “In vitro Plant Formation from Stem Explants of Rape,” *Physiol. Plant*, 31:217-220, 1974; Narasimhulu et al., “Species Specific Shoot Regeneration Response of Cotyledonary Explants of Brassicas,” *Plant Cell Reports* (Spring 1988); Sun et al., “Cotyledon-derived diploid and haploid protoplast culture and diploid plant regeneration in *Brassica napus* cv. ‘Topas’,” *Can. J. Bot.* 76: 530-541, 1998; Swanson, E., “Microspore Culture in Brassica,” *Methods in Molecular Biology*, Vol. 6, Chapter 17, p. 159, 1990.

[0331] Дальнейшее размножение этого сорта может быть осуществлено путем культивирования ткани и регенерации. Культура различных тканей сои и регенерация растений из них хорошо известны и широко описаны в литературе. См., например, Komatsuda et al., “Genotype X Sucrose Interactions for Somatic Embryogenesis in Soybeans,” *Crop Sci.* 31:333-337, 1991; Stephens et al., “Agronomic Evaluation of Tissue-Culture-Derived Soybean Plants,” *Theor. Appl. Genet.* 82:633-635, 1991; Komatsuda et al., “Maturation and Germination of Somatic Embryos as Affected by Sucrose and Plant Growth Regulators in Soybeans *Glycine gracilis* L. Skvortz and *Glycine max* L. Merr.” *Plant Cell, Tissue and Organ Culture*, 28:103-113, 1992; Dhir et al., “Regeneration of Fertile Plants from Protoplasts of Soybean (*Glycine max* L. Merr.); Genotypic Differences in Culture Response,” *Plant Cell Reports* 11:285-289, 1992; Pandey et al., “Plant Regeneration from Leaf and Hypocotyl Explants of *Glycine wightii* (W. and A.) VERDC. var. *longicauda*,” *Japan J. Breed.* 42:1-5, 1992; and Shetty et al., “Stimulation of In Vitro Shoot Organogenesis in *Glycine max* L. Merrill. by Allantoin and Amides,” *Plant Science* 81:245-251, 1992. Описание патента США No. 5024944, выданного Collins et al., 18 июня 1991, и патента США No. 5008200, выданного Ranch et al., 16 апреля 1991, включено в настоящее описание в полном объеме посредством ссылки.

[0332] В некоторых своих аспектах, настоящее изобретение также относится к растениям, полученным из растений, имеющих одну или более мутаций в нуклеиновой кислоте (например, в гене SHP) согласно изобретению. Так, например, растения, имеющие одну или более мутаций SHP, могут быть скрещены с одними и теми же или различными растениями для получения потомства растения F₁, где по меньшей мере один из родителей потомства растения F₁ имел одну или более мутаций SHP. Эти растения F₁ могут быть затем скрещены между собой или скрещены с другой линией растений, и полученное потомство F₂ может быть подвергнуто скринингу на наличие одной или более мутаций SHP.

Примеры

[0333] Нижеследующие примеры представлены для дополнительной иллюстрации аспектов настоящего изобретения. Эти примеры не являются ограничивающими и не должны рассматриваться как ограничение какого-либо аспекта настоящего изобретения.

Пример 1: Идентификация и характеристика генов SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A и SHP4C

Молекулярная характеристика генов BnSHP

[0334] Было охарактеризовано восемь генов BnSHP, имеющих в базе данных геномов Brassica (Genoscope). Гены канолы SHP обозначены здесь как BnSHP1A, BnSHP1C, BnSHP2A, BnSHP2C, BnSHP3A, BnSHP3C, BnSHP4A и BnSHP4C. Очевидно, что каждая половина генов была представлена родительскими субгеномами *B. rapa* (AA) и *B. oleracea* (CC). Соответствующее хромосомное расположение каждого гена в геноме *B. napus* представлено в Таблице 1; нуклеотидная последовательность кодирующих областей всех восьми генов SHP имеется в Genoscope (SEQ ID NO: 1-8, см. Таблицу 1), и последовательность геномной ДНК всех восьми генов SHP также имеется в Genoscope (SEQ ID NO: 9-16, см. Таблицу 1). Неполная последовательность 5'-промоторной области UTR, а также часть 5'-кодирующей последовательности каждого гена BnSHP были клонированы и секвенированы из геномной ДНК, полученной из линии растений BN2-SU. Было обнаружено, что все 8 генов имеют уникальные промоторы. 5'-UTR гена BnSHP1A выявил вставку размером ~5 т.п.о., которая, по-видимому, представляет собой транслоцируемый мобильный элемент, который не присутствует в эталонной последовательности в Genoscope, и может быть специфичным для данной линии. На фиг. 1A показано выравнивание неполных нуклеотидных последовательностей, а на фиг. 1B показаны соответствующие транслируемые аминокислотные последовательности, полученные после характеристики генов 5'-геномной области генов SHP BN2-SU, и их сравнение с соответствующими последовательностями AtSHP1 и AtSHP2 Arabidopsis. Уровень гомологии среди всех генов BnSHP и генов AtSHP на уровне нуклеотидов и аминокислот является очень высоким (>80%).

Таблица 1: Гены резистентности к осыпанию семян, их положение в хромосоме (источник: Genoscope), источник генома, нуклеотидная кодирующая последовательность и последовательность геномной ДНК

Ген SHP под названием Cibus	Хромосома	Положение в хромосоме	Геном	Кодирующая последовательность SEQ ID NO:	Последовательность геномной ДНК, SEQ ID NO:
BnSHP-1A	A09	3139589-3143645	A	1	9
BnSHP-1C	C08	29697543-29700152	C	2	10
BnSHP-2A	A07	14877625-14881448	A	3	11
BnSHP-2C	C06	19706214-19709351	C	4	12

BnSHP-3A	A04	1164261- 1165954	A	5	13
BnSHP-3C	C04	24456162- 24474692	C	6	14
BnSHP-4A	A05	1619288- 1623119	A	7	15
BnSHP-4C	Неизвестна	Неизвестно	C	8	16

Клонирование и дальнейшая молекулярная характеристика генов BnSHP

[0335] Гены SHP1 и 2 в *Arabidopsis thaliana* в высокой степени гомологичны генам SHP канолы, обладающим 80% нуклеотидной идентичностью. С использованием общедоступных кДНК и геномных последовательностей *Arabidopsis thaliana* SHP1 и SHP2 и последовательностей *Brassica napus* были сконструированы ПЦР-праймеры, которые были использованы для амплификации последовательностей генов BnSHP из геномной ДНК элитных линий канолы BN2 и BN-17. ПЦР-амплифицированные геномные фрагменты SHP клонировали и секвенировали. Дополнительное секвенирование фрагментов геномной ДНК проводили путем секвенирования следующего поколения для завершения этого анализа. Прямой и обратный праймеры, уникальные для каждого из восьми генов SHP, были использованы для ПЦР-амплификации фрагмента из 5'UTR до области генов SHP, которая кодирует домен MADS-бокса. Геномную ДНК, выделенную из гаплоидной линии BN2-SU *B. napus*, использовали для ПЦР-амплификации каждого фрагмента гена. Продукты клонировали в клонирующий вектор TOPO TA, переносили в компетентные бактериальные клетки и высевали на планшеты со средой LB. Для каждого гена клонировали и секвенировали минимум десять колоний, и полученную последовательность сравнивали с эталонными последовательностями, имеющимися в GenBank и Genoscope.

Анализ на экспрессию генов

[0336] Стручки канолы из 6 стадий развития собирали (фиг. 2А). Общую РНК выделяли из каждого образца с использованием геноспецифических ОТ-ПЦР-праймеров для амплификации экспрессируемых генов/аллелей SHP из кДНК, синтезированной из всех экстрактов РНК. Затем кДНК каждого образца секвенировали с помощью NGS для определения относительной экспрессии каждого гена SHP на каждой тестируемой стадии развития.

[0337] Экспрессию всех 8 генов BnSHP исследовали для того, чтобы понять, какие именно гены являются важными для фенотипа снижения степени осыпания стручков. Анализ на экспрессию с использованием геноспецифических праймеров позволил получить по меньшей мере 6 из 8 генов BnSHP, экспрессированных в развивающихся стручках (фиг. 2А и фиг. 2В). Последовательности, соответствующие генам BnSHP1A и

BnSHP4C, в этих данных не представлены.

Ссылки

Roeder, A.H.K. and Yanofsky, M.F. (2006) Fruit Development in Arabidopsis. In, The Arabidopsis Book, American Society of Plant Biologists, doi: e0075. 10.1199/tab.0075

Liljegren, S.J., Ditta, G.S., Eshed, Y., Savidge, B., Bowman, J.L., and Yanofsky, M.F. (2000) SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. Nature 404, 776-770.

Raman H, Raman R, Kilian A, Detering F, Carling J, et al. (2014) Genome-Wide Delineation of Natural Variation for Pod Shatter Resistance in Brassica napus. PLoS ONE 9(7): e101673. doi:10.1371/journal.pone.0101673.

Gururaj K. (2009) Brassica shatter-resistance research update. In, 16th Australian Research Assembly on Brassicas. Ballarat, Victoria, 2009

Chalhoub, B. et al., (2014) Early allopolyploid evolution in the post-Neolithic Brassica napus oilseed genome. Science 345, 950-953.

Пример 2: Получение линий нокаутированных генов резистентности к осыпанию семян в линии канолы BN2-SU-H путем доставки плазмиды CRISPR/Cas9 в протопласты

[0338] В этом примере были получены линии растений канолы, толерантные к сульфонилмочевине и содержащие нефункциональные гены (КО) резистентности к осыпанию семян (SHP) с использованием CRISPR/Cas9. Ген CRISPR/Cas9 и оцрПНК, содержащиеся в плаزمиде, доставляли в протопласты, выделенные из листьев гаплоидных растений для нокдауна генов BnSHP. оцрПНК представляет собой гибрид ПНК CRISPR (cr-ПНК) и транскриптивирующей cr-ПНК (tracr-ПНК). cr-ПНК направляет Cas9 к генам-мишеням, где Cas9 вводит двухцепочечный разрыв в каждый из генов BnSHP по сайт-направленному механизму. Двухцепочечные разрывы в гене BnSHP, если они были репарированы по универсальному пути NHEJ с вероятностью ошибки, будут вызывать InDel (инсерции или делеции нуклеотидов) возле сайта расщепления. Аллели потери функции генов BnSHP присутствуют в том случае, когда эти InDel, созданные Cas9, будут сдвигать рамку считывания генов SHP.

[0339] Процент образования InDel в каждой из 8 генов-мишеней SHP в побегах, регенерированных из протопластов, обработанных CRISPR/Cas9, составлял от 20 до 40%, как было определено путем секвенирования следующего поколения (фиг. 3). Наиболее распространенными идентифицированными InDel являются +1 инсерция (до 10%) и -1 и -2 делеций (5% и 1,5% соответственно; данные не приводятся). Побеги с мутациями в 1-8 генах были идентифицированы с различными частотами (Таблица 2). Приблизительно 70% побегов содержали InDel по меньшей мере в одном гене SHP. Анализ последовательности области-мишени (возле сайта расщепления Cas9) в каждом гене показал, что не все InDel в этих генах приводили к образованию нефункционального (КО) аллеля из-за сдвига рамки считывания. Большинство растений имели от 2 до 5 генов с КО (Таблица 2). В результате этого эксперимента было регенерировано 80 независимых

линий растений, содержащих InDel в 2-8 генах VnSHP, что свидетельствует о том, что Cas9 является активным и способен расщеплять все 8 генов-мишеней SHP в *V. parus* и образовывать InDel. Кроме того, эти InDel могут генерировать нефункциональные гены с КО после введения плазмиды CRISPR/Cas9 посредством сдвига рамки считывания генов SHP.

Таблица 2: Частота встречаемости генов VnSHP с InDel и нефункциональных генов с КО в регенерированных побегах, скринированных с помощью NGS

Гены SHP с InDel	Побеги		Ген SHP с КО*	Побеги	
	#	%		#	%
0	133	29	0	149	33
1	40	9	1	36	8
2	38	8	2	62	14
3	49	11	3	80	18
4	57	12	4	62	14
5	50	11	5	40	9
6	57	12	6	14	3
7	19	3	7	8	2
8	11	2	8	3	1
Всего побегов	454	100	Всего побегов	454	100

* Количество нефункциональных аллелей генов VnSHP после сдвига рамки считывания, вызванного InDel.

Методы

[0340] Протопласты канолы выделяли из листьев выращенных *in vitro* гаплоидных растений BN2-SU из культуры микроспор (Sun et al., 1998; Swanson E., 1990). Плазмиды, кодируемые CRISPR/Cas9, содержат pMas:Cas9 с терминатором rbcSE9 гороха и AtU6P::оцрPHK с терминатором поли-T₁₀. Эти элементы имеют такие последовательности, как: аминокислотная последовательность Cas9 (SEQ ID NO: 50), нуклеотидная последовательность промотора Mas (SEQ ID NO: 51), нуклеотидная последовательность терминатора rbcSE9 (SEQ ID NO: 52), нуклеотидная последовательность промотора AtU6 (SEQ ID NO: 53) и нуклеотидная последовательность терминатора поли-T₁₀ (SEQ ID NO: 54).

[0341] Плазмиды CRISPR/Cas9 вводили в протопласты путем ПЭГ-опосредованной доставки в конечной концентрации 0,05 мкг/мкл. Протопласты культивировали в жидкой среде ($2,5 \times 10^5$ клеток/мл) и инкубировали в темноте при 25°C. Образцы клеток получали через одну неделю и анализировали с помощью NGS. Через 6-8 недель, полученные из протопласта микрокалусы высевали на твердую среду для регенерации, и приблизительно через 2-4 недели из регенерированных калусов начали образовываться

побеги. Образцы листьев от полностью сформировавшихся побегов были проанализированы с помощью NGS для определения наличия InDel в целевых генах SHP. Удлиненные побеги подвергали микроразмножению, и прижившиеся растения переносили в почву и выдерживали в камере для культивирования в течение 2-4 недель до полного укоренения растений.

[0342] CRISPR/Cas9 состоит из двух компонентов: Cas9 *Streptococcus pyogenes*, оптимизированного по кодонам растений (*SpCas9*), и оцрРНК, которые экспрессировались из отдельных плазмид. оцрРНК представляет собой гибрид РНК CRISPR (*cr*-РНК) и транс-активирующей *cr*-РНК (*tracr*-РНК). Область *cr*-РНК содержит спейсерные последовательности, описанные в Таблице 3, которые были использованы для нацеливания нуклеазы Cas9 на ген-мишень. В этом эксперименте, CRISPR/Cas9 нацелен на гены *BnSHP*.

Таблица 3: рРНК, нацеленные на гены *BnSHP*

Название рРНК	Последовательность рРНК (5' - 3')	Ген-мишень <i>BnSHP</i>
<i>BnSHP</i> -2	GTAGCAAGAAGATAGGTAGA (SEQ ID NO: 42)	1A, 1C
<i>BnSHP</i> -3	GTAACAAGAAGCTAGTGAGA (SEQ ID NO: 43)	3A, 3C
<i>BnSHP</i> -4	GTAGCAAGAAGCTAGTAAGA (SEQ ID NO: 44)	2A, 2C, 3A, 3C
<i>BnSHP</i> -5	GCAGCAAGAAGATAGGGAGA (SEQ ID NO: 45)	4A, 4C
<i>BnSHP</i> -9	CAGAAGCAATGGATGAAGGT (SEQ ID NO: 46)	1A, 1C
<i>BnSHP</i> -10	CAGAATCAATGGAGGAAGGT (SEQ ID NO: 47)	2A, 2C, 3A, 3C
<i>BnSHP</i> -11	GGGTTGATATAAATGGAGGG (SEQ ID NO: 48)	4A, 4C
<i>BnSHP</i> -12	CAGAAGCAATGGATGAAAGT (SEQ ID NO: 49)	1A, 1C

Ссылки

Sun et al., "Cotyledon-derived diploid and haploid protoplast culture and diploid plant regeneration in *Brassica napus* cv. «Топас», *Can. J. Bot.* 76: 530-541, 1998.

Swanson, E., "Microspore Culture in Brassica", *Methods in Molecular Biology*, Vol. 6, Chapter 17, p. 159-69, 1990.

Пример 3: Получение линий нокаутированных генов резистентности к осыпанию семян в линии канолы путем доставки рибонуклеопротеина CRISPR/Cas9 (RNP) в протопласты

[0343] Как и в предыдущем примере 2, целью этого примера является получение SU-толерантных линий растений канолы с нефункциональными (КО) генами резистентности к осыпанию семян (SHP) с использованием CRISPR/Cas9. CRISPR/Cas9, используемый для нокдауна генов *BnSHP* в геноме *B. napus*, доставляют в протопласты листьев в виде белка Cas9, образующего комплекс с рРНК (RNP), в комбинации с одноцепочечными олигонуклеотидами (ssODN или GRON, Таблица 4), нацеленными на три специфические мутации InDel (+1 инсерция, -1 и -2 делеции) в целях дизрупции функции каждого из 8 паралогичных генов *BnSHP*. Перед доставкой в протопласты, рекомбинантный белок Cas9 (коммерчески доступный) подвергают реакции образования комплекса *in vitro* с рРНК (Таблица 5), которая синтезируется *in vitro* из матричной ДНК-плазмиды с образованием гибрида РНК CRISPR (cr-РНК) и транс-активирующей cr-РНК (tracr-РНК). cr-РНК, после ее проникновения в клетку, доставляет Cas9 к генам-мишеням, где Cas9 образует двухцепочечный разрыв в каждом из генов *BnSHP* по сайт-направленному механизму. В присутствии GRON, двухцепочечные разрывы в генах *BnSHP* могут быть подвергнуты репарации, помимо пути NHEJ, по пути прямой гомологичной репарации (HDR). Путь HDR может использовать GRON в качестве ДНК-матрицы и вводить целевые мутации в конкретные последовательности GRON (то есть, n+1, n-1 и n-2). Аллели потери функции генов *BnSHP* встречаются в тех случаях, когда эти InDel сдвигают рамку считывания генов SHP, что приводит к образованию усеченных и нефункциональных генных продуктов (мРНК и белков, Таблица 6).

Результаты

[0344] Образование InDel в 1 и до 8 генов SHP наблюдалось в более, чем 95% побегов, регенерированных из протопластов, обработанных CRISPR/Cas9, как было определено путем секвенирования следующего поколения. GRON-направленные мутации (+1, -1, -2 нуклеотидных инсерций или делеций) были обнаружены в более, чем 90% побегов с InDel в генах SHP. Побег с мутациями в каждом из 8 генов SHP и с комбинациями множества генов с КО были идентифицированы с различными частотами. Из 5395 побегов, отобранных в трех последовательных экспериментах, было регенерировано 1127 независимых линий растений, содержащих целевые InDel в 1-8 генах *BnSHP*, включая всего 153 уникальных генотипа с КО (из 255 возможных). Линии растений с КО, представляющие каждый из генотипов SHP с КО, были успешно перенесены в почву и выращены в теплице до созревания для проведения анализов на фенотипические признаки.

Методы

[0345] Протопласты канолы выделяли из листьев микроразмноженных гаплоидных растений *in vitro*. Белок Cas9, образующий комплекс с рРНК (Таблица 5), наряду с одноцепочечными олигонуклеотидами (ssODN; GRON), вводит точные геноспецифические мутации в каждый из 8 паралогичных генов *BnSHP* (Таблица 4). Белок Cas9, образующий комплекс с рРНК и GRON, вводили в протопласты путем ПЭГ-опосредованной доставки в конечной концентрации 0,05 мкг/мкл и 0,5 мкМ,

соответственно. Протопласты культивировали в жидкой среде ($1,25 \times 10^5$ клеток/мл) и инкубировали в темноте при 25°C. Образцы клеток получали через одну неделю и анализировали с помощью NGS. Через 6-8 недель, полученные из протопластов микрокалусы высевали на твердую среду для регенерации, и приблизительно через 2-4 недели из регенерированных каллусов начали образовываться побеги. Образцы листьев от полностью сформировавшихся побегов были проанализированы с помощью NGS для определения наличия InDel в целевых генах SHP. Удлиненные побеги подвергали микроразмножению, и укоренившиеся растения переносили в почву и выдерживали в камере для культивирования в течение 2-4 недель до полного укоренения растений.

[0346] GRON, используемые с RNP Cas9, содержат кодирующую последовательность целевых генов SHP рядом с сайтом конверсии и были помечены группой 2'-O-Me в первом 5'-основании GRON, которое является основанием РНК, вместо основания ДНК (Таблица 4). CRISPR/Cas9 состоит из двух компонентов: Cas9 *Streptococcus pyogenes*, оптимизированного по кодонам растений (SpCas9), и оцрРНК, которые экспрессировались в виде белка и РНК соответственно. оцрРНК *in vitro* транскрибировалась из ДНК-матрицы и представляет собой гибрид РНК CRISPR (cr-РНК) и транс-активирующей cr-РНК (tracr-РНК). Область cr-РНК содержит спейсерные последовательности, описанные в Таблице 5, которые были использованы для нацеливания нуклеазного белка Cas9 на ген-мишень. В этом примере, CRISPR/Cas9 нацелен на гены BnSHP.

Таблица 4: Последовательность ssODN (GRON)

ID	Последовательность (5'-3')
BnSHP(-1;CR2-5)/C/41	A*GAGAGYARCAAGAAGMTAGKAGAGGGAAGATAGAGAT AAA (SEQ ID NO: 55)
BnSHP(-2R;CR2-5)/C/41	A*GAGAGYARCAAGAAGMTAGAGAGGGAAGATAGAGATA AAG (SEQ ID NO: 56)
BnSHP(+1;CR2-5)/C/42	A*GAGAGYARCAAGAAGMTAGKDNAGAGGGAAGATAGAG ATAAA (SEQ ID NO: 57)

Таблица 5: рРНК, нацеленные на гены BnSHP

Название рРНК	Последовательность рРНК (5'-3')	Ген-мишень BnSHP
BnSHP-2	GTAGCAAGAAGATAGGTAGA (SEQ ID NO: 42)	1A, 1C
BnSHP-3	GTAACAAGAAGCTAGTGAGA (SEQ ID NO: 43)	3A, 3C

	43)	
BnSHP-4	GTAGCAAGAAGCTAGTAAGA (SEQ ID NO: 44)	2A, 2C, 3A, 3C
BnSHP-5	GCAGCAAGAAGATAGGGAGA (SEQ ID NO: 45)	4A, 4C

Таблица 6: Размер предсказанных усеченных нефункциональных белковых продуктов, полученных путем введения целевой мутации n-1, n-2 и n+1 (A, G, C и T) в каждый из генов SHP

Ген SHP	Стоп-кодон (AA#)				
	Эталон дикого типа	n-2	n-1	n+1 (A, G, C)	n+1 (T)
1A	276	48	21	49	18
1C	136	48	21	49	18
2A	250	37	21	38	18
2C	249	37	21	38	18
3A	277	48	21	49	18
3C	276	48	21	49	18
4A	245	48	21	49	18
4C	245	44	21	45	18

Пример 4: Фенотипирование линий канолы с мутацией генов резистентности к осыпанию семян

[0347] В растении *Arabidopsis thaliana*, гены SHATTERPROOF 1 и 2 (SHP1/SHP2) представляют собой факторы транскрипции, принадлежащие к членам семейства MADS-боксов, участвующих в формировании зоны растрескивания (DZ), то есть, слоя клеток между створками плодов, ответственными за осыпание зрелых стручков (Liljegen et al., 2000; Roeder and Yanofsky, 2006). Дифференциация DZ в развивающихся стручках характеризуется образованием слоя клеток с одревесневшими клеточными стенками и разделительного клеточного слоя. В двойных мутантах *thaliana shp1 shp2* не развиваются функциональные DZ, которые не подвергаются одревеснению или имеет определенный разделительный слой, и, как следствие, плоды не растрескиваются и не раскрываются в конце развития. В каноле (*Brassica napus*) заявители идентифицировали и охарактеризовали 8 генов, которые являются в высокой степени гомологичными *Arabidopsis SHP1/SHP2*. Гены BnSHP, по-видимому, также участвуют в дифференцировке DZ и играют аналогичную роль в предотвращении осыпания зрелых стручков канолы.

Результаты

[0348] Исследования потери функции показали, что SHP1 и SHP2 способствуют одревеснению субпопуляции клеток краев створки плодов *Arabidopsis*. Паттерн

одревеснения плодов, полученных из различных линий VnSHP KO C₀, анализировали и сравнивали с плодами дикого типа. При этом наблюдалось явное снижение одревеснения клеток краев створки у плодов линий KnSHP KO (фиг. 4). В мутированных линиях с 7-8 генами с KO VnSHP, какого-либо одревеснения краев створки у основании плодов не наблюдалась, тогда как одревесневшие клетки краев створки, окрашенные флороглюцинолом, присутствуют в плодах дикого типа, а также в плодах мутантных линий растений с низким числом генов с KO (фиг. 4).

[0349] Тест на разрушение стручка с использованием TissueLyser и Geno/Grinder был также проведен для анализа на резистентность к осыпанию семян из зрелых сухих стручков. Два теста выявили корреляцию между количеством генов с KO VnSHP и снижением степени осыпания стручков (фиг. 5). Также была обнаружена корреляция ($r=0,86$) между частотой встряхивания для оценки снижения степени осыпания стручков и балльными показателями окрашивания одревесневших слоев (Таблица 7А). Отдельный эксперимент с различными KO-линиями показал аналогичные результаты (Таблица 7В).

Таблица 7А: Сравнение результатов испытания на осыпание стручков и окрашивание одревесневшего слоя - Эксперимент 1

Линия-кандидат ID	Оценка окрашивания одревесневшего слоя (C ₀)	Частота встряхивания для оценки осыпания (C ₁)	Описание гена KO	Генотип
Контроль (WT)	1	16	N/A	N/A
A02-0230	1	18	2 KO	4A/4C
A01-1013	2	16	2 KO	1C/2C
A01-0151	1	16	3 KO	2A/3A/4C
A02-0232	1	20	4 KO	1A/1C/4A/4C
A01-0037	1	18	4 KO	1C/2A/3C/4C
A01-0164	4	18	6 KO	1A/2A/3A/3C/4A/4C
A01-1315	1	20	6 KO	1C/2C/3A/3C/4A/4C
A01-0069	4	30	7 KO	1C/2A/2C/3A/3C/4A/4C
A01-1291	4	24	7 KO	1C/2A/2C/3A/3C/4A/4C
A01-1166	3	30	7 KO	1A/2A/2C/3A/3C/

				4A/4C
A01-0222	5	30	8 КО	1A/1C/2A/2C/3A/ 3C/4A/4C
A01-1187	5	30	8 КО	1A/1C/2A/2C/3A/ 3C/4A/4C
A01-0025	4	30	8 КО	1A/1C/2A/2C/3A/ 3C/4A/4C
A01-0022	4	30	8 КО	1A/1C/2A/2C/3A/ 3C/4A/4C

Таблица 7В: Сравнение результатов испытания на осыпание стручков и окрашивание одревесневшего слоя - Эксперимент 2

Линия ID	Оценка окрашивания одревесневшего слоя (C₀)	Частота встряхивания для оценки осыпания (C₁)	Описание гена КО	Генотип (ген SHP с КО)
Контроль (WT)	1	14,0	N/A	N/A
A05_1098	1	14,3	1КО	1A
A05_1884	1	15,3	1КО	4C
A05_0085	1	15,0	2КО	1A/3A
A05_0641	1	14,3	2КО	3C/4A
A05_1094	1	14,0	2КО	3A/3C
A05_0375	1	14,0	3КО	1C/3A/3C
A05_2790	1	14,5	3КО	2A/4A/4C
A05_0102	2	16,0	4КО	3A/3C/4A/4C
A05_0129	1	16,0	4КО	1C/3A/3C/4A

A05_041 5	1	16,3	4КО	1C/2A/2C/3C
A06_038 7	1	16,5	4КО	2A/3A/3C/4C
A08_078 0	1	15,3	4КО	1C/2C/3C/4A
A05_075 1	3	18,3	5КО	1C/2A/3A/3C/4A
A05_160 8	4	19,0	5КО	2A/3A/3C/4A/4C
A05_189 4	2	16,5	5КО	2A/2C/3A/3C/4C
A08_006 8	2	17,0	5КО	1A/2C/3C/4A/4C
A05_027 7	2	18,0	6КО	1C/2A/2C/3A/3C/4A
A05_121 7	2	17,0	6КО	1C/2A/2C/3A/3C/4C
A05_034 2	3	18,0	7КО	1A/1C/2A/2C/3A/3C/4A
A05_163 5	4	24,0	7КО	1A/1C/2A/2C/3A/4A/4C
A05_348 4	3	22,0	7КО	1A/1C/2A/2C/3A/3C/4C

Методы

[0350] *Окрашивание одревесневшего клеточного слоя флороглюцинолом в зоне растрескивани створок.*

Развивающиеся плоды BN2-SU дикого типа и различных линий SHP с КО (растений C₀) собирали для оценки паттерна одревеснения плодов. Для окрашивания лигнина, поперечные срезы плодов получали с помощью лезвия бритвы, и срезы окрашивали в течение 2 минут в 2% растворе флороглюцинола в 95% этаноле, а затем фотографировали в 66% перхлорной кислоте. Интенсивность окрашивания флороглюцинолом одревесневшего клеточного слоя коррелирует с количеством лигнина в клеточных стенках (Liljegren et al., 2000; Roeder and Yanofsky, 2006), и такая интенсивность оценивается по баллам, как показано ниже в Таблице 8.

Таблица 8: Систематизированная оценка одревеснения

Результат окрашивания флороглюцином	Баллы
Очень темный окрашенный одревесневший слой	1
Одревесневший слой с умеренной темной окраской	2
Светлоокрашенный одревесневший слой	3
Очень светлая окраска - частично отсутствует одревесневший слой	4
Одревесневший слой полностью отсутствует	5

[0351] *Испытание на осыпание стручков с использованием TissueLyser.* Линии-кандидаты переносили в 3,5-дюймовые горшки со средой Sunshine Mix 4 и выдерживали при флуоресцентном освещении для роста T12 при 14-часовом фотопериоде. Через три недели, линии пересаживали в 5,5-дюймовые горшки и переносили в теплицу, где максимальная температура охлаждения установлена на уровне 78°F. Растения выращивали в стандартных условиях выращивания канолы, и для предотвращения ауткроссинга и загрязнения использовали перфорированные пакеты для пыльцы. Растения вынимали из воды при изменении цвета семян на 30% и продолжали сушить в теплице до тех пор, пока семена не достигнут изменения цвета на 100%, то есть, когда растения полностью высыхают. Стручки собирали и помещали в печь на 1 час при 30°C для гарантии обеспечения равномерных уровней влажности во всех образцах.

[0352] Нижеследующие линии-кандидаты выбирали для фенотипирования с помощью теста на резистентность к осыпанию семян (см. также фиг. 5): диплоид BN2 дикого типа (невысушенный), диплоид BN2 дикого типа (высушенный), SHP-A01-0151 (с 2 КО), SHP-A01-0037 (с 3 КО), SHP-A01-1098 (с 5 КО), SHP-A01-0154 (с 6 КО), SHP-A01-0222 (с 7 КО).

[0353] Фенотип резистентности к осыпанию семян определяли по уровню разделения створок, детектируемого при регулируемом встраивании стручков. Для тестирования разделения створок, отдельные стручки помещали в 96-луночный глубокий контейнер с поддоном и закрепляли на кронштейнах TissueLyser II (Qiagen, Germany). Образцы отдельных стручков обрабатывали на устройстве TissueLyser в течение 30 секунд на частотах 22, 23, 24, 25, 26, 27, 28, 29 и 30 Гц. Для каждой линии, четыре копии стручков тестировали на каждой частоте. Фенотип оценивали по шкале баллов 1-4. Неповрежденный стручок получил оценку один, частично расщепленный стручок с соединительными створками получил оценку два, балл три получил стручок с отделением одной створки, а балл четыре получил стручок, в котором обе створки были отделены от перегородки.

[0354] *Тест на осыпание стручков для линий поколения C₁ с КО.* Семена канолы (C₁) прорастали в поддоне, где они оставались в течение 3 недель. Затем растения переносили в 4-дюймовые горшки, а затем выдерживали в теплице в течение 14-часового фотопериода при температуре 23-25°C днем/19-21°C ночью. Стручки собирали с полностью зрелых растений, а затем их сушили в пластиковых контейнерах с отверстиями в течение примерно 3 недель. Перед тестированием, стручки были абсолютно сухими. Три

одинаковых стручка отбирали и помещали в 3 отдельных ячейки модифицированной коробки для встряхивания. Затем коробку загружали в TissueLyser II. Частота встряхивания была установлена на определенное значение (начиная с 12 Гц), а время встряхивания было установлено на 30 секунд. После завершения встряхивания, коробку разгружали, и стручки оценивали на осыпание семян по баллам. Каждую частоту тестировали для 12 стручков на растение. Средний балл для 12 стручков использовали в качестве конечного балла для каждой частоты. Когда показатель встряхивания превышал 2,5, то считалось, соответствующая частота представляет собой заданное значение частоты для осыпания стручков.

[0355] *Испытание на осыпание стручков с использованием Geno/Grinder 2010 (SPEX Sample Prep, США).* Линии-кандидаты переносили в 3,5-дюймовые горшки со средой Sunshine Mix 1 и выдерживали при освещении для роста T12 в течение 16-часового фотопериода и при температуре 21°C днем/19°C ночью для закалки. Через три недели, линии пересаживали в 5,5-дюймовые горшки и переносили в теплицу, где максимальная температура охлаждения была установлена на 78°F. Растения выращивали в стандартных условиях выращивания канолы, и для предотвращения ауткроссинга и загрязнения использовали перфорированные пакеты для пыльцы. Растения вынимали из воды при изменении цвета семян на 30% и продолжали сушить в теплице до тех пор, пока семена не достигнут изменения цвета на 100%, то есть, до тех пор, пока растения полностью не высохнут. Стручки каждого отдельного растения собирали в пластиковый контейнер с крышкой с отверстием в центре, и помещали в печь по меньшей мере на 12 часов при 40°C для гарантии обеспечения равномерных уровней влажности во всех образцах.

[0356] Фенотип резистентности к осыпанию семян определяли по уровню разделения створок, детектируемому при регулируемом встряхивании стручков. Для тестирования разделения створок, 12-24 стручка помещали в 96-луночный глубокий контейнер с поддоном и закреплял на кронштейнах Geno/Grinder 2010 (SPEX Sample Prep, США). Контейнеры, содержащие образцы стручков, обрабатывали в течение 20 секунд при вращении на различных оборотах в минуту (например, при 720, 750, 780, 810, 840, 870, 900, 930, 960, 990, 1020, 1050, 1080 об/мин). По окончании обработки, контейнер вынимали из устройства, и каждому стручку присваивали балл по осыпанию семян как показано в Таблице (Таблица 9). Когда средний показатель осыпания при определенных оборотах превышает 2,5, значение оборотов может рассматриваться как значение для осыпания стручков для указанной линии. Этот метод позволяет обрабатывать большее количество стручков одновременно, и работает гораздо быстрее, чем TissueLyser. Для подтверждения надежности метода Geno/Grinder, был использован негативный контроль и позитивный контроль вместе с обработкой нескольких линий с использованием TissueLyser и Geno/Grinder. Были получены соответствующие оценки осыпания. Значение R для обеих серий данных составляло 0,88, что указывает на то, что данные, собранные с применением обоих методов, в высокой степени коррелируют друг с другом (фиг. 6).

Таблица 9: Систематизированная оценка осыпания семян

Результат анализа со встряхиванием	Баллы
Неповрежденный стручок, который не имеет трещин или видимых повреждений	1
Стручок имеет небольшую трещину на конце, семена не осыпались	2
Есть видимая трещина, которая занимает больше половины стручка, две створки все еще прикреплены к перегородке. Семена могут осыпаться	3
Одна из створок была отделена от перегородки. Семена осыпались	4
Обе створки были отделены от перегородки, и стручок был полностью разрушен.	5

Пример 5: Получение и полевые испытания линий с КО-генами резистентности к осыпанию семян

[0357] В этом примере показаны оценка и выбор наиболее эффективных линий SHP с КО. Поскольку признаки разрушения стручка регулируются множеством генов и могут зависеть от факторов окружающей среды (например, биотических и абиотических), то важно оценить признак разрушения стручка во многих условиях окружающей среды и в течение многих лет.

Результаты

[0358] Резистентность к осыпанию семян из стручков для линий SHP с КО оценивали сначала в теплице (поколение C₀), а затем в двух разных местах в течение двухлетнего периода (поколения C₁ и C₂). Выбранные линии с 5-8 генами SHP с КО неизменно демонстрировали гораздо лучшую эффективность, чем негативный контроль дикого типа (фиг. 7). Линии с КО также показали сходные или лучшие фенотипы уменьшения степени осыпания семян из стручков по сравнению с коммерчески доступными контрольными линиями с резистентностью к осыпанию семян (позитивный контроль). Так, например, линии A05_1635, A05_0342, A05_0277, A05_2013 и A05_0071 были в высокой степени стабильными, и не показали какого-либо аномального фенотипа роста ни в одной из протестированных сред. Эти линии с КО можно рассматривать как линии с истинной резистентностью к осыпанию семян.

[0359] Гены SHP (мутантные или дикого типа) в каждой линии секвенировали для подтверждения наличия мутации в соответствующем гене SHP, где это возможно. Последовательности полного ампликона-мишени области вокруг области-мишени представлены ниже в Таблице 10.

Таблица 10: Последовательности SHP в мутантных линиях

Линия	Ген SHP SEQ ID NO:							
	1A	1C	2A	2C	3A	3C	4A	4C
A05_0071	62	72	82	92	102	112	122	132
A05_2013	63	73	83	93	103	113	123	133

A05_1635	64	74	84	94	104	114	124	134
A05_0342	65	75	85	95	105	115	125	135
A05_0113	66	76	86	96	106	116	126	136
A05_0277	67	77	87	97	107	117	127	137
A05_0272	68	78	88	98	108	118	128	138
A05_1600	69	79	89	99	109	119	129	139
A05_0751	70	80	90	100	110	120	130	140
A05_1894	71	81	91	101	111	121	131	141

Методы

[0360] Белок CRISPR/Cas9, образующий комплекс с рРНК (RNP, Таблица 5), наряду с одноцепочечными олигонуклеотидами (GRON) (Таблица 4), использовали для нокаутирования генов *VnSHP* в другой линии канолы с фоновым фенотипом дикого типа. CRISPR/Cas9 состоит из двух компонентов: Cas9 *Streptococcus pyogenes*, оптимизированного по кодонам растений (*SpCas9*), и оцрРНК, которые экспрессировались в виде белка и РНК, соответственно. оцрРНК *in vitro* транскрибировалась из ДНК-матрицы и представляла собой гибрид РНК CRISPR (*cr*-РНК) и транс-активирующей *cr*-РНК (*tracr*-РНК). Область *cr*-РНК содержит спейсерные последовательности, описанные в Таблице 5, которые были использованы для нацеливания нуклеазного белка Cas9 на каждый из генов-мишеней *SHP*. GRON содержат кодирующую последовательность целевых генов *SHP* рядом с сайтом конверсии и несут точные геноспецифические мутации (n+1, n-1 и n-2), а также помечены группой 2'-O-Me в первом 5'-основании, которое является основанием РНК, вместо основания ДНК (Таблица 4).

[0361] RNP и GRON были введены в протопласты посредством ПЭГ-опосредованной доставки в конечной концентрации 1,0 мкг/мкл и 0,05 мкМ, соответственно. Перед доставкой в протопласты, рекомбинантный белок Cas9 был подвергнут реакции образования комплекса *in vitro* с рРНК. Протопласты канолы выделяли из листьев растений, подвергнутых микроразмножению *in vitro*, в соответствии со стандартным протоколом. Протопласты культивировали в жидкой среде ($1,25 \times 10^5$ клеток/мл) и инкубировали в темноте при 25°C. Образцы клеток получали через одну или три недели и анализировали методом глубокого секвенирования для определения частоты мутаций в генах-мишенях. Через 6-8 недель, полученные из протопласта микрокалусы переносили в твердую среду для регенерации, в результате чего, приблизительно через 2-4 недели из регенерированных калусов начали образовываться побеги. Образцы листьев от полностью дифференцированных побегов были проанализированы с помощью NGS для определения встречаемости целевых мутаций в каждом из 8 генов *SHP*. Побеги с целевыми мутациями в отдельных генах и во множестве генов, включая все 255 возможных комбинаций генов или генотипов с КО, затем подвергали скринингу на ploидность. Диплоидные растения подвергали микроразмножению *in vitro* и переносили в

почву в камере для культивирования. Закаленные растения (C_0) переносили в теплицу и выращивали до созревания (завязывания семян).

[0362] Семена, собранные у растений C_0 , называются поколением C_1 . По меньшей мере 3 растения C_0 для каждой комбинации КО отбирали и выращивали, как указано в Примере 4. Во время закалки собирали образцы листьев, и генотипы растений C_0 подтверждали с помощью NGS. Семена C_1 прорастали в поддонах (5 растений на линию), и образцы листьев собирали через 10-12 дней после посадки для подтверждения генотипа. Через три недели после посадки, растения пересаживали в 5,5-дюймовые горшки и переносили в теплицу, где максимальная температура охлаждения была установлена на 78°F. Растения выращивали в стандартных условиях выращивания канолы, и для предотвращения ауткроссинга и загрязнения использовали перфорированные пакеты для пыльцы. Растения вынимали из воды при изменении цвета семян на 30% и продолжали сушить в теплице до тех пор, пока семена не достигнут изменения цвета на 100%, то есть, до тех пор, пока растения полностью не высохнут. Стручки растений C_0 или C_1 , выращенных в теплице, собирали и помещали в печь при 40°C по меньшей мере на 12 часов, чтобы обеспечить равномерный уровень влажности во всех образцах. Осыпание семян из стручков оценивали методом TissueLyser. Выбранные линии также тестировали в полевых условиях в двух различных регионах: одну в Калифорнии и одну в Северной Дакоте. Эксперименты осуществляли на делянках с абсолютно случайным расположением (RCBD) с тремя повторностями. Семена C_1 и C_2 обрабатывали фунгицидом (Helix, Bayer Crop Science) и высевали непосредственно в почву. В условиях выращивания, растения доводили до зрелости, и стручки собирали в виде полного образца для каждой линии и от каждого дубликата. Фенотипы осыпания семян из стручков оценивали на TissueLyser или на Geno/Grinder. Данные собирали и анализировали.

Ссылки

Bohanec B (2003) Ploidy Determination Using Flow Cytometry. In: Maluszynski M, Kasha KJ, Forster Bp, Szarejko I (Eds) Doubled Haploid Production In Crop Plants: A Manual. Kluwer, Dordrechts, Pp 397-403.

к осыпанию семян.

11. Растение или клетка растения, содержащие одну или более мутаций в одном или более генах SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C.

12. Выделенная нуклеиновая кислота, имеющая последовательность одного из SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C или их фрагментов;

или имеющая последовательность, которая по меньшей мере на 90% идентична последовательности одного из SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C;

или имеющая последовательность, которая по меньшей мере на 95% идентична последовательности одного из SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C;

или имеющая последовательность, которая по меньшей мере на 98% идентична последовательности одного из SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C;

или имеющая 10, 9, 8, 7, 6, 5, 4, 3, 2 или 1 нуклеотидных замен по сравнению с последовательностью одного из SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C.

13. Выделенная аминокислотная последовательность, кодируемая SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C или их фрагментами;

или кодируемая последовательностью, которая по меньшей мере на 90% идентична последовательности SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C;

или кодируемая последовательностью, которая по меньшей мере на 95% идентична последовательности SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C;

или кодируемая последовательностью, которая по меньшей мере на 98% идентична последовательности SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C;

или кодируемая последовательностью, которая дает 5, 4, 3, 2 или 1 аминокислотных замен по сравнению с последовательностью, кодируемой SHP1A, SHP1C, SHP2A, SHP2C, SHP3A, SHP3C, SHP4A или SHP4C.

14. Растение или клетка растения, содержащие ген резистентности к осыпанию семян, имеющий последовательность, которая отличается от любой из SEQ ID NO: 1-16.

15. Растение или клетка растения, содержащие по меньшей мере два гена резистентности к осыпанию семян, имеющих последовательность, которая отличается от любой из SEQ ID NO: 1-8.

16. Растение или клетка растения, содержащие по меньшей мере три гена резистентности к осыпанию семян, имеющих последовательность, которая отличается от любой из SEQ ID NO: 1-8.

17. Растение или клетка растения, содержащие по меньшей мере четыре гена резистентности к осыпанию семян, имеющих последовательность, которая отличается от любой из SEQ ID NO: 1-8.

18. Растение или клетка растения, содержащие по меньшей мере пять генов резистентности к осыпанию семян, имеющих последовательность, которая отличается от любой из SEQ ID NO: 1-8.

19. Растение или клетка растения, содержащие по меньшей мере шесть генов резистентности к осыпанию семян, имеющих последовательность, которая отличается от любой из SEQ ID NO: 1-8.

20. Растение или клетка растения, содержащие по меньшей мере семь генов резистентности к осыпанию семян, имеющих последовательность, которая отличается от любой из SEQ ID NO: 1-8.

21. Растение или клетка растения, содержащие по меньшей мере восемь генов резистентности к осыпанию семян, имеющих последовательность, которая отличается от любой из SEQ ID NO: 1-8.

22. Растение по любому из предшествующих пунктов, где указанными различиями в указанной последовательности или указанной мутацией являются инсерция или делеция.

23. Растение по любому из предшествующих пунктов, где указанными различиями в указанной последовательности или указанной мутацией является одна нуклеотидная замена.

24. Растение по любому из предшествующих пунктов, где указанные различия в указанной последовательности или указанная мутация включают множество нуклеотидных замен.

25. Растение по любому из предшествующих пунктов, где указанным различием в указанной последовательности или указанной мутацией является введение преждевременного стоп-кодона.

26. Растение по любому из предшествующих пунктов, где указанным различием в указанной последовательности или указанной мутацией является снижение восприимчивости растения к растрескиванию плодов до сбора урожая.

27. Растение или клетка растения или способ по любому из предшествующих пунктов, где активность или экспрессия белка, экспрессируемого модифицированным или мутированным геном резистентности к осыпанию семян, является сниженной или отсутствует по сравнению с соответствующим полноразмерным белком резистентности к осыпанию семян дикого типа.

28. Растение или клетка растения или способ по любому из предшествующих пунктов, где указанное растение или клетка растения не содержат какого-либо трансгена.

29. Способ, где указанный способ включает введение мутации в ген резистентности к осыпанию семян в растении.

30. Способ, где указанный способ включает контактирование клетки с ДНК-разрезающим ферментом, структура которого является подходящей для специфического одноцепочечного разрыва или разрезания гена резистентности к осыпанию семян.

31. Способ, где указанный способ включает контактирование клетки с CRISPR, TALEN, цинковым пальцем или мегануклеазой, структура которых является подходящей

для специфического одноцепочечного разрыва или разрезания гена резистентности к осыпанию семян.

32. Способ индуцирования генетической модификации в клетке растения, где указанный способ включает обработку указанной клетки GRON, кодирующим по меньшей мере одну мутацию в гене резистентности к осыпанию семян.

33. Способ предотвращения или снижения растрескивания плодов перед сбором урожая, где указанный способ включает мутацию по меньшей мере одного эндогенного гена резистентности к осыпанию семян в клетке указанного растения.

34. Способ получения растения или клетки растения, включающих мутацию в гене SHP, где указанный способ включает

(1) введение в клетки растения олигонуклеосложения для репарации генов с целевой мутацией в гене SHP для получения клеток растений с мутантным геном SHP; и

(2) регенерацию растения, имеющего мутированный ген SHP, из указанной выбранной клетки растения.

35. Способ введения мутации в ген SHP, где указанный способ включает обработку клетки ДНК-разрезающим ферментом.

36. Способ введения мутации в ген SHP, где указанный способ включает обработку клетки ДНК-разрезающим ферментом и GRON.

37. Способ введения мутации в ген SHP, где указанный способ включает обработку клетки ДНК-разрезающим ферментом и GRON, где указанный GRON модифицирован одной или более группами Су3, группами ЗPS и 2'О-метильными группами.

38. Клетка растения, которая включает ДНК-разрезающий фермент и GRON (такой как GRON, который связывается с геном SHP и/или модифицирует этот ген), например, где GRON модифицирован такими группами, как группа Су3, группа ЗPS, 2'О-метильная группа или другая модификация. В некоторых вариантах осуществления изобретения, ДНК-разрезающий фермент представляет собой один или более ферментов, выбранных из CRISPR, TALEN, цинкового пальца, мегануклеазы и ДНК-разрезающего антибиотика.

39. Способ по любому из предшествующих пунктов, где указанный способ не включает контактирования указанного растения или клетки растения с каким-либо трансгеном.

40. Способ по любому из предшествующих пунктов, где растение, выращенное указанным способом, является нетрансгенным.

41. Растение или способ по любому из предшествующих пунктов, где растением является растение капусты, *Brassica napus* (канола, масличный рапс), *Brassica rapa* (например, турнепс, капуста китайская), *Brassica oleracea* (брокколи, капуста белокочанная, цветная капуста и т.п.), *Brassica juncea* (горчица) или *Raphanus sativus* (редька обыкновенная), а также многие ценные бобовые культуры, такие как горох, фасоль, чечевица и соя.

42. Способ или композиция по любому из предшествующих пунктов, где указанная мутация в гене SHP, если она присутствует, снижает или устраняет активность или

экспрессию гена SHP.

43. Способ или композиция по любому из предшествующих пунктов, где указанная мутация в гене SHP, если она присутствует, представляет собой инсерцию или делецию.

44. Способ или композиция по любому из предшествующих пунктов, где указанная мутация в гене SHP, если она присутствует, представляет собой инсерцию или делецию, которая снижает или устраняет активность или экспрессию гена SHP.

45. Способ или композиция по любому из предшествующих пунктов, где указанная мутация в гене SHP, если она присутствует, представляет собой модификацию или замену нуклеотида.

46. Способ или композиция по любому из предшествующих пунктов, где указанная мутация в гене SHP, если она присутствует, представляет собой модификацию или замену нуклеотида, которая снижает или устраняет активность или экспрессию гена SHP.

47. Растение или его часть, содержащие по меньшей мере одну мутацию по меньшей мере в пяти, по меньшей мере в шести, по меньшей мере в семи или в восьми последовательностях нуклеиновых кислот, кодирующих гены SHATTERPROOF (SHP).

48. Растение или его часть по п. 47, где последовательности нуклеиновой кислоты имеют последовательность, которая по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 98% или по меньшей мере на 99% идентична последовательностям нуклеиновой кислоты, выбранным из группы, состоящей из SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 и SEQ ID NO: 8.

49. Растение или его часть по п. 47, где последовательности нуклеиновой кислоты выбраны из группы, состоящей из SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 и SEQ ID NO: 8.

50. Растение или его часть по любому из п.п. 47-49, где мутация представляет собой мутацию со сдвигом рамки считывания.

51. Растение или его часть по п. 50, где мутация со сдвигом рамки считывания приводит к одной или более инсерциям или делециям нуклеотидов по сравнению с соответствующим эндогенным геном без мутации со сдвигом рамки считывания.

52. Растение или его часть по п. 50 или 51, где мутация со сдвигом рамки считывания приводит к встраиванию преждевременного стоп-кодона.

53. Растение или его часть по любому из п.п. 50-52, где мутация снижает или устраняет экспрессию гена SHP и/или полипептида SHP.

54. Растение или его часть по любому из п.п. 47-53, где растение обладает пониженной восприимчивостью к растрескиванию плодов до сбора урожая.

55. Растение по любому из п.п. 47-54, где растение выбрано из группы, состоящей из видов *Brassica napus*, *Brassica rapa*, *Brassica oleracea*, *Brassica juncea*, *Brassica*, *Raphanus sativus*, *Pisum sativum*, *Phaseolus vulgaris*, *Lens culinaris*, *Glycine max* и *Fabaceae*.

56. Способ выращивания растения по любому из п.п. 47-55, включающий стадии:

а) введения мутаций в клетки растения, где мутации представляют собой по меньшей мере одну мутацию по меньшей мере в пяти, по меньшей мере в шести, по

меньшей мере в семи или в восьми последовательностях нуклеиновых кислот, кодирующих гены SHP;

б) отбора клеток растения, содержащих мутации; и

с) регенерации растения, имеющего мутации;

где растение обладает пониженной восприимчивостью к растрескиванию плодов до сбора урожая.

57. Способ по п. 56, где мутации вводят с использованием одного или более векторов, где векторы содержат компоненты редактирования генов, выбранные из группы, состоящей из системы CRISPR/Cas9, TALEN, цинкового пальца и мегануклеазы, предназначенных для нацеливания на последовательность нуклеиновой кислоты, кодирующую ген SHP.

58. Способ по п. 56, где мутации вводят с использованием системы GRON, предназначенной для нацеливания на последовательность нуклеиновой кислоты, кодирующую ген SHP.

59. Способ по п. 58, где система GRON включает одну или более модификаций, выбранных из группы, состоящей из группы $Cu3$, группы 3PS и 2'О-метильной группы.

60. Способ по любому из п.п. 56-59, где последовательности нуклеиновой кислоты имеют последовательность, которая по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 98% или по меньшей мере на 99% идентична последовательностям нуклеиновой кислоты, выбранным из группы, состоящей из SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 и SEQ ID NO: 8.

61. Способ по любому из п.п. 56-59, где последовательности нуклеиновой кислоты выбраны из группы, состоящей из SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 и SEQ ID NO: 8.

62. Способ по любому из п.п. 56-61, где мутация выбрана из группы, состоящей из мутации со сдвигом рамки считывания, мутации со сдвигом рамки считывания, приводящей к одной или более инсерциям или делециям нуклеотидов по сравнению с соответствующим эндогенным геном без мутации со сдвигом рамки считывания, и мутации со сдвигом рамки считывания, приводящей к встраиванию преждевременного стоп-кодона, и где мутация уменьшает или устраняет экспрессию гена SHP и/или полипептида SHP.

63. Способ по любому из п.п. 56-62, где растение выбрано из группы, состоящей из видов *Brassica napus*, *Brassica rapa*, *Brassica oleracea*, *Brassica juncea*, *Brassica*, *Raphanus sativus*, *Pisum sativum*, *Phaseolus vulgaris*, *Lens culinaris*, *Glycine max* и *Fabaceae*.

64. Растение F_1 , где растение F_1 представляет собой растение по любому из п.п. 47-55 как родительское растение.

65. Способ получения семян растений, где указанный способ включает скрещивание растения по любому из п.п. 47-55 с другим растением и сбор семян этих растений.

66. Способ получения растения по любому из п.п. 47-55, где указанный способ

включает отбор семян после скрещивания растения по любому из п.п. 47-55 с растением по любому из п.п. 47-55 с получением растения по любому из п.п. 47-55.

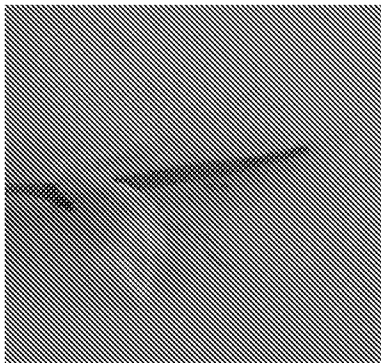
67. Растение, полученное путем выращивания семян по п.п. 65 или 66, где растение обладает всеми физиологическими и морфологическими свойствами растения по любому из п.п. 47-55.

По доверенности

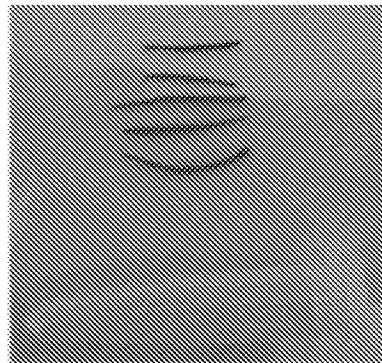
ФИГ. 1В

```
AtSHP1 MEEGGSSHDAESSKKLGSGKIEIKRIENTTNRQVTFCKRRNGLLKKA YELSVLCDAEVALVIFSTRGRLYEYANNR...
AtSHP2 MEGGASNEVAESSKKIGRGKIEIKRIENTTNRQVTFCKRRNGLLKKA YELSVLCDAEVALVIFSTRGRLYEYANNR...
BnSHP-1A MEEGGSSHDAESSKKIGRGKIEIKRIENTTNRQVTFCKRRNGLLKKA YELSVLCDAEVALVIFSTRGLLYEYASNR...
BnSHP-1C MEEGGSSHDAESSKKIGRGKIEIKRIENTTNRQVTFCKRRNGLLKKA YELSVLCDAEVALVIFSTRGELYEYASNR...
BnSHP-2A MEEGGSSHDAESSKKLVRSKIEIKRIENTTNRQVTFCKRRNGLLKKA YELSVLCDAEVALVIFSTRGRLYEYANNR...
BnSHP-2C MEEGGSSHDAESSKKLVRSKIEIKRIENTTNRQVTFCKRRNGLLKKA YELSVLCDAEVALVIFSTRGRLYEYANNR...
BnSHP-3C MEEGGSSHDAESSKKLVRSKIEIKRIENTTNRQVTFCKRRNGLLKKA YELSVLCDAEVALVIFSTRGELYEYANNR...
BnSHP-3A MEEGGSSHDAESSKKLVRSKIEIKRIENTTNRQVTFCKRRNGLLKKA YELSVLCDAEVALVIFSTRGRLYEYANNR...
BnSHP-4A MEGGASDEVAESSKKIGRGKIEIKRIENTTNRQVTFCKRRNGLLKKA YELSVLCDAEVALVIFSTRGRLYEYANNR...
BnSHP-4C MEGGASDEVAESSKKIGRGKIEIKRIENTTNRQVTFCKRRNGLLKKA YELSVLCDAEVALVIFSTRGRLYEYANNR...
*: ..*.. **.*: *****.*****.*****.*****.*****.*****.*
```

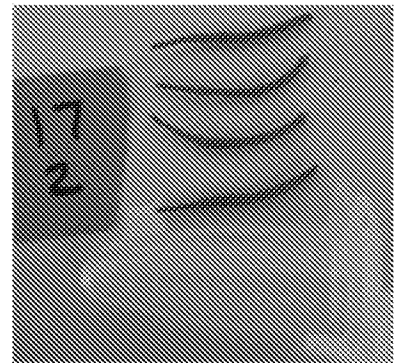
ФИГ. 2А



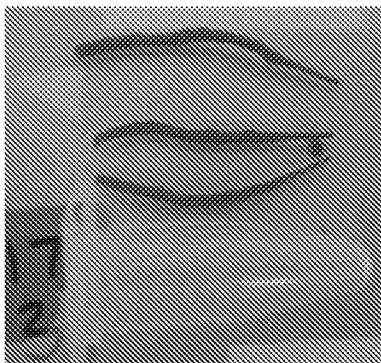
13



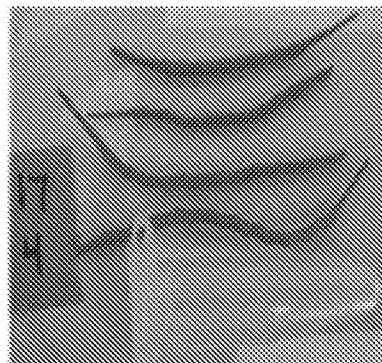
17-1



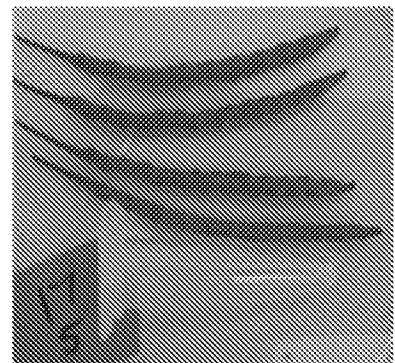
17-2



17-3

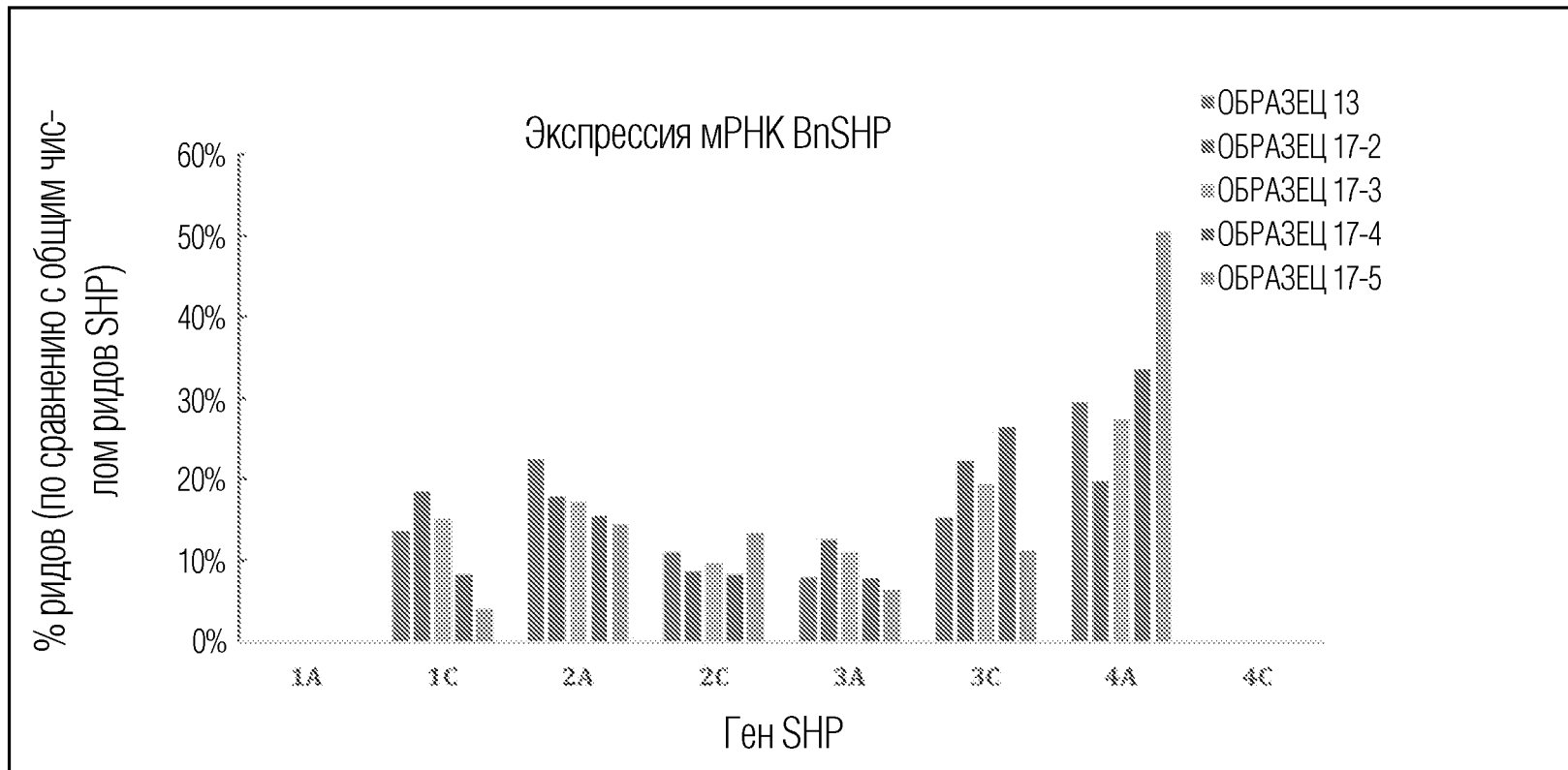


17-4

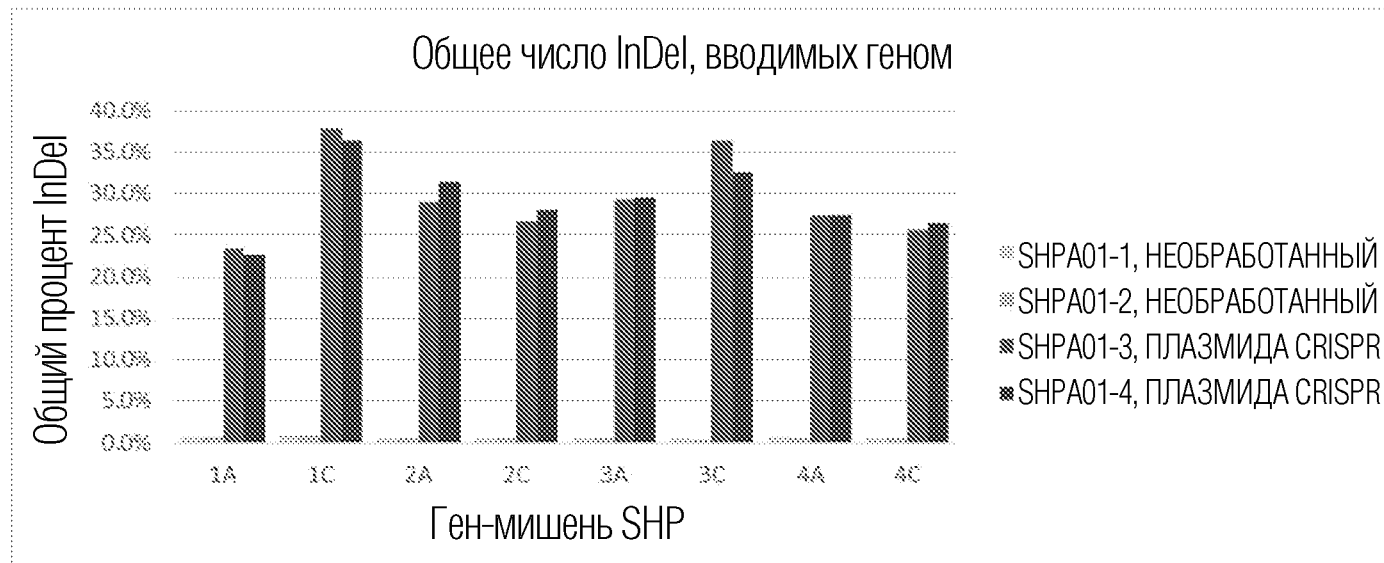


17-5

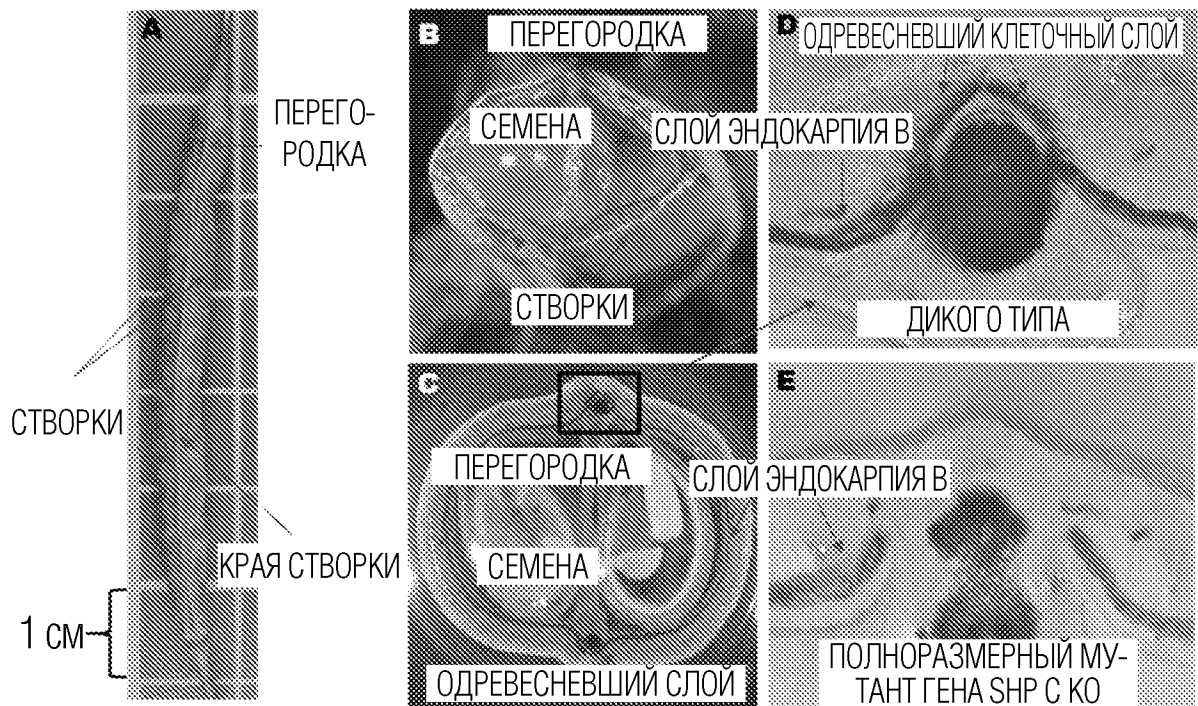
ФИГ. 2В



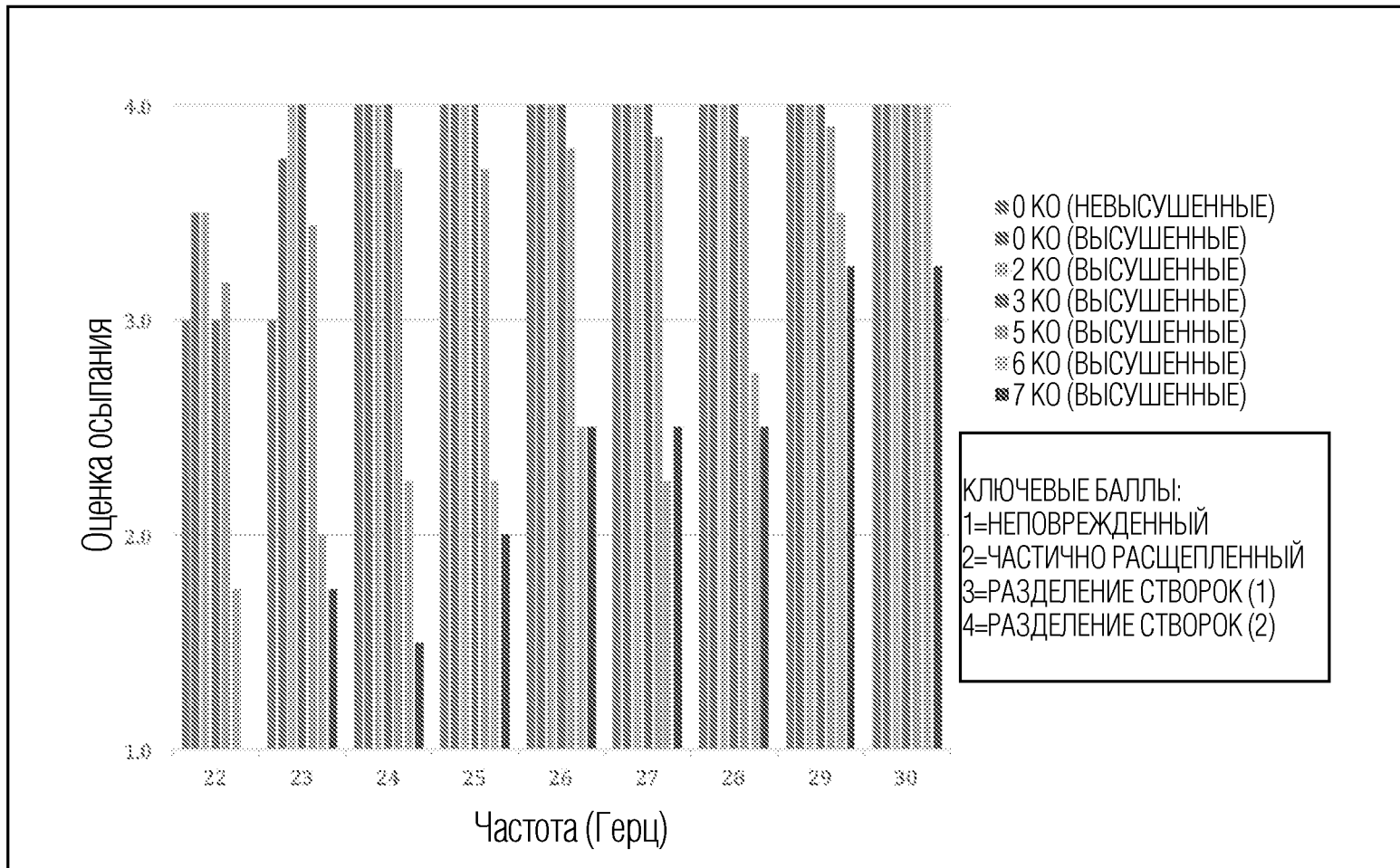
ФИГ. 3



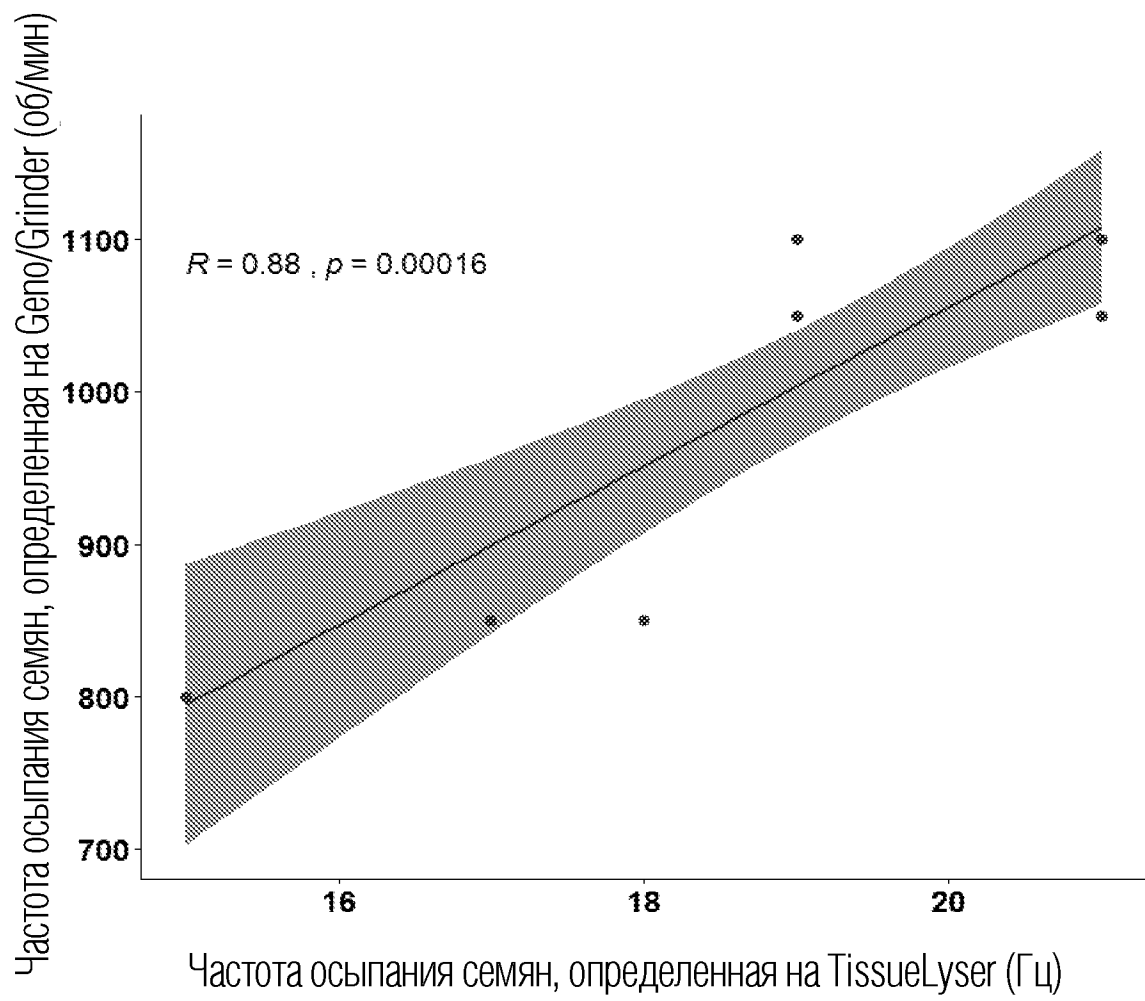
ФИГ. 4



ФИГ. 5



ФИГ. 6



ФИГ. 7

ЛИНИЯ	КОНФИГУРАЦИЯ ГЕНА SHP С КО (ГЕНОТИП)								ОПИСАНИЕ ГЕНА С КО	СО В ТЕПЛИЦЕ	С1 В ТЕПЛИЦЕ	1-Й ПОЛЕВОЙ ТЕСТ ДЛЯ С1	2-Й ПОЛЕВОЙ ТЕСТ ДЛЯ С1/С2
	1A	1C	2A	2C	3A	3C	4A	4C		ЧАСТОТА НА TISSUE-LYSER (Гц)	ЧАСТОТА НА TISSUE-LYSER (Гц)	ЧАСТОТА НА TISSUE-LYSER (Гц)	УСТРОЙСТВО GENO/GRINDER (ОБ./МИН.)
A05_0071	1	1	1	-2	1	1	1	-2	8КО	17	22	n/a	1005
A05_2013	1	1	1	1	1	1	1	1	8КО	15.3	20	n/a	1065
A05_1635	1	1	1	1	-1	wt	1	1	7КО	24	26	n/a	996
A05_0342	1	1	1	-2	1	1	-2	wt	7КО	18	22	20.3	1078
A05_0113	1	-2	-1	1	1	1	1	wt	7КО	16	20	19	1033
A05_0277	wt	1	-2	1	1	1	1	wt	6КО	18	25	20.3	990
A05_0272	wt	1	-2	1	1	1	1	wt	6КО	16.7	19.3	19	978
A05_1600	wt	1	1	1	wt	1	-2	wt	5КО	15.7	20	n/a	900
A05_0751	wt	1	1	wt	-1	1	1	wt	5КО	18.3	19.3	20.3	818
A05_1894	wt	wt	1	1	1	1	wt	-1	5КО	16.5	19.3	n/a	966
КОНТРОЛЬНАЯ КОММЕРЧЕСКИ ДОСТУПНАЯ ЛИНИЯ 1									ПОЗИТИВНЫЙ КОНТРОЛЬ	17.8	25.3	19.7	995
КОНТРОЛЬНАЯ КОММЕРЧЕСКИ ДОСТУПНАЯ ЛИНИЯ 2									ПОЗИТИВНЫЙ КОНТРОЛЬ	17.5	25.3	19.7	970
ДИКОГО ТИПА									НЕГАТИВНЫЙ КОНТРОЛЬ	14.3	16.7	16.7	744