

(19)



**Евразийское
патентное
ведомство**

(11) **047686**(13) **B9**

**(12) ИСПРАВЛЕННОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ К
ЕВРАЗИЙСКОМУ ПАТЕНТУ**

(15) Информация об исправлении
Версия исправления: 1 (W1 B1)
исправления в формуле: п.6

(51) Int. Cl. *C12N 15/82* (2006.01)
A01H 6/46 (2018.01)

(48) Дата публикации исправления
2024.09.06, Бюллетень №9'2024

(45) Дата публикации и выдачи патента
2024.08.26

(21) Номер заявки
202193108

(22) Дата подачи заявки
2020.07.01

**(54) ГЕН DMP ИНДУКЦИИ ПАРТЕНОГЕНЕТИЧЕСКОГО ГАПЛОИДА И ЕГО
ПРИМЕНЕНИЕ**

(31) **201910445082.3**

(32) **2019.05.27**

(33) **CN**

(43) **2022.04.05**

(86) **PCT/CN2020/099680**

(87) **WO 2020/239137 2020.12.03**

(71)(73) Заявитель и патентовладелец:
**ЧАЙНА АГРИКАЛЧУРАЛ
ЮНИВЕРСИТИ (CN)**

(72) Изобретатель:
**Чэнь Шаоцзян, Чжун Юй, Лю
Чэньсюй, Ци Сяолун, Ли Мэнжэнь,
Чэнь Баоцзянь, Цзяо Яньянь, Лю
Цзункай (CN)**

(74) Представитель:
**Поликарпов А.В., Соколова М.В.,
Путинцев А.И., Черкас Д.А., Игнатъев
А.В., Билык А.В., Дмитриев А.В.,
Бучака С.М., Бельтюкова М.В. (RU)**

(56) Yu Zhong et al. "A DMP-triggered in vivo Maternal Haploid Induction System in the Dicotyledonous Arabidopsis" Nature Plants, Vol. 6, No. 5, 15 May 2020 (2020-05-15), ISSN: 2055-0278, see entire document

Yu Zhong et al. "Mutation of ZmDMP Enhances Haploid Induction in Maize" Nature Plants, Vol. 5, No. 6, 10 June 2019 (2019-06-10), ISSN: Nature Plants, see entire document

Philipp Cyprys et al. "Gamete Fusion is Facilitated by Two Sperm Cell-Expressed DUF679 Membrane Proteins" Nature Plants, Vol. 5, No. 3, 08 March 2019 (2019-03-08), ISSN: 2055-0278, see the abstract

Theologis, A. et al. "Accession number: NM_001035933.2" GenBank, 14 February 2019 (2019-02-14), see sequence list

Tabata, S. et al. "Accession number: NM_123327.2" GenBank, 14 February 2019 (2019-02-14), see sequence list

"Accession number: XM_004239348.3" GenBank, 08 August 2018 (2018-08-08), see sequence list

CN-A-109628480

CN-A-108192912

CN-A-106164254

(57) Предложены ген DMP индукции партеногенетического гаплоида и его применение. Гены AtDMP8 и AtDMP9 индукции партеногенетического гаплоида клонированы из *Arabidopsis thaliana*. Эксперименты показали то, что мутации AtDMP8 и AtDMP9 могут давать индуцибельность партеногенетического гаплоида с обеспечением индукции двудольных сельскохозяйственных культур к образованию гаплоидов посредством партеногенетических способов. Изобретение было дополнительно подтверждено на томатах, и у томатов также обнаружили, что мутация SIDMP может давать индуцибельность партеногенетического гаплоида. Изобретение служит важной основой для расширения применения технологии селекции гаплоидов на двудольные растения и выявления биологического механизма получения партеногенетических гаплоидов. Принимая во внимание универсальность применения технологии селекции гаплоидов в современном промышленном разведении, изобретение имеет очень широкую сферу применения и перспективы на рынке.

B9**047686****047686****B9**

Область изобретения

Настоящее изобретение относится к области сельскохозяйственной биотехнологии и генетической селекции сельскохозяйственных культур на основе технологии редактирования генома, в частности, к способу получения линии-индуктора материнского гаплоидного растения и ее применению, в частности, к применению мутанта по гену DMP индукции партеногенетического гаплоида, полученного с использованием технологии редактирования генома, в качестве линии-индуктора гаплоидного растения в индукции растения к генерации материнского гаплоида.

Предшествующий уровень техники

Селекция элитных самоопыляемых линий является основой и ключом для селекции элитного гибрида с использованием гетерозиса сельскохозяйственных культур. Однако, для получения более стабильной самоопыляемой линии требуется 7-8 поколений при использовании традиционных способов селекции и только 2 поколения при использовании способа гаплоидной селекции (Weber D F, 2014), что значительно сокращает цикл селекции. В настоящее время методика культуры пыльников *in vitro* является наиболее предпочтительной для стимулирования продукции гаплоида у двудольных сельскохозяйственных культур, которая имеет недостатки, заключающиеся в низкой эффективности, высокой зависимости от генотипа материалов и трудности при крупномасштабном применении. Хотя введение генетически модифицированного варианта центромер-специфичного гистона CENH3 в мутант *cenh3 Arabidopsis thaliana* и индуцирует продукцию гаплоидов, данный подход дает большое число эуплоидов во время индукции (Ravi, M, et al., 2010).

В настоящее время система технологии редактирования генов достигла высокого уровня и стабильности, однако она не может далее оказывать свой максимальный эффект в виду ограничения генетической трансформации из-за зависимости от материала и генотипа. Хотя редактирование может быть достигнуто посредством введения вектора редактирования генов в какой-либо легко трансформируемый материал-реципиент и его дальнейшей гибридизации с подлежащей редактированию мишенью (Li C et al., 2017), возвратное скрещивание для восстановления фона занимает длительное время и не может достигать 100%. Индукцию гаплоидов и методики редактирования генов объединяют для достижения редактирования генов целевого гаплоидного материала для получения гомозиготных редактируемых линий, что значительно сокращает по времени и расширяет диапазон применения гаплоидов и методик редактирования генов (Kelliher, T et al., 2019; Wang, B et al., 2019; Hu, N et al., 2019). Однако у двудольных растений индукция гаплоидов не может быть объединена с методиками редактирования генов из-за отсутствия способа получения гаплоида посредством биологической индукции.

Краткое изложение сущности изобретения

Задача настоящего изобретения заключается в том, чтобы предложить способ получения линии-индуктора гаплоидного растения.

Способ получения линии-индуктора гаплоидного растения, предложенный в данном изобретении, является следующим из А1) или А2):

А1) сайленсинг или ингибирование экспрессии и/или активности гена DMP в геноме растения или нокаутирование гена DMP с получением линии-индуктора гаплоидного растения;

А2) ингибирование активности белка DMP в растении с получением линии-индуктора гаплоидного растения;

данное растение является двудольным;

белок DMP представляет собой белок, представленный следующими из Б1), или Б2), или Б3), или Б4):

Б1) белок с аминокислотной последовательностью, показанной в SEQ ID NO: 2, или SEQ ID NO: 4, или SEQ ID NO: 6;

Б2) слитый белок, полученный путем присоединения метки к N-концу и/или C-концу белка, показанного в SEQ ID NO: 2, или SEQ ID NO: 4, или SEQ ID NO: 6;

Б3) белок, имеющий такую же функцию, полученный путем замены и/или делеции и/или вставки одного или нескольких аминокислотных остатков в аминокислотной последовательности, показанной в SEQ ID NO: 2, или SEQ ID NO: 4, или SEQ ID NO: 6;

Б4) белок, имеющий гомологию 75% или более с аминокислотной последовательностью, показанной в SEQ ID NO: 2, или SEQ ID NO: 4, или SEQ ID NO: 6, и имеющий такую же функцию;

данный белок DMP имеет следующие функции: двудольное растение становится линией-индуктором гаплоидного растения при ингибировании в нем активности белка DMP. Активность белка DMP ингибируется таким образом, что данный белок DMP не экспрессируется или является неактивным.

Метка в вышеупомянутом Б2) относится к полипептиду или белку, экспрессируемому при слиянии с интересующим белком с использованием генной инженерии ДНК *in vitro*, для облегчения экспрессии, выявления, отслеживания и/или очистки интересующего белка. Белковая метка может представлять собой метку Flag, метку His, метку MBP (основной белок миеллина), метку HA, метку Мус, метку GST (глутатион-S-трансфераза) и/или метку SUMO.

Замена и/или делеция и/или вставка одного или нескольких аминокислотных остатков в вышеупомянутом Б3) представляет собой замену и/или делецию и/или вставку не более чем 10 аминокислотных

остатков.

Гомология 75% или более в вышеупомянутом Б4) может составлять 80, 85, 90 или 95% или более.

Согласно способу получения линии-индуктора гаплоидного растения, ген DMP представляет собой ген, показанный в следующих из В1), или В2), или В3), или В4):

В1) молекула кДНК или молекула геномной ДНК, показанная в SEQ ID NO: 1, или SEQ ID NO: 3, или SEQ ID NO: 5;

В2) молекула кДНК или молекула геномной ДНК, имеющая идентичность 75% или более с нуклеотидной последовательностью, определенной В1);

В3) молекула кДНК или молекула геномной ДНК, происходящая из двудольного растения и имеющая идентичность 75% или более с нуклеотидной последовательностью, определенной в В1);

В4) молекула кДНК или молекула геномной ДНК, которая гибридизуется с нуклеотидной последовательностью, определенной в В1), или В2), или В3), в жестких условиях.

Ген DMP имеет следующие функции: двудольное растение становится линией-индуктором гаплоидного растения, когда ген DMP подвергается сайленсингу, или ингибированию, или нокаутированию в нем.

Термин "идентичность", как он использован здесь, относится к сходству последовательности с последовательностью природной нуклеиновой кислоты. "Идентичность" включает нуклеотидные последовательности, имеющие идентичность 75% или более, или 85% или более, или 90% или более, или 95% или более с нуклеотидной последовательностью, кодирующей белок, состоящий из аминокислотной последовательности, показанной в SEQ ID NO: 2, или SEQ ID NO: 4, или SEQ ID NO: 6 по настоящему изобретению. Идентичность можно оценивать невооруженным глазом или с помощью компьютерной программы. С использованием компьютерной программы идентичность между двумя или более последовательностями может быть выражена в виде процентной доли (%), которую можно использовать для оценки идентичности между родственными последовательностями.

В способе получения линии-индуктора гаплоидного растения способ сайленсинга или ингибирования экспрессии и/или активности гена DMP или нокаутирования гена DMP в геноме растения включает уменьшение уровня экспрессии гена DMP в геноме растения или подвергание гена DMP в геноме растения делеционной мутации, или инсерционной мутации, или замене основания.

Кроме того, когда двудольное растение представляет собой *Arabidopsis thaliana*, ген DMP представляет собой ген AtDMP8 и/или ген AtDMP9; белок DMP представляет собой белок AtDMP8 и/или белок AtDMP9;

данный способ включает сайленсинг или ингибирование экспрессии и/или активности гена AtDMP8 и/или гена AtDMP9 в геноме *Arabidopsis thaliana*, или нокаутирование гена AtDMP8 и/или гена AtDMP9, или ингибирование активности белка AtDMP8 и/или белка AtDMP9 с получением линии-индуктора гаплоида *Arabidopsis thaliana*.

Когда двудольное растение представляет собой томат, ген DMP представляет собой ген SIDMP; белок DMP представляет собой белок SIDMP; данный способ включает сайленсинг или ингибирование гена SIDMP, или нокаутирование гена SIDMP, или ингибирование активности белка SIDMP в геноме томата с получением линии-индуктора гаплоида томата.

Кроме того, сайленсинг или ингибирование экспрессии и/или активности гена AtDMP8 и/или гена AtDMP9 в геноме *Arabidopsis thaliana*, или нокаутирование гена AtDMP8 и/или гена AtDMP9 включает уменьшение экспрессии гена AtDMP8 и/или гена AtDMP9 в геноме *Arabidopsis thaliana* или подвергание гена AtDMP8 и/или гена AtDMP9 в геноме *Arabidopsis thaliana* делеционной мутации, или инсерционной мутации, или замене основания.

Сайленсинг или ингибирование экспрессии и/или активности гена SIDMP в геноме томата или нокаутирование гена SIDMP включает уменьшение уровня экспрессии гена SIDMP в геноме томата или подвергание гена SIDMP в геноме томата делеционной мутации, или инсерционной мутации, или замене основания.

Способ уменьшения экспрессии гена DMP в геноме растения включает РНКи интерференцию. РНКи интерференция включает интерференцию одноцепочечной РНК, такой как миРНК (микроРНК), или интерференцию двухцепочечной РНК, такой как киРНК (короткая интерферирующая РНК), дцРНК (двухцепочечная РНК), кшРНК (короткая шпилечная РНК) и т.д.

Способ подвергания гена DMP в геноме растения делеционной мутации, или инсерционной мутации, или замене основания включает CRISPR/Cas9, или TELLEN, или вставку Т-ДНК, или EMS (этилметансульфонат) мутагенез.

Кроме того, способ подвергания гена DMP в геноме растения делеционной мутации, или инсерционной мутации, или замене основания включает CRISPR/Cas9. Данный способ подвергания гена DMP в геноме растения делеционной мутации, или инсерционной мутации, или замене основания включает следующие стадии: введение вектора CRISPR/Cas9, содержащего целевую последовательность, в растение с получением линии-индуктора гаплоидного растения.

В конкретном примере настоящего изобретения, когда двудольное растение представляет собой *Arabidopsis thaliana*, целевая последовательность CRISPR/Cas9 находится в положениях 98-117 SEQ ID NO:

1, положениях 290-309 SEQ ID NO: 3, положениях 368-387 SEQ ID NO: 3 и положениях 509-528 SEQ ID NO: 1. Вектор CRISPR/Cas9, содержащий целевую последовательность, представляет собой рекомбинантный вектор, полученный соединением молекулы ДНК, представленной в SEQ ID NO: 7 (экспрессионный элемент енРНК (единая направляющая РНК)), молекулы ДНК, представленной в SEQ ID NO: 9 (экспрессионный элемент Cas9) и молекулы ДНК, представленной в SEQ ID NO: 10 (экспрессионный элемент флуоресцентного белка), с вектором pICSL4723 методом "золотых ворот".

В конкретном примере настоящего изобретения, когда двудольное растение представляет собой томат, целевые последовательности CRISPR/Cas9 находятся в положениях 76-95 SEQ ID NO: 5 и положениях 247-266 SEQ ID NO: 5. Вектор CRISPR/Cas9, содержащий целевую последовательность, представляет собой рекомбинантный вектор, полученный соединением молекулы ДНК, представленной в SEQ ID NO: 8 (экспрессионный элемент енРНК), молекулы ДНК, представленной в SEQ ID NO: 11 (экспрессионный элемент Cas9), молекулы ДНК, представленной в SEQ ID NO: 10 (экспрессионный элемент флуоресцентного белка), и молекулы ДНК, представленной в SEQ ID NO: 12 (экспрессионный элемент NptII), с вектором pICSL4723 методом "золотых ворот".

Другая задача настоящего изобретения заключается в том, чтобы предложить способ получения линии-индуктора гаплоидного растения, включающего стадию самоопыления линии-индуктора гаплоидного растения по меньшей мере один раз, в частности один раз.

Способ получения линии-индуктора гаплоидного растения дополнительно включает стадию скрининга мутантов по гену DMP. Мутант по гену DMP представляет собой растение с мутацией гена DMP, и мутант по гену DMP представляет собой гетерозиготный мутант по гену DMP или гомозиготный мутант по гену DMP.

Согласно способу получения линии-индуктора гаплоидного растения, линия-индуктор гаплоидного растения конкретно включает мутантную линию T1-34 *Arabidopsis thaliana*, мутантную линию T1-6 *Arabidopsis thaliana*, мутантную линию T1-11 *Arabidopsis thaliana*, мутантную линию T1-19 *Arabidopsis thaliana*, мутантную линию T1-24 *Arabidopsis thaliana*, мутантную линию T1-25 *Arabidopsis thaliana*, мутантную линию T1-28 *Arabidopsis thaliana*, мутантную линию T1-32 *Arabidopsis thaliana*, мутантную линию T2-33 *Arabidopsis thaliana*, мутантную линию T2-38 *Arabidopsis thaliana*, мутантную линию *sltmp-1* томата или мутантную линию *sltmp-2* томата;

различие между геномной ДНК мутантной линии T1-34 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имеет место делеция фрагмента в одной хромосоме в положениях 115-512 SEQ ID NO: 1 и имеет место вставка основания (T) в другой хромосоме в положениях 114-115 SEQ ID NO: 1;

различие между геномной ДНК мутантной линии T1-6 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имеет место делеция фрагмента в одной хромосоме в положениях 115-512 SEQ ID NO: 1 и имеет место делеция фрагмента в другой хромосоме в положениях 113-114 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имеет место вставка фрагмента в обеих хромосомах в положениях 160-161 SEQ ID NO: 3, где нуклеотидная последовательность вставленного фрагмента показана как SEQ ID NO: 13;

различие между геномной ДНК мутантной линии T1-11 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имеет место вставка основания (T) в положениях 114-115 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имеет место вставка основания (T) в одной хромосоме в положениях 160-161 SEQ ID NO: 3 и имеет место вставка фрагмента в другой хромосоме в положениях 160-161 SEQ ID NO: 3, где нуклеотидная последовательность вставленного фрагмента показана как SEQ ID NO: 13;

различие между геномом мутантной линии T1-19 *Arabidopsis thaliana* и геномом *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имеет место делеция основания (T) в одной хромосоме в положении 114 SEQ ID NO: 1 и имеет место делеция фрагмента в другой хромосоме в положениях 115-511 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имеет место делеция фрагмента в одной хромосоме в положениях 161-560 SEQ ID NO: 3 и имеет место делеция фрагмента в другой хромосоме в положениях 161-564 SEQ ID NO: 3;

различие между геномной ДНК мутантной линии T1-24 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имеет место делеция фрагмента в одной хромосоме в положениях 115-512 SEQ ID NO: 1 и имеет место вставка основания (T) в другой хромосоме в положениях 114-115 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имеет место делеция фрагмента в одной хромосоме в положениях 161-560 SEQ ID NO: 3 и имеет место делеция фрагмента в другой хромосоме в положениях 159-160 SEQ ID NO: 3;

различие между геномной ДНК мутантной линии T1-25 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имеет место делеция основания (T) в одной хромосоме в положении 114 SEQ ID NO: 1 и имеет место вставка фрагмента CGT в другой хромосоме в положениях 114-115 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имеет место делеция фрагмента в одной хромосоме в положениях 161-162 SEQ ID NO: 3 и имеет место вставка основания (A) в другой хромосоме в положениях 160-161 SEQ ID NO: 3;

различие между геномной ДНК мутантной линии T1-28 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имеет место делеция фрагмента в одной хромосоме в положениях 115-512 SEQ ID NO: 1 и имеет место делеция основания (Т) в другой хромосоме в положении 114 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имеет место вставка основания (А) в обеих хромосомах в положениях 160-161 SEQ ID NO: 3;

различие между геномной ДНК мутантной линии T1-32 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имеет место делеция фрагмента в одной хромосоме в положениях 115-511 SEQ ID NO: 1 и имеет место вставка основания (Т) в другой хромосоме в положениях 114-115 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имеет место делеция основания (С) в обеих хромосомах в положении 161 SEQ ID NO: 3;

отличие геномной ДНК мутантной линии T2-33 *Arabidopsis thaliana* и *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP9, имеет место вставка основания (Т) в обеих хромосомах в положениях 160-161 SEQ ID NO: 3 и имеет место вставка фрагмента в обеих хромосомах в положениях 561-562 SEQ ID NO: 3, где нуклеотидная последовательность вставленного фрагмента показана как SEQ ID NO: 14;

различие между геномной ДНК мутантной линии T2-38 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имеет место делеция фрагмента в обеих хромосомах в положениях 115-127 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имеет место вставка основания (Т) и основания (G) в обеих хромосомах в положениях 160-161 SEQ ID NO: 3 и в положениях 562-563 SEQ ID NO: 3, соответственно;

различие между геномной ДНК мутантной линии томата *sldmp-1* и геномной ДНК томата дикого типа AC заключается только в том, что в гене, кодирующем белок SIDMP, имеет место вставка основания (С) в обеих хромосомах в положениях 92-93 SEQ ID NO: 5;

различие между геномной ДНК мутантной линии томата *sldmp-2* и геномной ДНК томата дикого типа AC заключается только в том, что в гене, кодирующем белок SIDMP, имеет место делеция фрагмента в обеих хромосомах в положениях 93-249 SEQ ID NO: 5.

Другая задача настоящего изобретения заключается в том, чтобы предложить способ получения гаплоидного растения.

Способ получения гаплоидного растения, предложенный в изобретении, включает следующие стадии: самоопыление линии-индуктора гаплоидного растения или его потомства, полученного данным способом, или скрещивание линии-индуктора гаплоидного растения или его потомства с другими растительными материалами в качестве родителя мужского пола с получением самоопыленного потомства или межвидового потомства, а именно гаплоидного растения; где данное растение представляет собой двудольное растение.

Кроме того, способ получения гаплоидного растения дополнительно включает следующие стадии: проведение идентификации флуоресцентного мечения, и/или идентификации признаков гаплоида, и/или идентификации пloidии листьев, и/или идентификации молекулярного маркера на самоопыленном потомстве или межвидовом потомстве одного растения и отбор одиночных растений потомства, идентифицированных как гаплоиды по меньшей мере одним способом, в качестве гаплоидных растений.

Кроме того, способ идентификации флуоресцентного мечения может быть осуществлен следующим образом: скрещивание линии-индуктора гаплоидного растения, несущей экспрессионный элемент флуоресцентного белка в качестве родителя мужского пола, с родителем женского пола с получением межвидового потомства и проведение оценки на предмет того, является ли семя, подлежащее анализу, гаплоидом или диплоидом посредством выявления наличия флуоресцентного сигнала у данного семени межвидового потомства; где, если семя, подлежащее анализу, не показывает флуоресценции или показывает слабую флуоресценцию, данное семя является гаплоидом или кандидатом; если семя, подлежащее анализу, показывает сильную флуоресценцию, данное семя является диплоидом или кандидатом. Кроме того, являются ли семена, подлежащие анализу, флуоресцентными или нет, выявляют посредством флуоресцентного излучения. Кроме того, родитель мужского пола несет экспрессионный элемент флуоресцентного белка TagRFP, управляемый промотором AtOLEO1, таким образом, что можно оценить, является ли он гаплоидом или диплоидом, в зависимости от того, имеет ли гибридное потомство красную флуоресценцию.

Способ идентификации признаков гаплоида может быть осуществлен следующим образом: если растение, подлежащее анализу, имеет признаки карлика, узкие листья, растущие вверх, компактный тип растений и мужскую стерильность, данное растение является гаплоидом или кандидатом; если растение, подлежащее анализу, имеет признаки высокого растения, широкие и рассредоточенные листья и нормальную фертильность, данное растение является диплоидом или кандидатом.

Способ идентификации пloidии листьев может быть осуществлен следующим образом: выделение клеточных ядер молодых листьев растения, подлежащего анализу, и взятие листьев диплоидного растения в качестве контроля; выявление сигнала посредством проточного цитометра: сначала выявление сигнала диплоидного ядра и принятие пика сигнала диплоидного ядра за 50 (поскольку генетические материалы в диплоидных клетках находятся в двойном количестве относительно гаплоидных клеток, пик

сигнала гаплоидного ядра появляется около 25). Если пик сигнала ядра растения, подлежащего анализу, оказывается около 25, данное растение является гаплоидом или кандидатом; и если пик сигнала растения, подлежащего анализу, оказывается около 50, и положение обогащения интенсивности сигнала растения, подлежащего анализу, является таким же, как и у диплоидного ядра, данное растение является диплоидом или кандидатом.

Идентификация молекулярного маркера может быть осуществлена следующим образом: проведение ПЦР-амплификации посредством применения полиморфного праймера между родителем мужского пола (линией-индуктором гаплоида - родителем женского пола) и родителем женского пола, и проведение оценки на предмет того, является ли растение, подлежащее анализу, гаплоидом или диплоидом согласно продукту ПЦР-амплификации; где, если данный продукт амплификации растения, подлежащего анализу, имеет только картину полос родителя женского пола без картины полос родителя мужского пола, данное растение является гаплоидом или кандидатом; если данный продукт амплификации растения, подлежащего анализу, имеет гетерозиготную картину полос родителя мужского пола и родителя женского пола, данное растение является диплоидом или кандидатом.

Линия-индуктор гаплоидного растения и гаплоидное растение, полученные данным способом, также входят в объем данного изобретения.

Данная линия-индуктор гаплоидного растения и данное гаплоидное растение включают не только клетки, ткани и органы, происходящие от линии-индуктора гаплоидного растения и гаплоидного растения, такие как семена, листья, плоды, стебли, цветки, корни, но также материалы для размножения, полученные от линии-индуктора гаплоидного растения и гаплоидного растения, такие как группа, состоящая из пыльцы, завязей, семязачатков, зародышей, эндоспермов, яйцеклеток, эмбрионов на стадии дробления, корней, кончиков корней, гипокотилей, семядолей, стеблей, листьев, цветков, пыльников, семян, меристематических клеток, протопластов и культур ткани.

Ген AtDMP8 в мутантной линии *Arabidopsis thaliana*, или ген AtDMP9 в мутантной линии *Arabidopsis thaliana*, или ген SIDMP в мутантной линии томата в приведенном выше способе также входят в объем настоящего изобретения.

Любое из следующих применений Г1) - Г5) также входит в объем настоящего изобретения:

Г1) применение линии-индуктора гаплоидного растения, полученной данным способом, в культивировании гаплоидных растений;

Г2) применение вещества для сайленсинга или ингибирования экспрессии и/или активности гена AtDMP8 и/или гена AtDMP9 в геноме растения или нокаутирования гена AtDMP8 и/или гена AtDMP9 или вещества для ингибирования активности белка AtDMP8 и/или белка AtDMP9 в культивировании линий-индукторов гаплоидного растения или гаплоидных растений;

Г3) применение вещества для сайленсинга или ингибирования экспрессии и/или активности гена SIDMP в геноме растения или нокаутирования гена SIDMP или вещества для ингибирования активности белка SIDMP в культивировании линий-индукторов гаплоидного растения или гаплоидных растений;

Г4) применение белка AtDMP8 или родственного ему биологического материала и/или белка AtDMP9 или родственного ему биологического материала в регулировании показателя индукции линий-индукторов гаплоидного растения, или в улучшении показателя индукции линий-индукторов гаплоидного растения, или в культивировании линий-индукторов гаплоидного растения или гаплоидных растений;

Г5) применение белка SIDMP или родственных ему биологических материалов в регулировании показателя индукции линий-индукторов гаплоидного растения, или в улучшении показателя индукции линий-индукторов гаплоидного растения, или в культивировании линий-индукторов гаплоидного растения или гаплоидных растений.

В любом из вышеупомянутых применений или способов белок AtDMP8 является белком, представленным в а1), или а2), или а3), или а4):

а1) белок с аминокислотной последовательностью, показанной в SEQ ID NO: 2;

а2) слитый белок, полученный путем присоединения метки к N-концу и/или C-концу белка, показанного в SEQ ID NO: 2;

а3) белок, имеющий такую же функцию, полученный путем замены и/или делеции и/или вставки одного или нескольких аминокислотных остатков в аминокислотной последовательности, показанной в SEQ ID NO: 2;

а4) белок, имеющий гомологию 75% или более с аминокислотной последовательностью, показанной в SEQ ID NO: 2, и имеющий такую же функцию;

белок AtDMP9 является белком, представленным в следующих из б1), или б2), или б3), или б4):

б1) белок с аминокислотной последовательностью, показанной в SEQ ID NO: 4;

б2) слитый белок, полученный путем присоединения метки к N-концу и/или C-концу белка, показанного в SEQ ID NO: 4;

б3) белок, имеющий такую же функцию, полученный путем замены и/или делеции и/или вставки одного или нескольких аминокислотных остатков в аминокислотной последовательности, показанной в SEQ ID NO: 4;

б4) белок, имеющий гомологию 75% или более с аминокислотной последовательностью, показан-

ной в SEQ ID NO: 4, и имеющий такую же функцию.

Белок SIDMP является белком, представленным в следующих из в1), или в2), или в3), или в4):

в1) белок с аминокислотной последовательностью, показанной в SEQ ID NO: 6;

в2) слитый белок, полученный путем присоединения метки к N-концу и/или C-концу белка, показанного в SEQ ID NO: 6;

в3) белок, имеющий такую же функцию, полученный путем замены и/или делеции и/или вставки одного или нескольких аминокислотных остатков в аминокислотной последовательности, показанной в SEQ ID NO: 6;

в4) белок, имеющий гомологию 75% или более с аминокислотной последовательностью, показанной в SEQ ID NO: 6, и имеющий такую же функцию.

Биоматериал, родственный белку AtDMP8, или AtDMP9, или SIDMP, представляет собой любое из следующего г1) - г12):

г1) молекула нуклеиновой кислоты, кодирующая белок AtDMP8, или AtDMP9, или SIDMP;

г2) экспрессионная кассета, содержащая молекулу нуклеиновой кислоты по г1);

г3) рекомбинантный вектор, содержащий молекулу нуклеиновой кислоты по г1);

г4) рекомбинантный вектор, содержащий экспрессионную кассету по г2);

г5) рекомбинантный микроорганизм, содержащий молекулу нуклеиновой кислоты по г1);

г6) рекомбинантный микроорганизм, содержащий экспрессионную кассету по г2);

г7) рекомбинантный микроорганизм, содержащий рекомбинантный вектор по г3);

г8) рекомбинантный микроорганизм, содержащий рекомбинантный вектор по г4);

г9) трансгенная линия растительных клеток, содержащая молекулу нуклеиновой кислоты по г1);

г10) трансгенная линия растительных клеток, содержащая экспрессионную кассету по г2);

г11) трансгенная линия растительных клеток, содержащая рекомбинантный вектор по г3);

г12) трансгенная линия растительных клеток, содержащая рекомбинантный вектор по г4).

Кроме того, молекула нуклеиновой кислоты по г1) представляет собой ген, показанный в следующих из 1), или 2), или 3), или 4):

1) молекула кДНК или молекула геномной ДНК с кодирующей последовательностью, показанной в SEQ ID NO: 1, или SEQ ID NO: 3, или SEQ ID NO: 5;

2) молекула кДНК или молекула геномной ДНК, имеющая идентичность 75% или более с нуклеотидной последовательностью, определенной в 1), и кодирующая белок AtDMP8, или AtDMP9, или SIDMP;

3) молекула кДНК или молекула геномной ДНК, происходящая из двудольного растения и имеющая идентичность 75% или более с нуклеотидной последовательностью, определенной в 1), и кодирующая белок AtDMP8, или AtDMP9, или SIDMP;

4) молекула кДНК или молекула геномной ДНК, гибридизирующаяся с нуклеотидной последовательностью, определенной в 1), или 2), или 3), в жестких условиях и кодирующая белок AtDMP8, или AtDMP9, или SIDMP;

где SEQ ID NO: 1 представляет собой молекулу нуклеиновой кислоты, кодирующую белок AtDMP8; SEQ ID NO: 3 представляет собой молекулу нуклеиновой кислоты, кодирующую белок AtDMP9; и SEQ ID NO: 5 представляет собой молекулу нуклеиновой кислоты, кодирующую белок SIDMP.

Кроме того, вещество, осуществляющее нокаутирование гена AtDMP8 и/или гена AtDMP9, или вещество, осуществляющее нокаутирование гена SIDMP, представляет собой вектор CRISPR/Cas9, содержащий целевую последовательность, как описано выше.

В любом из вышеупомянутых применений, или способов, или растений, где данные растения являются двудольными, в частности, *Arabidopsis thaliana* или томат; в частности, *Arabidopsis thaliana* может представлять собой *Arabidopsis thaliana* дикого типа (Col-0) или ms1; в частности, томат может представлять собой томат дикого типа AC или Micro-Tom или материал поколения F₁, полученный скрещиванием томата AC с Micro-Tom.

Краткое описание графических материалов

На фиг. 1 представлен график, на котором показано сравнение гаплоидов и диплоидов *Arabidopsis thaliana*. Растения, указанные стрелками, являются гаплоидами.

На фиг. 2 представлен график, на котором показано сравнение гаплоидов и диплоидов *Arabidopsis thaliana* после стрелкования.

На фиг. 3 представлен график, на котором показано сравнение результатов проточной цитометрии гаплоидов и диплоидов *Arabidopsis thaliana*.

На фиг. 4 показан график идентификации флуоресценции гаплоидов *Arabidopsis thaliana*. На панели (а) показан фенотип семян *Arabidopsis thaliana* под белым светом, на панели (б) показан фенотип семян *Arabidopsis thaliana* при флуоресценции, и семена, показанные стрелками, являются гаплоидами.

На фиг. 5 показано изображение геля подтверждения молекулярного маркера гаплоидов и диплоидов *Arabidopsis thaliana*. М означает молекулярный маркер 2К, I представляет картину полос материала родителя женского пола, II представляет картину полос материала родителя мужского пола, III представ-

ляет картину полос гаплоида, и IV представляет картину полос диплоида.

На фиг. 6 представлен график, на котором показано сравнение растений и плодов томата дикого типа и мутантов *sldmp*. На панелях (а) и (б) показаны растения томата дикого типа и мутанта *sldmp*, соответственно, и на панелях (в) и (г) показаны полученные в результате самоопыления плоды томата дикого типа и мутанта *sldmp*, соответственно.

На фиг. 7 показано изображение флуоресцентной экспрессии семян томата. На панелях (а) и (б) показаны характеристики непроросших и проросших семян томата, сфотографированных в светлом поле, и на панелях (в) и (г) показаны характеристики непроросших и проросших семян томата, сфотографированных при флуоресценции.

На фиг. 8 представлен график, на котором показано сравнение результатов проточной цитометрии гаплоидов и диплоидов томата.

На фиг. 9 показана схематичная диаграмма структуры главных элементов в рекомбинантном векторе. На панелях (а) и (б) показана схематичная диаграмма структуры вектора *Arabidopsis thaliana* и томата, соответственно.

Наилучший способ осуществления изобретения

Следующие примеры способствуют лучшему пониманию настоящего изобретения, но не ограничивают настоящее изобретение. Экспериментальные способы в следующих примерах являются традиционными, если не указано иное. Экспериментальные материалы, используемые в следующих примерах, приобретали у поставщиков стандартных биохимических реагентов, если не указано иное. В количественных экспериментах следующих примеров эксперименты проводят в трех повторностях, и результаты приводят к среднему значению.

Вектор pICSL4723 в следующих примерах описан в документе "Castel, B. and L. Tomlinson, et al. (2019). Optimization of T-DNA architecture for Cas9-mediated mutagenesis in *Arabidopsis thaliana*. PloS one, 14(1).", общедоступном от Китайского сельскохозяйственного университета, данный экспериментальный материал используют только для повторения релевантных экспериментов по настоящему изобретению и не может быть использован в других целях.

Arabidopsis thaliana дикого типа Col-0 и *ms1* в следующих примерах описан в литературе "Rosso, M. G. and Y. Li, et al. (2003). An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. Plant Mol Biol 53 (1-2): 247-59.", общедоступной от Китайского сельскохозяйственного университета, данный биологический материал использован только для повторения релевантных экспериментов по настоящему изобретению и не может быть использован в других целях.

Томат дикого типа AC в следующих примерах описан в литературе "Yuan, G. and C. Jia, et al. (2010). Effect of brassinosteroids on drought resistance and abscisic acid concentration in tomato under water stress. Scientia Horticulturae 126 (2): 103-108.", общедоступной от Китайского сельскохозяйственного университета, данный биологический материал использован только для повторения релевантных экспериментов по настоящему изобретению и не может быть использован в других целях.

Томат дикого типа Micro-Tom в следующих примерах описан в литературе "Sun, H. and S. Uchii, et al. (2006). A highly efficient transformation protocol for Micro-Tom, a model cultivar for tomato functional genomics. Plant and Cell Physiology 47 (3): 426-431.", общедоступной от Китайского сельскохозяйственного университета, данный биологический материал использован только для повторения релевантных экспериментов по настоящему изобретению и не может быть использован в других целях.

CDS (кодирующая) последовательность гена AtDMP8 в данном изобретении показана в положениях 95-826 SEQ ID NO: 1 в перечне последовательностей, и аминокислотная последовательность белка, кодируемого геном AtDMP8, показана в SEQ ID NO: 2.

CDS последовательность гена AtDMP9 в данном изобретении показана в положениях 141-875 SEQ ID NO: 3 в перечне последовательностей, и аминокислотная последовательность белка, кодируемого геном AtDMP9, показана в SEQ ID NO: 4.

CDS последовательность гена SIDMP в данном изобретении показана в положениях 1-678 SEQ ID NO: 5 в таблице последовательностей, и аминокислотная последовательность белка, кодируемого геном SIDMP, показана в SEQ ID NO: 6.

Пример 1. Получение нокаутированных по гену AtDMP8 и/или по гену AtDMP9 мутантов *Arabidopsis thaliana* и их применения.

1. Гены AtDMP8 и/или AtDMP9 нокаутировали с использованием системы CRISPR/Cas9.

Гены AtDMP8 и/или AtDMP9 в *Arabidopsis thaliana* нокаутировали посредством системы CRISPR/Cas9 с получением нокаутированных по гену AtDMP8 и/или по гену AtDMP9 мутантов *Arabidopsis thaliana*. Конкретные стадии были следующими.

1. Выбор последовательностей енРНК.

Последовательности целевого сайта разрабатывали на гене AtDMP8 и/или гене AtDMP9, соответственно, и их длина составляла 20 п.о. (пар оснований).

Целевой сайт 1 находился в положениях 98-117 SEQ ID NO: 1 и 144-163 SEQ ID NO: 3, и последовательность целевого сайта 1 енРНК представляла собой GAGAAAACAGAGGAAAGCGT (SEQ ID NO: 15).

Целевой сайт 2 находился в положениях 290-309 SEQ ID NO: 3 и последовательность целевого сайта 2 енРНК представляла собой AAGAGGTGCGAAAACGTCGCA (SEQ ID NO: 16).

Целевой сайт 3 находился в положениях 368-387 SEQ ID NO: 3 и последовательность целевого сайта 3 енРНК представляла собой TCAAGAGTGTTCCTGTCCGGA (SEQ ID NO: 17).

Целевой сайт 4 находился в положениях 509-528 SEQ ID NO: 1 и 558-577 SEQ ID NO: 3, и последовательность целевого сайта 4 енРНК представляла собой ATGACAACCGCGAGTCCACG (SEQ ID NO: 18).

2. Конструирование вектора CRISPR/Cas9.

Вектор CRISPR/Cas9 представляет собой рекомбинантный вектор (структурная диаграмма вектора показана на фиг. 9a), полученный соединением молекулы ДНК, показанной в SEQ ID NO: 7 (экспрессионный элемент енРНК), молекулы ДНК, показанной в SEQ ID NO: 9 (экспрессионный элемент Cas9), и молекулы ДНК, показанной в SEQ ID NO: 10 (экспрессионный элемент флуоресцентного белка) с вектором pICSL4723 с использованием метода "золотых ворот". Молекула ДНК, показанная в SEQ ID NO: 7, последовательно содержит кодирующую последовательность енРНК, нацеленной на целевой сайт 1, кодирующую последовательность енРНК, нацеленной на целевой сайт 2, кодирующую последовательность енРНК, нацеленной на целевой сайт 3, и кодирующую последовательность енРНК, нацеленной на целевой сайт 4, с промотором AtU6-26, используемым для инициации экспрессии кодирующей последовательности енРНК, находящейся около начала каждой кодирующей последовательности енРНК.

3. Получение трансгенных растений.

Вектор CRISPR/Cas9, полученный на стадии 2, трансформировали в компетентные клетки агробактерии GV3101 посредством теплового шока (компетентные клетки агробактерии GV3101 приобретали у Beijing Aosen Dingxin Biotechnology Co., Ltd., и они являются общедоступными) с получением рекомбинантного штамма GV3101/CRISPR/Cas9.

Затем цветonos Arabidopsis thaliana дикого типа Col-0 инфицировали рекомбинантным штаммом GV3101/CRISPR/Cas9 посредством применения способа трансформации инфицированием цветonos Arabidopsis thaliana агробактерией (т.е. рекомбинантную агробактерию подвергали расширенному размножению при 28°C, и цветonos Arabidopsis thaliana инфицировали раствором размноженной агробактерии), и поколение T1 трансгенного растения Arabidopsis thaliana получили после скрининга красной флуоресценции.

4. Идентификация трансгенных растений с мутацией гена AtDMP8 и/или гена AtDMP9.

Листья трансгенного растения Arabidopsis thaliana поколения T1, полученного на стадии 3, отбирали, выделяли геномную ДНК в качестве матрицы, и проводили ПЦР-амплификацию посредством применения следующих двух пар праймеров, соответственно, с получением продуктов ПЦР-амплификации разных линий.

Последовательности праймеров для выявления мутантной последовательности гена AtDMP8 были следующими:

DMP8F1: TGCGAAATGAGATTGGTTTGGG (SEQ ID NO: 19);

DMP8R1: AAACACCCTGTGACTCTCCG (SEQ ID NO: 20).

Последовательности праймеров для выявления мутантной последовательности гена AtDMP9 были следующими:

DMP9F1: ATAACCGTCAATAACCGCCG (SEQ ID NO: 21);

DMP9R2: CCAGTCATGCAACCAACACC (SEQ ID NO: 22).

Продукты ПЦР-амплификации разных линий подвергали секвенированию по Сэнджеру, и результаты секвенирования сравнивали с AtDMP8 и AtDMP9 Arabidopsis thaliana дикого типа Col-0, соответственно. Генотипы AtDMP8 и AtDMP9, соответственно, идентифицировали согласно следующим принципам.

Если последовательность из последовательности целевого сайта имела характеристику дублета, генотип данной линии был гетерозиготным (т.е. ген AtDMP8 и/или ген AtDMP9 мутировал на одной из двух гомологичных хромосом и не мутировал на другой хромосоме), и данная линия представляла собой гетерозиготную мутантную линию трансгенного Arabidopsis thaliana поколения T1;

Для последовательности с характеристиками специфичного синглета из последовательности целевого сайта, если она была такой же, как последовательность гена AtDMP8 и AtDMP9 Arabidopsis thaliana дикого типа Col-0, генотип данной линии был генотипом дикого типа, то есть, не было мутации на последовательности гена AtDMP8 и AtDMP9; если она отличалась от последовательности гена AtDMP8 и/или AtDMP9 Arabidopsis thaliana дикого типа Col-0, генотип данной линии был гомозиготным (ген AtDMP8 и/или ген AtDMP9 мутировал на обеих гомологичных хромосомах), данная линия представляла собой гомозиготную мутантную линию трансгенного Arabidopsis thaliana поколения T1.

Результаты идентификации показаны в табл. 1 и 2 (в табл. 1 и 2 описаны мутации генов AtDMP8 и AtDMP9 трансгенного Arabidopsis thaliana поколения T1, соответственно): среди 41 трансгенного растения Arabidopsis thaliana поколения T1 имелось 33 трансгенных растения поколения T1 с мутацией гена AtDMP8, из которых 2 растения имели гомозиготную мутацию в гене AtDMP8, и 17 растений имели би-

аллельную мутацию в гене AtDMP8. Имелось 28 трансгенных растений поколения T1 с мутацией гена AtDMP9, из которых 5 растений имели гомозиготную мутацию в гене AtDMP9, и 7 растений имели биаллельную мутацию в гене AtDMP9. Имелось 12 растений с гомозиготными/биаллельными мутациями в AtDMP8 и AtDMP9. Индивиды, которые возникали в результате мутации со сдвигом рамки считывания (делеция, не кратная 3 основаниям), среди гомозиготных/биаллельных мутантных индивидов отбирали далее на идентификацию фенотипа. Имелось три типа гомозиготной мутации/биаллельной мутации AtDMP8, гомозиготной мутации/биаллельной мутации AtDMP9 и гомозиготной мутации/биаллельной мутации AtDMP8 и AtDMP9.

Таблица 1
Типы мутаций гена AtDMP8 в трансгенном *Arabidopsis thaliana* поколения T1

| Серийный номер | <i>AtDMP8</i> | аллель 1 | аллель 2 |
|----------------|------------------------|---|---|
| T1-1 | Биаллельная мутация | GAAAG----- | GGGGCGTCGGAATCA(SEQ ID NO:23)..GACTCCACGT(SEQ ID NO:24) |
| T1-2 | Гетерозиготная мутация | -----CTCGC | WT (дикий тип) |
| T1-3 | Биаллельная мутация | GAAAGACGTCGGAATCA(SEQ ID NO:25)..GACTCCACGT(SEQ ID NO:26) | GGGTCGTCGGAATCA(SEQ ID NO:27)..GACTCCACGT(SEQ ID NO:28) |
| T1-4 | Дикий тип | WT | WT |
| T1-5 | Дикий тип | WT | WT |
| T1-6 | Биаллельная мутация | GAAAG----- | GGAA—GTCGGAATCA(SEQ ID NO:29)..GACTCCACGT(SEQ ID NO:30) |
| T1-7 | Гетерозиготная мутация | GAAAGACGTCGGAATCA(SEQ ID NO:31)..GACTCCACGT(SEQ ID NO:32) | WT |
| T1-8 | Гетерозиготная | GAAAGGCGTCGGAATCA(SEQ | WT |

| | | | |
|-------|------------------------|---|---|
| | мутация | ID NO:33)..GACTCCACGT(SEQ ID NO:34) | |
| T1-9 | Биаллельная мутация | GAAAGTCGTCGGAATCA(SEQ ID NO:35)..GACTCCACGT(SEQ ID NO:36) | GGGTCG--GGAATCA..GACTCCACGT (SEQ ID NO:37) |
| T1-10 | Биаллельная мутация | GAAAG----- | GGGGCGTCGGAATCA(SEQ ID NO:38)..GACTCCACGT(SEQ ID NO:39) |
| T1-11 | Гомозиготная мутация | GAAAGGCGTCGGAATCA(SEQ ID NO:40)..GACTCCACGT(SEQ ID NO:41) | GGGGCGTCGGAATCA(SEQ ID NO:42)..GACTCCACGT(SEQ ID NO:43) |
| T1-12 | Биаллельная мутация | GAAAGGCGTCGGAATCA(SEQ ID NO:44)..GACTCCACGT(SEQ ID NO:45) | GGG-GTCGGAATCA(SEQ ID NO:46)..GACTCCACGT(SEQ ID NO:47) |
| T1-13 | Дикий тип | WT | WT |
| T1-14 | Гетерозиготная мутация | GAAAGGCGTCGGAATCA(SEQ ID NO:48)..GACTCCACGT(SEQ ID NO:49) | WT |
| T1-15 | Биаллельная мутация | GAAAG----- | GGGGCGTCGGAATCA(SEQ ID NO:50)..GACTCCACGT(SEQ ID NO:51) |
| T1-16 | Гетерозиготная мутация | GAA-----(+49 п.н.)---TTCATGAA | WT |
| T1-17 | Дикий тип | WT | WT |
| T1-18 | Гетерозиготная мутация | GAAA-CGTCGGAATCA(SEQ ID NO:52)..GACTCCACGT(SEQ ID NO:53) | WT |
| T1-19 | Биаллельная мутация | GAAA-CGTCGGAATCA(SEQ ID NO:54).. CTCCACGTG-ACTC | GGG----- |
| T1-20 | Гетерозиготная мутация | GAAAG--TCGGAATCA..GACTCCACGT (SEQ ID NO:55) | WT |
| T1-21 | Гетерозиготная мутация | GAAAG-----GACTCGC | WT |
| T1-22 | Дикий тип | WT | WT |
| T1-23 | Биаллельная мутация | GAAAG----- | GGGTCGTCGGAATCA(SEQ ID NO:56) |
| T1-24 | Биаллельная | GAAAG-----GACTCGC | GGGTCGTCGGAATCA(SEQ |

| | | | |
|-------|------------------------|--|---|
| | мутация | | ID NO:57) |
| T1-25 | Биаллельная мутация | GAAA-CGTCGGAATCA(SEQ ID NO:58)..GACTCCACGT(SEQ ID NO:59) | GGG(+3 п.н.)CGTCGGAATCA(SEQ ID NO:60) |
| T1-26 | Дикий тип | WT | WT |
| T1-27 | Биаллельная мутация | GAAAG----- | GG-CGTCGGAATCA (SEQ ID NO:61) |
| T1-28 | Биаллельная мутация | GAAAG-----GACTCGC | GG-CGTCGGAATCA(SEQ ID NO:62) |
| T1-29 | Дикий тип | WT | WT |
| T1-30 | Биаллельная мутация | GAAAGGCGTCGGAATCA(SEQ ID NO:63)..GACTCCACGT(SEQ ID NO:64) | GGG-----ACTCGC |
| T1-31 | Гетерозиготная мутация | GAAAGACGTTCGGAATCA(SEQ ID NO:65)..GACTCCACGT(SEQ ID NO:66) | WT |
| T1-32 | Биаллельная мутация | GAAAG----- | GGGTCGTCGGAATCA(SEQ ID NO:67) |
| T1-33 | Гетерозиготная мутация | GAAAG-----(+202 п.н.)---- | WT |
| T1-34 | Биаллельная мутация | GAAAG-----GACTCGC | GGGTCGTCGGAATCA(SEQ ID NO:68) |
| T1-35 | Биаллельная мутация | GAAAGACGTTCGGAATCA(SEQ ID NO:69)..GACTCCACGT(SEQ ID NO:70) | GGGGCGTCGGAATCA(SEQ ID NO:71)..GACTCCACGT(SEQ ID NO:72) |
| T1-36 | Гетерозиготная мутация | GAAAG-----GACTCGC | WT |
| T1-37 | Гетерозиготная мутация | GAAAGTCGTCGGAATCA(SEQ ID NO:73)—GACTCCACGT(SEQ ID NO:74) | WT |
| T1-38 | Гомозиготная мутация | GAAAG-----GTTTACA—CCACGTGGACT(SEQ ID NO:75) | GGG-----GTTTACA..CCACGTGG |
| T1-39 | Дикий тип | WT | WT |
| T1-40 | Гетерозиготная мутация | GAAAG----- | WT |
| T1-41 | Гетерозиготная мутация | GAAAGGCGTCGGAATCA(SEQ ID NO:76)..GACTCCACGT(SEQ ID NO:77) | WT |

Примечание: "-" означает наличие делеций оснований, "." означает опущенные основания.

Таблица 2

Типы мутаций гена *AtDMP9* в трансгенном *Arabidopsis thaliana* поколения T1

| Серийный номер | <i>AtDMP9</i> | аллель 1 | аллель 2 |
|----------------|------------------------|--------------------------------------|--------------------------------------|
| T1-1 | Биаллельная мутация | GGAAAGACGTCGG(SEQ ID NO:78)..GACGCCA | GGAAAG-GTCGG..GACGCCA |
| T1-2 | Дикий тип | WT | WT |
| T1-3 | Гомозиготная мутация | GGAAA-CGTCGG..GACGCCA | GGAAA-CGTCGG..GACGCCA |
| T1-4 | Гетерозиготная мутация | GGAA-GCGTCGG..GACGCCA | WT |
| T1-5 | Дикий тип | WT | WT |
| T1-6 | Гомозиготная мутация | GGAAAG(+16 п.н.)CGTCGGA..GACGCCA | GGAAAG(+16 п.н.)CGTCGG..GACGCCA |
| T1-7 | Гетерозиготная мутация | GGAAAGACGTCGG(SEQ ID NO:79)..GACGCCA | WT |
| T1-8 | Дикий тип | WT | WT |
| T1-9 | Дикий тип | WT | WT |
| T1-10 | Биаллельная мутация | GGAAAGCG-----CAATGTC..GCCA | GGAAAGACGTCGG(SEQ ID NO:80)..GACGCCA |
| T1-11 | Биаллельная мутация | GGAAAGGCGTCGG(SEQ ID NO:81)..GACGCCA | GGAAAG(+16 п.н.)CGTCGG..GACGCCA |
| T1-12 | Гетерозиготная мутация | GGAAAGGCGTCGG(SEQ ID NO:82)..GACGCCA | WT |
| T1-13 | Дикий тип | WT | WT |
| T1-14 | Гетерозиготная мутация | GGAAAGACGTCGG(SEQ ID NO:83)..GACGCCA | WT |
| T1-15 | Гетерозиготная мутация | GGAAAGGCGTCGG(SEQ ID NO:84)..GACGCCA | WT |
| T1-16 | Гетерозиготная мутация | GGAAAG-GTCGG..GACGCCA | WT |
| T1-17 | Дикий тип | WT | WT |
| T1-18 | Дикий тип | WT | WT |
| T1-19 | Биаллельная мутация | GGAAAG-----GGA | GGAAAG-----TCGCGGTGTTC(SEQ ID NO:85) |
| T1-20 | Гетерозиготная мутация | GGA---CGTCGG..GACGCCA | WT |
| T1-21 | Гетерозиготная мутация | GGAAAGACGTCGG(SEQ ID NO:86)..GACGCCA | WT |

| | | | |
|-------|------------------------|--|--|
| T1-22 | Гетерозиготная мутация | GGAAAG-----GGA | WT |
| T1-23 | Гетерозиготная мутация | GGAAAGACGTCGG(SEQ ID NO:87)..GACGCCA | WT |
| T1-24 | Биаллельная мутация | GGAAAG-----GGA | GGAA--CGTCGG..GACGCCA |
| T1-25 | Биаллельная мутация | GGAAAG--TCGG..GACGCCA | GGAAAGACGTCGG(SEQ ID NO:88)..GACGCCA |
| T1-26 | Гетерозиготная мутация | GGAAAG-GTCGG..GACGCCA | WT |
| T1-27 | Гетерозиготная мутация | GG----CGTCGG..GACGCCA | WT |
| T1-28 | Гомозиготная мутация | GGAAAGACGTCGG(SEQ ID NO:89)..GACGCCA | GGAAAGACGTCGG(SEQ ID NO:90)..GACGCCA |
| T1-29 | Дикий тип | WT | WT |
| T1-30 | Биаллельная мутация | GGAAAG-----GGA | GGAAA-CGTCGG..GACGCCA |
| T1-31 | Гетерозиготная мутация | GGAAA-CGTCGG..GACGCCA | WT |
| T1-32 | Гомозиготная мутация | GGAAAG-GTCGG..GACGCCA | GGAAAG-GTCGG..GACGCCA |
| T1-33 | Гетерозиготная мутация | GGAAAGACG..CCACGTG(+401 п.н.)GA | WT |
| T1-34 | Дикий тип | WT | WT |
| T1-35 | Гетерозиготная мутация | GGAA--CGTCGG..GACGCCA | WT |
| T1-36 | Гетерозиготная мутация | GGAAAGCGTCGG(SEQ ID NO:91)..GACGCCAA | WT |
| T1-37 | Дикий тип | WT | WT |
| T1-38 | Гомозиготная мутация | GGAAAGTCGTCGG(SEQ ID NO:92)..GACGCCACGTGGGA (SEQ ID NO:93) | GGAAAGTCGTCGG(SEQ ID NO:94)..GACGCCACGTGGGA (SEQ ID NO:95) |
| T1-39 | Дикий тип | WT | WT |
| T1-40 | Дикий тип | WT | WT |
| T1-41 | Дикий тип | WT | WT |

Примечание: "-" означает наличие делеций оснований, "." означает опущенные основания.

Полученные мутантные линии по гену AtDMP8 трансгенного *Arabidopsis thaliana* поколения T1 включали T1-34, и конкретные мутации были следующими:

идентификация секвенированием показала, что: различие между геномной ДНК мутантных линий T1-34 по гену AtDMP8 трансгенного *Arabidopsis thaliana* поколения T1 и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имела место делеция фрагмента в одной хромосоме в положениях 115-512 SEQ ID NO: 1 и имела место вставка основания (T) в другой хромосоме в положениях 114-115 SEQ ID NO: 1.

Полученные линии двойных мутантов по генам AtDMP8 и AtDMP9 трансгенного *Arabidopsis thaliana* поколения T1 включали T1-6, T1-11, T1-19, T1-24, T1-25, T1-28 и T1-32, и конкретные мутации были следующими:

идентификация секвенированием показала, что: различие между геномной ДНК двойных мутантных линий T1-6 по генам AtDMP8 и AtDMP9 трансгенного *Arabidopsis thaliana* поколения T1 и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имела место делеция фрагмента в одной хромосоме в положениях 115-512 SEQ ID NO: 1, и имела место делеция фрагмента в другой хромосоме в положениях 113-114 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имела место вставка фрагмента в положениях 160-161 SEQ ID NO: 3, где нуклеотидная последовательность вставленного фрагмента показана как GTTTACACGGCGACTC (SEQ ID NO: 13).

Идентификация секвенированием показала, что: различие между геномной ДНК двойных мутантных линий T1-11 по генам AtDMP8 и AtDMP9 трансгенного *Arabidopsis thaliana* поколения T1 и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имела место вставка основания (T) в положениях 114-115 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имела место вставка основания (T) в одной хромосоме в положениях 160-161 SEQ ID NO: 3 и имела место вставка фрагмента в другой хромосоме в положениях 160-161 SEQ ID NO: 3, где нуклеотидная последовательность вставленного фрагмента показана как GTTTACACGGCGACTC (SEQ ID NO: 13).

Идентификация секвенированием показала, что: различие между геномной ДНК двойных мутант-

ных линий T1-19 по генам AtDMP8 и AtDMP9 трансгенного *Arabidopsis thaliana* поколения T1 и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имела место делеция основания (T) в одной хромосоме в положении 114 SEQ ID NO: 1 и имела место делеция фрагмента в другой хромосоме в положениях 115-511 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имела место делеция фрагмента в одной хромосоме в положениях 161-560 SEQ ID NO: 3 и имела место делеция фрагмента в другой хромосоме в положениях 161-564 SEQ ID NO: 3.

Идентификация секвенированием показала, что: различие между геномной ДНК двойных мутантных линий T1-24 по генам AtDMP8 и AtDMP9 трансгенного *Arabidopsis thaliana* поколения T1 и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имела место делеция фрагмента в одной хромосоме в положениях 115-512 SEQ ID NO: 1 и имела место вставка основания (T) в другой хромосоме в положениях 114-115 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имела место делеция фрагмента в одной хромосоме в положениях 161-560 SEQ ID NO: 3 и имела место делеция фрагмента в другой хромосоме в положениях 159-160 SEQ ID NO: 3.

Идентификация секвенированием показала, что: различие между геномной ДНК двойных мутантных линий T1-25 по генам AtDMP8 и AtDMP9 трансгенного *Arabidopsis thaliana* поколения T1 и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имела место делеция основания (T) в одной хромосоме в положении 114 SEQ ID NO: 1 и имела место вставка фрагмента CGT в другой хромосоме в положениях 114-115 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имела место делеция фрагмента в одной хромосоме в положениях 161-162 SEQ ID NO: 3 и имела место вставка основания (A) в другой хромосоме в положениях 160-161 SEQ ID NO: 3.

Идентификация секвенированием показала, что: различие между геномной ДНК двойных мутантных линий T1-28 по генам AtDMP8 и AtDMP9 трансгенного *Arabidopsis thaliana* поколения T1 и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имела место делеция фрагмента в одной хромосоме в положениях 115-512 SEQ ID NO: 1 и имела место делеция основания (T) в другой хромосоме в положении 114 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имела место вставка основания (A) в положениях 160-161 SEQ ID NO: 3.

Идентификация секвенированием показала, что: различие между геномной ДНК двойных мутантных линий T1-32 по генам AtDMP8 и AtDMP9 трансгенного *Arabidopsis thaliana* поколения T1 и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имела место делеция фрагмента в одной хромосоме в положениях 115-511 SEQ ID NO: 1 и имела место вставка основания (T) в другой хромосоме в положениях 114-115 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имела место делеция основания (C) в положении 161 SEQ ID NO: 3.

5. Идентификация трансгенного *Arabidopsis thaliana* поколения T2 посредством генотипирования

Мутантные линии T1-19, T1-33 и T1-38 по гену AtDMP8 и/или гену AtDMP9 трансгенного *Arabidopsis thaliana* поколения T1, полученные на стадии 4, подвергали самоопылению, после сбора семени высевали с получением трансгенного *Arabidopsis thaliana* поколения T2. Конкретный способ идентификации генотипов генов AtDMP8 и AtDMP9 трансгенного *Arabidopsis thaliana* поколения T2 был следующим: с геномной ДНК трансгенного *Arabidopsis thaliana* T2 в качестве матрицы идентифицировали генотипы генов AtDMP8 и AtDMP9 в трансгенном *Arabidopsis thaliana* поколения T2 посредством использования праймеров для определения мутантной последовательности AtDMP8 (DMP8F1 и DMP8R1) и AtDMP9 (DMP9F1 и DMP9R2) согласно способу на стадии 4, соответственно.

Полученные мутантные гомозиготные линии по гену AtDMP8 трансгенного *Arabidopsis thaliana* поколения T2 включали T2-33-1, T2-33-2 и T2-33-3, которые имеют такие же мутантные последовательности, и конкретные мутации были следующими:

идентификация секвенированием показала, что: различие между геномной ДНК гомозиготных мутантных линий T2-33-1, T2-33-2 и T2-33-3 по гену AtDMP8 трансгенного *Arabidopsis thaliana* поколения T2 и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, фрагмент в положениях 115-511 SEQ ID NO: 1 был заменен фрагментом размера 202 п.о. В частности, нуклеотидная последовательность фрагмента размером 202 п.о. была следующей:

```
GAAATTGACGAGCATTGATGTCTTCGAAACCGTTTTTGAACSTTTTCGCCACCA
TGCGACGTTTTCTACSTTTTCTCCTCCCGCGGCGGCTCTGCCGGAAGCATAGGC
AGTGAAGAGAGAGGGACAGGTTTGGGCGACCGAGACGATGTTGGTGACGGATTT
TGCGTCGTTGTCGTCGTGTAACSTCTGATTCGACG (SEQ ID NO: 96).
```

Полученные гомозиготные линии по гену AtDMP9 трансгенного *Arabidopsis thaliana* поколения T2 включали T2-33-4, T2-33-5 и T2-33-6, которые имеют такие же мутантные последовательности, и конкретные мутации были следующими:

идентификация секвенированием показала, что: различие между геномной ДНК гомозиготных мутантных линий T2-33-1, T2-33-2 и T2-33-3 по гену AtDMP9 трансгенного *Arabidopsis thaliana* поколения T2 и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP9, имела место вставка основания (T) в положениях 160-161 SEQ ID NO: 3 и имела

место вставка фрагмента в положениях 561-562 SEQ ID NO: 3, где данная нуклеотидная последовательность показана как:

```
CGTCGGAATCAGAGTTTACACGGCGACTCCGCCGCAAAAACCATCACCATCACCA
CCTTCTCGTTCACCAAAAACCGTCTTAATCTTTCATTGCCTTCCCTCCCGTCAGG
AGCCGCCGCTGGAGGAGGAAGAGGTCGAAAACGTCGCATGGTGGCGCAAGGAGT
TCAAAAAACGGTTTCGAAGACATCAATGCTCGTCAACTCCTTCCGACAGGAACA
CTCTTGATGTTTCGAAATGGTTCTTCCATCAATATACCGTGACGGAGACTGTAACG
GAATCAACACACTCATGATTATCTCCTTCTTGTCTTTGCGCAATGTCTTGTCTT
TCTTCCATTTTACCGACAGTTTCAAAGCATCCGATGGGAAGATCTACTACGGTTTC
GTGACGCCACGTG (SEQ ID NO: 14).
```

Полученные гомозиготные мутантные линии по генам AtDMP8 и AtDMP9 трансгенного *Arabidopsis thaliana* поколения T2 включали T2-19, T2-38-1 и T2-38-2, причем типы мутаций каждой линии являются следующими:

идентификация секвенированием показала, что: различие между геномной ДНК гомозиготных мутантных линий T2-19 по генам AtDMP8 и AtDMP9 трансгенного *Arabidopsis thaliana* поколения T2 и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имела место делеция фрагмента в положениях 115-511 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имела место делеция фрагмента в положениях 161-564 SEQ ID NO: 3.

Идентификация секвенированием показала, что: различие между геномной ДНК гомозиготных мутантных линий T2-38-1 и T2-38-2 по генам AtDMP8 и AtDMP9 трансгенного *Arabidopsis thaliana* поколения T2 и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, кодирующем белок AtDMP8, имела место делеция фрагмента в положениях 115-127 SEQ ID NO: 1; и в гене, кодирующем белок AtDMP9, имела место вставка фрагмента в положениях 160-161 SEQ ID NO: 3 и в положениях 562-563 SEQ ID NO: 3.

Описанные выше мутантные линии трансгенного *Arabidopsis thaliana* поколения T1 и T2 были отобраны для следующих экспериментов по анализу способности к индукции гаплоида.

II. Применение нокаутированного мутанта *Arabidopsis thaliana* по гену AtDMP8 и/или гену AtDMP9 в индукции получения гаплоида.

(A) Идентификация способности самоопыления нокаутированных мутантов *Arabidopsis thaliana* по AtDMP8 и AtDMP9 индуцировать гаплоид.

Три типа мутантов, полученных по генам AtDMP8 и AtDMP9, соответственно, самоопыляли с получением самоопыленного потомства, и идентификацию гаплоидов проводили на самоопыленном потомстве следующим способом (поскольку *Arabidopsis thaliana* дикого типа Col представлял собой гомозиготную полученную самоопылением линию, мутантное самоопыленное потомство, полученное нокаутированием генов AtDMP8 и AtDMP9 на этом фоне, не может идентифицировать гаплоидов посредством молекулярных маркеров):

1. Идентификация фенотипа растения.

После посадки семян, образовавшихся в результате самоопыления, наблюдали фенотипы одиночных растений, причем при гаплоидности наблюдались карликовость, узкие листья, растущие вверх, компактный тип растения и мужская стерильность, тогда как при диплоидности наблюдались высокие растения, широкие и рассредоточенные листья и нормальная фертильность (фиг. 1, 2).

2. Идентификация на листьях посредством проточной цитометрии.

Проточную цитометрию проводили на растении с характеристиками гаплоидности, полученном на стадии 1, где конкретный способ был следующим: ядра клеток молодых листьев растения, подлежащего анализу, выделяли, и диплоидные листья *Arabidopsis thaliana* использовали в качестве контроля; сигнал затем выявляли проточным цитометром, сперва выявляли сигнал диплоидных ядер, и пиковое положение сигнала диплоидного ядра принимали за 50 (поскольку генетические материалы в диплоидных клетках находятся в двойном количестве относительно гаплоидных клеток, пиковое положение сигнала гаплоидного ядра оказывается около 25). И если пиковое положение сигнала ядра растения, подлежащего анализу, оказывается около 25, данное растение, подлежащее анализу, считается гаплоидным растением. Если пик сигнала растения, подлежащего анализу, оказывается около 50, считали, что положение обогащения интенсивности сигнала растения, подлежащего анализу, является таким же, как и у диплоидного ядра, и растение, подлежащее анализу, было диплоидом (фиг. 3).

Результаты идентификации подсчитывали, и показатель индукции рассчитывали согласно следующей формуле: показатель индукции (%) = (число гаплоидных растений/общее число растений) × 100. Как видно, после одновременной мутации генов AtDMP8 и AtDMP9 в самоопыленном потомстве могут быть получены гаплоиды.

Таблица 3

Статистика по показателю индукции гаплоидов у самоопыленного потомства мутантов *dmp*

| Генотип | Номер растений | Общее число растений | Число гаплоидов | Показатель индукции гаплоидов (%) |
|-----------------|----------------|----------------------|-----------------|-----------------------------------|
| WT | Col-0 | 523 | 0 | 0 |
| <i>dmp8</i> | T2-33-1 | 270 | 0 | 0 |
| <i>dmp9</i> | T2-33-4 | 183 | 0 | 0 |
| <i>dmp8dmp9</i> | T2-38-1 | 165 | 6 | 3,64 |

А. Идентификация способности к индукции гибридизации мутантов *Arabidopsis thaliana*, нокаутированных по *AtDMP8* и *AtDMP9*.

Три типа мутантов, полученных по генам *AtDMP8* и *AtDMP9*, скрещивали с материалами *Arabidopsis thaliana* *ms1* с получением межвидового потомства, и гаплоиды в межвидовом потомстве идентифицировали следующими способами.

1. Идентификация флуоресцентного мечения.

Вектор CRISPR/Cas9 несет экспрессионный элемент TagRFP (*Entasmea quadricolor*) управляемый промотором *AtOLEO1*. Поскольку промотор *AtOLEO1* специфично экспрессировался в зародышах зрелых семян, флуоресцентный сигнал TagRFP можно было наблюдать посредством флуоресцентного излучения. Следовательно, мутант, несущий данный экспрессионный элемент, в качестве родителя мужского пола гибридизовали с другими нефлуоресцентными материалами родителя женского пола с получением семян, где зародыши диплоидных семян демонстрировали сильную красную флуоресценцию из-за наличия генома родителя мужского пола, тогда как зародыши гаплоидных семян не демонстрировали флуоресценции или демонстрировали слабую флуоресценцию из-за происхождения от родителя женского пола (фиг. 4).

2. Идентификация молекулярного маркера.

Семена без флуоресценции и со слабой флуоресценцией, идентифицированные на стадии 1, высаживали далее, выделяли геномную ДНК, и проводили ПЦР-амплификацию с использованием полиморфного праймера мутанта *Arabidopsis thaliana*, нокаутированного по *AtDMP8* и *AtDMP9*, 092B02-F(092B02-F: CAGCTGAGATGAACGAGTTGTCTT) (SEQ ID NO: 97), 092B02-R(092B02-R: TCTTTTGAGTCAC-TCCGTATGTCC) (SEQ ID NO: 98) и LB-08474(LB-08474: ATAATAACGCTGCGGACATCTACATTTT) (SEQ ID NO: 99), и амплифицированный продукт подвергали выявлению картины полос в агарозе, если размер амплифицированного продукта индивидуального растения, подлежащего анализу, составлял 500 п.о., демонстрируя 1 полосу, считали, что полоса индивидуального растения является картиной полос *Arabidopsis thaliana* *ms1* без картины полос материала родителя мужского пола, данное индивидуальное растение представляло собой родителя гаплоидного женского пола. Если размер амплифицированных продуктов индивидуального растения, подлежащего анализу, составлял 500 п.о. и 1094 п.о., демонстрируя 2 полосы, считали, что полоса индивидуального растения является гетерозиготной картиной полос *Arabidopsis thaliana* *ms1* и трансгенной мутантной линии *Arabidopsis thaliana*, данное индивидуальное растение представляло собой потомство нормального гибрида и было диплоидом (фиг. 5).

3. Идентификация фенотипа зрелых растений.

Фенотипы растений, идентифицированные на стадиях 1 и 2, дополнительно осматривали, причем гаплоиды демонстрировали карликовость, узкие листья, растущие вверх, компактный тип растения и мужскую стерильность, тогда как диплоиды были высокими растениями с широкими и рассредоточенными листьями и нормальной фертильностью.

4. Идентификация на листьях посредством проточной цитометрии.

Проточную цитометрию проводили на растениях с характеристиками гаплоида, полученном на стадии 3, где конкретный способ был следующим: выделяли ядра клеток молодых листьев растения, подлежащего анализу, и листья диплоидного *Arabidopsis thaliana* использовали в качестве контроля; сигнал затем выявляли проточным цитометром, сначала выявляли сигнал диплоидных ядер, и пиковое положение сигнала диплоидного ядра принимали за 50 (поскольку генетические материалы в диплоидных клетках находились в двойном количестве относительно гаплоидных клеток, пиковое положение сигнала гаплоидного ядра оказывается около 25). И если пиковое положение сигнала ядра растения, подлежащего анализу, оказывается около 25, данное растение, подлежащее анализу, считается гаплоидным растением. Если пик сигнала растения, подлежащего анализу, оказывается около 50, считали, что положение обогащения интенсивности сигнала растения, подлежащего анализу, является таким же, как и у диплоидного ядра, и растение, подлежащее анализу, было диплоидом.

Результаты идентификации подсчитывали, и показатель индукции рассчитывали согласно следующей формуле: показатель индукции (%) = (число материнских гаплоидных растений/общее число растений) × 100. Как видно, после мутации гена *AtDMP8*, мутации гена *AtDMP9* и одновременной мутации генов *AtDMP8* и *AtDMP9* они подвергались гибридизации с другими материалами, и в потомстве мог быть получен гаплоидный родитель женского пола.

Таблица 4

Статистика по показателю индукции гаплоидов в межвидовом потомстве мутантов *dmp*

| Комбинация гибридизации | Номер растений | Общее число растений | Число гаплоидов | Показатель индукции гаплоидов (%) |
|------------------------------|----------------|----------------------|-----------------|-----------------------------------|
| <i>ms1</i> × WT | T1-13 | 1171 | 0 | 0,00 |
| | T1-17 | 1339 | 0 | 0,00 |
| | Всего | 2510 | 0 | 0,00 |
| <i>ms1</i> × <i>dmp8</i> | T1-34 | 1309 | 1 | 0,08 |
| | T2-33-1 | 1006 | 0 | 0,00 |
| | T2-33-2 | 719 | 0 | 0,00 |
| | Всего | 3034 | 1 | 0,03 |
| <i>ms1</i> × <i>dmp9</i> | T2-33-4 | 844 | 3 | 0,36 |
| | T2-33-5 | 598 | 3 | 0,50 |
| | T2-33-6 | 589 | 2 | 0,34 |
| | Всего | 2031 | 8 | 0,39 |
| <i>ms1</i> × <i>dmp8dmp9</i> | T1-6 | 146 | 2 | 1,37 |
| | T1-11 | 169 | 4 | 2,37 |
| | T1-19 | 194 | 6 | 3,09 |
| | T1-24 | 68 | 3 | 4,41 |
| | T1-25 | 361 | 11 | 3,05 |
| | T1-28 | 327 | 3 | 0,92 |
| | T1-32 | 31 | 1 | 3,23 |
| | T2-38-1 | 851 | 17 | 2,00 |
| | T2-38-2 | 559 | 10 | 1,79 |
| | Всего | 2706 | 57 | 2,11 |

Пример 2. Получение нокаутированных по *SIDMP* мутантов томата и их применение.

1. Ген *SIDMP* нокаутировали с использованием системы CRISPR/Cas9.

Ген *SIDMP* у томата нокаутировали с использованием системы CRISPR/Cas9 с получением нокаутированного по гену *SIDMP* мутанта томата. Конкретные стадии были следующими.

1. Выбор последовательностей енРНК

Последовательность целевого сайта разрабатывали на гене *SIDMP*, и она имела длину 20 п.о.

Целевой сайт 1 находился в положениях 76-95 SEQ ID NO: 5, и последовательность целевого сайта 1 енРНК представляла собой TATCTCACTAATTACCACA (SEQ ID NO: 100).

Целевой сайт 2 находился в положениях 247-266 SEQ ID NO: 5, и последовательность целевого сайта 2 енРНК представляла собой TCTCCTTTTACCAAATACTGA (SEQ ID NO: 101).

2. Конструирование вектора CRISPR/Cas9.

Вектор CRISPR/Cas9 представляет собой рекомбинантный вектор (структурная диаграмма вектора показана на фиг. 9а), полученный соединением молекулы ДНК, представленной в SEQ ID NO: 8 (экспрессионный элемент енРНК), молекулы ДНК, представленной в SEQ ID NO: 11 (экспрессионный элемент Cas9), молекулы ДНК, представленной в SEQ ID NO: 10 (экспрессионный элемент флуоресцентного белка) и молекулы ДНК, представленной в SEQ ID NO: 12 (экспрессионный элемент *NptII*), с вектором pICSL4723 с использованием метода "золотых ворот". Молекула ДНК, показанная в SEQ ID NO: 8, последовательно содержит кодирующую последовательность енРНК, нацеленной на целевой сайт 1, кодирующую последовательность енРНК, нацеленной на целевой сайт 2, с промотором *AtU6-26*, используемым для инициации экспрессии кодирующей последовательности енРНК, находящейся около начала каждой кодирующей последовательности енРНК.

3. Получение трансгенных растений.

Вектор CRISPR/Cas9, полученный на стадии 2, трансформировали в компетентные клетки агробактерии GV3101 посредством теплового шока (компетентные клетки агробактерии GV3101 приобретали у Beijing Aosen Dingxin Biotechnology Co., Ltd., и они являются общедоступными) с получением рекомбинантного штамма GV3101/CRISPR/Cas9.

Затем эксплант семядоли томата дикого типа AC инфицировали рекомбинантным штаммом GV3101/CRISPR/Cas9 посредством применения способа трансформации инфицированием экспланта семядоли томата агробактерией (т.е. рекомбинантную агробактерию подвергали расширенному размножению при 28°C, и эксплант семядоли томата инфицировали раствором размноженной агробактерии), и поколение T0 трансгенного растения томата из экспланта семядоли получали после скрининга на резистентность к канамицину.

4. Идентификация трансгенных растений с мутацией гена *SIDMP*.

Листья трансгенного растения томата поколения T0, полученного на стадии 3, отбирали, геномную ДНК выделяли в качестве матрицы, и проводили ПЦР-амплификацию посредством применения следую-

ших праймеров с получением продуктов ПЦР-амплификации разных линий.

Последовательности праймеров для выявления последовательности мутанта по гену SIDMP были следующими:

SIDMPF2: ACTGCTTAGGATATTAAGTACCC (SEQ ID NO: 102);

SIDMPR1: TTTTGGCACATCGACACCAAG (SEQ ID NO: 103).

Продукты ПЦР-амплификации разных линий подвергали секвенированию по Сэнджеру, и результаты секвенирования сравнивали с геном SIDMP томата дикого типа AC. Генотип SIDMP идентифицировали согласно следующим принципам.

Если последовательность из последовательности целевого сайта имела характеристику дублета, генотип данной линии был гетерозиготным генотипом (т.е. ген SIDMP мутировал на одной из двух гомологичных хромосом и не мутировал на другой хромосоме), и данная линия представляла собой гетерозиготную мутантную линию трансгенного томата поколения T0.

Для последовательности с характеристиками специфичного синглета из последовательности целевого сайта, если она была такой же, как последовательность гена SIDMP томата дикого типа AC, генотип данной линии был геотипом дикого типа, то есть, не было мутации на последовательности гена SIDMP; если она отличалась от последовательности гена SIDMP томата дикого типа AC, генотип данной линии был гомозиготным (гены SIDMP мутировали на обеих гомологичных хромосомах), данная линия представляла собой гомозиготную мутантную линию трансгенного томата поколения T0.

Таблица 5

Типы мутаций гена SIDMP в трансгенном томате поколения T0

| Серийный номер | Генотип | аллель 1 | аллель 2 |
|----------------|------------------------|--|--|
| T0-1 | Дикий тип | WT | WT |
| T0-2 | Дикий тип | WT | WT |
| T0-3 | Дикий тип | WT | WT |
| T0-4 | Дикий тип | WT | WT |
| T0-5 | Дикий тип | WT | WT |
| T0-6 | Дикий тип | WT | WT |
| T0-6 | Дикий тип | WT | WT |
| T0-7 | Дикий тип | WT | WT |
| T0-8 | Дикий тип | WT | WT |
| T0-9 | Гетерозиготная мутация | TATCCTACTAATTTAC(SEQ ID NO:104)(-3п.н.)AAGGGGAAAAA (SEQ ID NO:105) | WT |
| T0-10 | Биаллельная мутация | TATCCTACTA(SEQ ID NO:106)(-167п.н.)TTTGGTAAAGG(SEQ ID NO:107) | TACTAATTTACC(SEQ ID NO:108)(-158п.н.)TATTTGGTAA(SEQ ID NO:109) |
| T0-11 | Гетерозиготная мутация | WT | TAATTTACC(C)ACAAGGT---ACTTCCC(-3п.н.)GTATTTG |
| T0-12 | Гетерозиготная мутация | WT | TAATTTACC(C)ACAAGGT---TCCCTCAGT(-3п.н.)TGGTAAAG |
| T0-12 | Гетерозиготная мутация | WT | TAATTTACC(C)ACAAGGT---TCCCTCAGT(-3п.н.) |

| | | | |
|-------|------------------------|---|--|
| | | | TGGTAAAG |
| T0-13 | Дикий тип | WT | WT |
| T0-14 | Гетерозиготная мутация | ACTAATTTACC(SEQ ID NO:110)(-157п.н.)GTATTTGGTAA(SEQ ID NO:111) | WT |
| T0-15 | Дикий тип | WT | WT |
| T0-16 | Дикий тип | WT | WT |
| T0-16 | Дикий тип | WT | WT |
| T0-17 | Дикий тип | WT | WT |
| T0-17 | Дикий тип | WT | WT |
| T0-18 | Дикий тип | WT | WT |
| T0-18 | Дикий тип | WT | WT |
| T0-19 | Дикий тип | WT | WT |
| T0-20 | Дикий тип | WT | WT |
| T0-21 | Дикий тип | WT | WT |
| T0-21 | Дикий тип | WT | WT |
| T0-22 | Дикий тип | WT | WT |
| T0-22 | Дикий тип | WT | WT |
| T0-24 | Дикий тип | WT | WT |
| T0-24 | Дикий тип | WT | WT |
| T0-25 | Дикий тип | WT | WT |
| T0-26 | Дикий тип | WT | WT |
| T0-27 | Дикий тип | WT | WT |
| T0-28 | Дикий тип | WT | WT |
| T0-29 | Дикий тип | WT | WT |
| T0-31 | Дикий тип | WT | WT |
| T0-32 | Дикий тип | WT | WT |
| T0-33 | Дикий тип | WT | WT |
| T0-34 | Биаллельная мутация | TACTAATTTACC(SEQ ID NO:112)(-157п.н.)GTATTTGGTAA(SEQ ID NO:113) | AATTTACC(C)ACAAGGT--- CTTCCCTCA(A) GTATT |

Примечание: "-" означает делецию основания.

5. Идентификация трансгенного томата поколения T1 генотипированием Мутантные линии T0-33 и T0-34 по гену SIDMP трансгенного томата T0, полученные на стадии 4, самоопыляли, после сбора семян засеивали с получением трансгенного томата поколения T1. Конкретный способ идентификации генотипа гена SIDMP трансгенного томата поколения T1 был следующим: с геномной ДНК трансгенного томата T1 в качестве матрицы идентифицировали генотип гена SIDMP трансгенного томата T1 посредством применения праймеров для определения мутантной последовательности SIDMP (SIDMPF2 и SIDMPR1) согласно способу на стадии 4, соответственно.

Полученные, наконец, гомозиготные мутантные линии по гену SIDMP трансгенного томата поколения T1 включали sldmp-1 и sldmp-2 со следующим типом мутации каждой линии:

Идентификация секвенированием показала, что: различие между геномной ДНК гомозиготных мутантных линий sldmp-1 по генам SIDMP трансгенного томата поколения T1 и геномной ДНК томата дикого типа AC заключается только в том, что в гене, кодирующем белок SIDMP, имела место вставка основания (C) в положениях 92-93 SEQ ID NO: 5.

Идентификация секвенированием показала, что: различие между геномной ДНК гомозиготных мутантных линий sldmp-2 по генам SIDMP трансгенного томата поколения T1 и геномной ДНК томата дикого типа AC заключается только в том, что в гене, кодирующем белок SIDMP, имела место делеция фрагмента в положениях 93-249 SEQ ID NO: 5.

II. Применение нокаутированного мутанта томата по гену SIDMP, индуцирующего получение гаплоида.

(A) Эффективность растения и показатель завязывания семян нокаутированных мутантов томата по гену SIDMP.

При сравнении эффективности мутантных растений томата с геном SIDMP, нокаутированным относительно томата дикого типа AC, и AC обнаружили, что нокаутирование гена SIDMP не влияет на рост растения (фиг. 6). Число семян на плод самоопыленного томата подсчитывали, причем среднее значение составило примерно 79,5 семян на плод самоопыленного для дикого типа, тогда как мутант sldmp продуцировал только 17 семян, что значительно меньше, чем продукция дикого типа (табл. 6). Данные результаты показали то, что мутация гена SIDMP приводила к снижению показателя завязывания семян, свидетельствуя о том, что мутант sldmp имеет способность к индукции гаплоида.

Таблица 6

Статистика по числу семян мутантов *sldmp* после самоопыления

| Материалы | Статистика по числу соцветий | Среднее число закладываемых семян |
|----------------|------------------------------|-----------------------------------|
| WT | 12 | 79,5±24,1 |
| <i>sldmp-1</i> | 14 | 19,7±7,1 |
| <i>sldmp-2</i> | 23 | 15,3±6,6 |

Примечание: WT представляет собой томат дикого типа AC.

(Б) Идентификация способности к индукции гибридизации нокаутированных мутантов томата по гену SIDMP.

Мутанты *sldmp* скрещивали с материалом поколения F₁, полученным посредством скрещивания томата AC и Micro-Tom, с получением межвидового потомства, и гаплоидов в данном межвидовом потомстве идентифицировали следующими способами.

1. Идентификация флуоресцентного мечения.

Вектор CRISPR/Cas9 несет экспрессионный элемент TagRFP (*Entanmaea quadricolor*), управляемый промотором AtOLEO1. Поскольку промотор AtOLEO1 специфично экспрессировался в зародышах зрелых семян, флуоресцентный сигнал TagRFP можно было наблюдать посредством флуоресцентного излучения. Следовательно, мутант, несущий данный экспрессионный элемент в качестве родителя мужского пола, гибридизовали с материалами другого нефлуоресцентного родителя женского пола с получением семян, где зародыши диплоидных семян демонстрировали красную флуоресценцию из-за наличия генома родителя мужского пола, тогда как зародыши гаплоидных семян не демонстрировали флуоресценции или демонстрировали слабую флуоресценцию из-за происхождения от родителя женского пола (фиг. 7).

2. Идентификация молекулярного маркера.

Семена без флуоресценции, идентифицированные на стадии 1, высаживали далее, выделяли геномную ДНК и проводили ПЦР-амплификацию с применением полиморфных праймеров SIDMPF2+SIDMPR1 между поколением F₁, полученным скрещиванием томата AC с Micro-Tom, и мутантом томата, нокаутированным по гену SIDMP, и амплифицированный продукт подвергали выявлению картины полос в агарозе или секвенированию. Если амплифицированный продукт одного растения, подлежащего анализу, демонстрировал 1 полосу, или результат секвенирования показывал график с одним пиком, считали, что одна полоса растения представляет собой картину полос родителя женского пола без картины полос материала родителя мужского пола, данное единичное растение представляло собой родителя гаплоидного женского пола. А если амплифицированный продукт одного растения, подлежащего анализу, демонстрировал две полосы, или результат секвенирования показывал график с гетерозиготным пиком, считали, что полоса одного растения является гетерозиготной картиной полос поколения F₁, полученного скрещиванием томата AC и Micro-Tom, и мутантного томата, нокаутированного по гену SIDMP, и данное единичное растение было нормальным межвидовым потомством и диплоидом.

3. Идентификация фенотипа зрелых растений.

Фенотипы растений, идентифицированных на стадиях 1 и 2, дополнительно осматривали, причем гаплоиды демонстрировали карликовость, узкие листья, растущие вверх, компактный тип растения и мужскую стерильность, тогда как диплоиды были высокими растениями с широкими и рассредоточенными листьями и нормальной фертильностью.

4. Идентификация на листьях посредством проточной цитометрии.

Проточную цитометрию проводили на растении с характеристиками гаплоида, полученном на стадии 3, где конкретный способ был следующим: выделяли ядра клеток молодых листьев растения, подлежащего анализу, и листья диплоидного *Arabidopsis thaliana* использовали в качестве контроля; сигнал затем выявляли проточным цитометром, сперва выявляли сигнал диплоидных ядер, и пиковое положение сигнала диплоидного ядра принимали за 100 (поскольку генетические материалы в диплоидных клетках находились в двойном количестве относительно гаплоидных клеток, пиковое положение сигнала гаплоидного ядра оказывалось около 50). И если пиковое положение сигнала ядра растения, подлежащего анализу, оказывается около 50, данное растение, подлежащее анализу, считается гаплоидным растением. Если пик сигнала растения, подлежащего анализу, оказывается около 100, считали, что положение обогащения интенсивности сигнала растения, подлежащего анализу, является таким же, как и у диплоидного ядра, и растение, подлежащее анализу, было диплоидом (фиг. 8).

Результаты идентификации подсчитывали, и показатель индукции рассчитывали согласно следующей формуле: показатель индукции (%) = (число материнских гаплоидных растений/общее число растений) × 100. После мутации гена SIDMP проводили скрещивание с другими материалами, и гаплоидный женский родитель мог быть получен в данном межвидовом потомстве.

Таблица 7.

Статистика по показателю индукции гаплоидов в межвидовом потомстве мутантов *sldpm*

| Комбинация гибридизации | Общее число растений | Число гаплоидов | Показатель индукции гаплоидов (%) |
|---------------------------------|----------------------|-----------------|-----------------------------------|
| F ₁ × WT | 954 | 0 | 0,00 |
| F ₁ × <i>sldmp-1</i> | 323 | 2 | 0,62 |
| F ₁ × <i>sldmp-2</i> | 629 | 3 | 0,99 |

Примечание: F₁ представляет собой межвидовое потомство между томатом AC и Micro-Tom, и WT представляет собой томат дикого типа AC.

Промышленные применения

Гены индукции партеногенетического гаплоида AtDMP8 и AtDMP9 клонированы из *Arabidopsis thaliana*. Эксперименты показали то, что мутации AtDMP8 и AtDMP9 могут дать индуцибельность партеногенетического гаплоида с обеспечением индукции двудольных сельскохозяйственных культур к продуцированию гаплоидов партеногенетическим способом. Настоящее изобретение было дополнительно подтверждено на томатах, и у томатов также обнаружили то, что мутация SIDMP может дать индуцибельность партеногенетического гаплоида. Данное изобретение закладывает важную основу для расширения применения технологии селекции гаплоидов на двудольных растениях и выявления биологического механизма продуцирования партеногенетического гаплоида. Принимая во внимание универсальность применения технологии селекции гаплоидов в современном промышленном разведении, данное изобретение имеет очень широкую сферу применения и перспективы на рынке.

ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

<110>China Agricultural University
 <120>Parthenogenetic haploid induction gene DMP and application thereof
 <130>GNPLY20045
 <160>113
 <170>PatentIn version 3.5

<210>1
 <211>856
 <212>DNA
 <213>*Arabidopsis thaliana*
 <400>1
 ccgtcgtagt aaagatagtt cgaacaaaaa tccaaaacac atttctcatt tcacaaaaca
 60
 aacgaaaact attgagtcac aaaacacaga gagaatggag aaaacagagg aaagcgtcgg
 120
 aatcagagtt tacacgacga caacgacgca aaatccgtca ccaacatcgt ctcggtcgcc
 180
 caaacctgtc cctctctctt cactgcctat gcttccggca ggagccgccc cgggaggagg
 240
 aaaaggtaga aaacgtcgca tggtagcgaa aggagttcaa aaaacggttt cgaagacatc
 300
 aatgctcgtc aatttccttc cgacaggaac tctcttgatg ttcgaaatgg ttcttccgac
 360
 aatctaccgt gacggagact gtaacggaat caacacactc atgattcatc ttctcttctt
 420
 tctttgcgca atgtcttggt tcttcttcca tttcaccgac agtttcaaag cctccgatgg
 480
 gaaaatttac tacggttttg tgactccacg tggactcgcg gtgttcatga aaccaccctc
 540
 gccgggggtt ggaggcggag atgtgattgc agagaaggag attccgggtga cggatgagag
 600
 gtataagttg agggttaatg actttgtgca ttcagtgatg agtgttttgg tttttatggc
 660
 gatcgcgttt tcggatcgga gagtcacagg gtgtttggtt ccaggaaaag agaaggagat
 720
 ggatcaagtt atggagagtt ttccgttaat ggtcgggaatt gtttgcagtg ctttgtttct

047686

780

tgtttttccg accagtcgat acggtgtcgg atgcatgtct acataataac taatttcact

840

tttctgttgt tttgtg

856

<210>2

<211>243

<212>PRT

<213>Arabidopsis thaliana

<400>2

Met Glu Lys Thr Glu Glu Ser Val Gly Ile Arg Val Tyr Thr Thr Thr
1 5 10 15

Thr Thr Gln Asn Pro Ser Pro Thr Ser Ser Arg Ser Pro Lys Pro Val
20 25 30

Pro Leu Ser Ser Leu Pro Met Leu Pro Ala Gly Ala Ala Ala Gly Gly
35 40 45

Gly Lys Gly Arg Lys Arg Arg Met Val Ala Lys Gly Val Gln Lys Thr
50 55 60

Val Ser Lys Thr Ser Met Leu Val Asn Phe Leu Pro Thr Gly Thr Leu
65 70 75 80

Leu Met Phe Glu Met Val Leu Pro Thr Ile Tyr Arg Asp Gly Asp Cys
85 90 95

Asn Gly Ile Asn Thr Leu Met Ile His Leu Leu Leu Leu Cys Ala
100 105 110

Met Ser Cys Phe Phe Phe His Phe Thr Asp Ser Phe Lys Ala Ser Asp
115 120 125

Gly Lys Ile Tyr Tyr Gly Phe Val Thr Pro Arg Gly Leu Ala Val Phe
130 135 140

Met Lys Pro Pro Ser Pro Gly Phe Gly Gly Gly Asp Val Ile Ala Glu
145 150 155 160

Lys Glu Ile Pro Val Thr Asp Glu Arg Tyr Lys Leu Arg Val Asn Asp
165 170 175

Phe Val His Ser Val Met Ser Val Leu Val Phe Met Ala Ile Ala Phe
180 185 190

Ser Asp Arg Arg Val Thr Gly Cys Leu Phe Pro Gly Lys Glu Lys Glu

047686

195 200 205

Met Asp Gln Val Met Glu Ser Phe Pro Leu Met Val Gly Ile Val Cys
 210 215 220

Ser Ala Leu Phe Leu Val Phe Pro Thr Ser Arg Tyr Gly Val Gly Cys
 225 230 235 240

Met Ser Thr

<210>3
 <211>992
 <212>DNA
 <213>Arabidopsis thaliana
 <400>3
 cgatgaaact ctataaccgt caataaccgc cgcgatttga aaaataaata caaaaatctc
 60
 attctccatt aaaccattta acacaacata cgaaaaactg gtaaatcaca aaagaaaaaa
 120
 acagagagaa acacacgaaa atggagaaaa cagaggaaag cgtcggaatc agagtttaca
 180
 cggcgactcc gccgcaaaaa ccatcaccat caccaccttc tcgttcacca aaaccctctc
 240
 taatctcttc attgccttcc ctcccgtcag gagccgccgc tggaggagga agaggtcgaa
 300
 aacgtcgcac ggtggcgcaa ggagttcaaa aaacggtttc gaagacatca atgctcgtca
 360
 acttccttcc gacaggaaca ctcttgatgt tcgaaatggc tcttccatca atataccgtg
 420
 acggagactg taacggaatc aacacactca tgattcatct cctcttgctt ctttgcgcaa
 480
 tgtcttgctt cttcttccat tttaccgaca gtttcaaagc atccgatggg aagatctact
 540
 acggtttcgt gacgccacgt ggactcgcgg tgttcatgaa accgccgcct ccagagtttg
 600
 gtggcggaga tggtatagcg gaggcagaga ttccgggtgac tgatgatagg tataagttga
 660
 cggttaatga ctttgttcat gcagtgatga gcgttttggc gtttatggcg attgcgtttt
 720
 cggatcgaag agtcacggga tgtttgtttc cagggaaaga gaaagagatg gatcaagtta
 780
 tggagagttt tccaataatg gttggaattg tttgtagtgc tttgtttctt gtttttccga
 840
 ccaactcgata tgggtgttggc tgcattgactg gtttaattat tttcactttt ctgttttttg
 900
 gtttgtttta gtatctatct cgtattttcc ttgtaattta gattatctga tttaaactcc
 960
 aatggttctac tatttttatc tattgtattg tt
 992

<210>4
 <211>244
 <212>PRT
 <213>Arabidopsis thaliana
 <400>4
 Met Glu Lys Thr Glu Glu Ser Val Gly Ile Arg Val Tyr Thr Ala Thr
 1 5 10 15

047686

Pro Pro Gln Lys Pro Ser Pro Ser Pro Pro Ser Arg Ser Pro Lys Pro
 20 25 30

Val Leu Ile Ser Ser Leu Pro Ser Leu Pro Ser Gly Ala Ala Ala Gly
 35 40 45

Gly Gly Arg Gly Arg Lys Arg Arg Met Val Ala Gln Gly Val Gln Lys
 50 55 60

Thr Val Ser Lys Thr Ser Met Leu Val Asn Phe Leu Pro Thr Gly Thr
 65 70 75 80

Leu Leu Met Phe Glu Met Val Leu Pro Ser Ile Tyr Arg Asp Gly Asp
 85 90 95

Cys Asn Gly Ile Asn Thr Leu Met Ile His Leu Leu Leu Leu Cys
 100 105 110

Ala Met Ser Cys Phe Phe Phe His Phe Thr Asp Ser Phe Lys Ala Ser
 115 120 125

Asp Gly Lys Ile Tyr Tyr Gly Phe Val Thr Pro Arg Gly Leu Ala Val
 130 135 140

Phe Met Lys Pro Pro Pro Pro Glu Phe Gly Gly Gly Asp Val Ile Ala
 145 150 155 160

Glu Ala Glu Ile Pro Val Thr Asp Asp Arg Tyr Lys Leu Thr Val Asn
 165 170 175

Asp Phe Val His Ala Val Met Ser Val Leu Val Phe Met Ala Ile Ala
 180 185 190

Phe Ser Asp Arg Arg Val Thr Gly Cys Leu Phe Pro Gly Lys Glu Lys
 195 200 205

Glu Met Asp Gln Val Met Glu Ser Phe Pro Ile Met Val Gly Ile Val
 210 215 220

Cys Ser Ala Leu Phe Leu Val Phe Pro Thr Thr Arg Tyr Gly Val Gly
 225 230 235 240

Cys Met Thr Gly

047686

<210>5
 <211>757
 <212>DNA
 <213>Solanum lycopersicum
 <400>5
 atggagcaaa ctagtgaagg aattggaata aaaatgtata gtacatcgaa acgcgtcgat
 60
 aattcatcgt ctatgtatcc tactaattta ccacaagggtg aaataatccc agaattacca
 120
 ataccaattg gtggtaaaaa aagaagagca atggcaaatg gtgtacaaaa aacactttca
 180
 aaaacttcat tacttggttaa ttttctccca acgggaaccc ttttaacatt tgaaatgta
 240
 cttccctcag tatttggtaa aggagattgt tcaccaatta ctacatttat gattttaacg
 300
 ttacttggac tttgtacttt gtcattgttt ttctttcatt ttaccgatag ttttcgaggt
 360
 cctgatggta aaatttacta tggttttggt acaccaagag gtttgaaagt tttcaagact
 420
 ggacttgggtg tcgatgtgcc aaaagatgaa aggtacattg tgggagtgc acatttttga
 480
 catgcaatga tgtctgtttt ggtgtttgtg gcaattgcat tttctgatca cagagtaaca
 540
 ctttgtctat ttctggaca tgctaaagaa cttgatgaaa ttatgaggag ttttccatta
 600
 atggttggag ttattttagt tggacttttt cttgtttttc ctaattcgag atatggtgtt
 660
 ggatgtatgt ctgcttagat atctacgttt acattctttc atgtattgta atatacgttg
 720
 gttgatattt atttaaaatc tatctatctt gtattga
 757

<210>6
 <211>225
 <212>PRT
 <213>Solanum lycopersicum
 <400>6
 Met Glu Gln Thr Ser Glu Gly Ile Gly Ile Lys Met Tyr Ser Thr Ser
 1 5 10 15
 Lys Arg Val Asp Asn Ser Ser Ser Met Tyr Pro Thr Asn Leu Pro Gln
 20 25 30
 Gly Glu Ile Ile Pro Glu Leu Pro Ile Pro Ile Gly Gly Lys Lys Arg
 35 40 45
 Arg Ala Met Ala Asn Gly Val Gln Lys Thr Leu Ser Lys Thr Ser Leu
 50 55 60
 Leu Val Asn Phe Leu Pro Thr Gly Thr Leu Leu Thr Phe Glu Met Leu
 65 70 75 80
 Leu Pro Ser Val Phe Gly Lys Gly Asp Cys Ser Pro Ile Thr Thr Phe
 85 90 95

Met Ile Leu Thr Leu Leu Gly Leu Cys Thr Leu Ser Cys Phe Phe Phe
 100 105 110

His Phe Thr Asp Ser Phe Arg Gly Pro Asp Gly Lys Ile Tyr Tyr Gly
 115 120 125

Phe Val Thr Pro Arg Gly Leu Lys Val Phe Lys Thr Gly Leu Gly Val
 130 135 140

Asp Val Pro Lys Asp Glu Arg Tyr Ile Val Gly Val Thr Asp Phe Val
 145 150 155 160

His Ala Met Met Ser Val Leu Val Phe Val Ala Ile Ala Phe Ser Asp
 165 170 175

His Arg Val Thr Leu Cys Leu Phe Pro Gly His Ala Lys Glu Leu Asp
 180 185 190

Glu Ile Met Arg Ser Phe Pro Leu Met Val Gly Val Ile Cys Ser Gly
 195 200 205

Leu Phe Leu Val Phe Pro Asn Ser Arg Tyr Gly Val Gly Cys Met Ser
 210 215 220

Ala
 225

<210>7

<211>1447

<212>DNA

<213>Artificial Sequence

<220>

<223>Синтезированная

<400>7

catcttcatt cttaagatat gaagataatc ttcaaaaggc ccctgggaat ctgaaagaag
 60

agaagcaggc ccatttatat gggaaagaac aatagtatctt cttatatagg cccatttaag
 120

ttgaaaacaa tcttcaaaag tcccacatcg cttagataag aaaacgaagc tgagtttata
 180

tacagctaga gtcgaagtag tgattgagaa aacagaggaa agcgtgtttt agagctagaa
 240

atagcaagtt aaaataaggc tagtccgtta tcaacttgaa aaagtggcac cgagtcggtg
 300

cttttttct agaccagct ttcttgatac aagttggcat tacgctttac gaattcccat
 360

ggggagcatc tcattctta agatatgaag ataatcttca aaaggcccct gggaatctga
 420

aagaagagaa gcaggcccat ttatatggga aagaacaata gtatttctta tataggccca
 480

tttaagttga aaacaatctt caaaagtccc acatcgctta gataagaaaa cgaagctgag

540
 tttatataca gctagagtcg aagtagtgat tgaagaggtc gaaaacgtcg cagttttaga
 600
 gctagaaata gcaagttaaa ataaggctag tccgttatca acttgaaaaa gtggcaccga
 660
 gtcggtgctt tttttctaga cccagctttc ttgtacaaag ttggcattac gctcagagaa
 720
 ttcgcatgcg gagcatcttc attcttaaga tatgaagata atcttcaaaa ggcccctggg
 780
 aatctgaaag aagagaagca ggcccattta tatgggaaag aacaatagta tttcttatat
 840
 aggcccattt aagttgaaaa caatcttcaa aagtcccaca tcgcttagat aagaaaacga
 900
 agctgagttt atatacagct agagtcgaag tagtgattgt caagagtgtt cctgtcggag
 960
 ttttagagct agaaatagca agttaaata aggctagtcc gttatcaact tgaaaaagtg
 1020
 gcaccgagtc ggtgcttttt ttctagacc agctttcttg taaaaagttg gcattacgct
 1080
 tgtggaattc gcatgcgagg catcttcatt cttaaagatat gaagataatc ttcaaaaggc
 1140
 ccctgggaat ctgaaagaag agaagcaggc ccatttatat gggaaagaac aatagtattt
 1200
 cttatatagg cccatttaag ttgaaaaca tcttcaaaaag tcccacatcg cttagataag
 1260
 aaaacgaagc tgagtttata tacagctaga gtcgaagtag tgattgatga acaccgcgag
 1320
 tccacggttt tagagctaga aatagcaagt taaaataagg ctagtccggt atcaacttga
 1380
 aaaagtggca ccgagtcggt gctttttttc tagaccagc tttcttgtag aaagttggca
 1440
 ttacgct
 1447

<210>8

<211>714

<212>DNA

<213>Artificial Sequence

<220>

<223>Синтезированная

<400>8

catcttcatt cttaaagatat gaagataatc ttcaaaaggc ccctgggaat ctgaaagaag
 60
 agaagcaggc ccatttatat gggaaagaac aatagtattt cttatatagg cccatttaag
 120
 ttgaaaaca tcttcaaaaag tcccacatcg cttagataag aaaacgaagc tgagtttata
 180
 tacagctaga gtcgaagtag tgattgtatc ctactaattt accacagttt tagagctaga
 240
 aatagcaagt taaaataagg ctagtccggt atcaacttga aaaagtggca ccgagtcggt
 300
 gctttttttc tagaccagc tttcttgtag aaagttggca ttacgcttgt ggaattcctc
 360
 gagggagcat cttcattctt aagatatgaa gataatcttc aaaaggcccc tgggaatctg
 420
 aaagaagaga agcaggccca tttatatggg aaagaacaat agtatttctt atataggccc
 480
 atttaagttg aaaacaatct tcaaaagtcc cacatcgctt agataagaaa acgaagctga
 540
 gtttatatac agctagagtc gaagtagtga ttgtctcctt taccaaatc tgagttttag

600
agctagaaat agcaagttaa aataaggcta gtccggtatc aacttgaaaa agtggcaccg
660
agtcggtgct ttttttctag acccagcttt cttgtacaaa gttggcatta cgct
714

<210>9

<211>6137

<212>DNA

<213>Artificial Sequence

<220>

<223>Синтезированная

<400>9

ctcaactttt gattcgctat ttgcagtgca cctgtggcgt tcatcacatc ttttgtggca
60
ctgtttgcac tggtcattgc tattacaaag gaccttcctg atgttgaagg agatcgaaag
120
taagtaactg cacgcataac cattttcttt ccgctctttg gctcaatcca tttgacagtc
180
aaagacaatg ttaaccagc tccgtttgat atattgtctt tatgtgtttg ttcaagcatg
240
tttagttaat catgcctttg attgatcttg aataggttcc aaatatcaac cctggcaaca
300
aaacttggag tgagaaacat tgcattcctc ggttctggac ttctgctagt aaattatggt
360
tcagccatat cactagcttt ctacatgcct cagggtgaatt catctatttc cgtcttaact
420
atctcggtta attaaagcac gaacaccatt actgcatgta gaagcttgat aaactatcgc
480
caccaattta tttttgttgc gatattgtta ctttcctcag tatgcagctt tgaaaagacc
540
aacctctta tcctttaaca atgaacaggc ttttagaggc agcttgatga ttcctgcaca
600
tgtgatcttg gcttcaggct taattttcca ggtaaagcat tatgagatac tcttatatct
660
cttacatact tttgagataa tgcacaagaa cttcataact atatgcttta gtttctgcat
720
ttgacactgc caaattcatt aatctcta atctttggtt ttgatctttg gtagacatgg
780
gtactagaaa aagcaaacta caccaaggta aaatactttt gtacaaacat aaactcgta
840
tcacggaaca tcaatggagt gtatatctaa cggagtgtag aacatttga ttattgcagg
900
aagctatctc aggatattat cggtttatat ggaatctctt ctacgcagag tatctgttat
960
tccccttctc ctagctttca atttcatggt gaggatatgc agttttcttt gtatatcatt
1020
cttcttcttc tttgtagctt ggagtcaaaa tcggttcctt catgtacata catcaaggat
1080
atgtccttct gaatttttat atcttgcaat aaaaatgctt gtaccaattg aaacaccagc
1140
tttttgagtt ctatgatcac tgacttgggt ctaaccaaaa aaaaaaaaaat gtttaattta
1200
catatctaaa agtaggttta gggaaaccta aacagtaaaa tatttgtata ttattcgaat
1260
ttcactcatc ataaaaactt aaattgcacc ataaaatttt gtttactat taatgatgta
1320
atttgtgtaa ctaagataa aaataatatt ccgtaagtta accggctaaa accacgtata
1380
aaccagggaa cctgttaaac cggttcttta ctggataaag aatgaaagc ccatgtagac

1440
 agctccatta gagcccaaac cctaaatttc tcatctatat aaaaggagtg acattagggt
 1500
 ttttgttcgt cctcttaaag cttctcgttt tctctgccgt ctctctcatt cgcgcgacgc
 1560
 aaacgatctt caggtgatct tctttctcca aatcctctct cataactctg atttcgtact
 1620
 tgtgtatattg agctcacgct ctgtttctct caccacagcc tactgtatatt ttacaacaat
 1680
 taccaacaac aacaacaac aaacaacatt acaattacta ttacaatta caatggataa
 1740
 gaagtactct atcggactcg atatcggaac taactctgtg ggatgggctg tgatcaccga
 1800
 tgagtacaag gtgccatcta agaagttaa ggttctcgga aacaccgata ggcactctat
 1860
 caagaaaaac cttatcggtg ctctcctctt cgattctggg gaaactgctg aggctaccag
 1920
 actcaagaga accgctagaa gaaggtacac cagaagaaag aacaggatct gctacctcca
 1980
 agagatcttc tctaacgaga tggctaaagt ggatgattca ttcttcaca ggctcgaaga
 2040
 gtcattcctc gtggaagaag ataagaagca cgagaggcac cctatcttcg gaaacatcgt
 2100
 tgatgagggtg gcataccacg agaagtacc tactatctac cacctcagaa agaagctcgt
 2160
 tgattctact gataaggctg atctcaggct catctacctc gctctcgctc acatgatcaa
 2220
 gttcagagga cacttcctca tcgaggggtga tctcaaccct gataactctg atgtggataa
 2280
 gttgttcatc cagctcgtgc agacctaaa ccagcttttc gaagagaacc ctatcaacgc
 2340
 ttcaggtgtg gatgctaagg ctatcctctc tgctaggctc tctaagtcaa gaaggcttga
 2400
 gaacctcatt gctcagctcc ctggtgagaa gaagaacgga cttttcggaa acttgatcgc
 2460
 tctctctctc ggactcacc ctaacttcaa gtctaacttc gatctcgctg aggatgcaaa
 2520
 gctccagctc tcaaaggata cctacgatga tgatctcgat aacctcctcg ctcagatcgg
 2580
 agatcagtac gctgatttgt tcctcgtcgc taagaacctc tctgatgcta tcctcctcag
 2640
 tgatattcctc agagtgaaca ccgagatcac caaggctcca ctctcagctt ctatgatcaa
 2700
 gagatacgat gagcaccacc aggatctcac acttctcaag gctcttgta gacagcagct
 2760
 cccagagaag taaaagaga ttttcttcga tcagtctaag aacggatacg ctggttacat
 2820
 cgatgggtgt gcatctcaag aagagttcta caagttcatc aagcctatcc tcgagaagat
 2880
 ggatggaacc gaggaactcc tcgtgaagct caatagagag gatcttctca gaaagcagag
 2940
 gaccttcgat aacggatcta tcctcatca gatccacctc ggagagttgc acgctatcct
 3000
 tagaaggcaa gaggatttct acccattcct caaggataac agggaaaaga ttgagaagat
 3060
 tctcaccttc agaatccctt actacgtggg acctctcgct agaggaaact caagattcgc
 3120
 ttggatgacc agaaagtctg aggaaacat cacccttggt aacttcgaag aggtggtgga
 3180
 taaggggtgct agtgctcagt ctttcatcga gaggatgacc aacttcgata agaaccttc

3240
 aaacgagaag gtgctcccta agcactcttt gctctacgag tacttcaccg tgtacaacga
 3300
 gttgaccaag gttaagtacg tgaccgaggg aatgaggaag cctgcttttt tgtcaggtga
 3360
 gcaaaaagaag gctatcgttg atctcttggt caagaccaac agaaaggtga ccgtgaagca
 3420
 gctcaaagag gattacttca agaaaatcga gtgcttcgat tcagttgaga tttctggtgt
 3480
 tgaggatagg ttcaacgcat ctctcggaac ctaccacgat ctctcaaga tcattaagga
 3540
 taaggatttc ttggataacg aggaaaacga ggatatcttg gaggatatcg ttcttacct
 3600
 caccctcttt gaagatagag agatgattga agaaaggctc aagacctacg ctcatctctt
 3660
 cgatgataag gtgatgaagc agttgaagag aagaagatac actggttggg gaaggctctc
 3720
 aagaaagctc attaacggaa tcagggataa gcagtctgga aagacaatcc ttgatttctc
 3780
 caagtctgat ggattcgcta acagaaactt catgcagctc atccacgatg attctctcac
 3840
 ctttaaagag gatatccaga aggctcaggt ttcaggacag ggtgatagtc tccatgagca
 3900
 tatcgctaac ctcgctggat ctctgcaat caagaaggga atcctccaga ctgtgaaggt
 3960
 tgtggatgag ttggtgaagg tgatgggaag gcataagcct gagaacatcg tgatcgaat
 4020
 ggctagagag aaccagacca ctcagaaggg acagaagaac tctagggaaa ggatgaagag
 4080
 gatcgaggaa ggtatcaaag agcttggatc tcagatcctc aaagagcacc ctggtgagaa
 4140
 cactcagctc cagaatgaga agctctacct ctactacctc cagaacggaa gggatatgta
 4200
 tgtggatcaa gagttggata tcaacaggct ctctgattac gatggtgatc atatcgtgcc
 4260
 acagtcattc ttgaaggatg attctatcga taacaagggtg ctaccaggt ctgataagaa
 4320
 caggggtaag agtgataacg tgccaagtga agaggttgtg aagaaaatga agaactattg
 4380
 gaggcagctc ctcaacgcta agctcatcac tcagagaaag ttcgataact tgactaaggc
 4440
 tgagagggga ggactctctg aattggataa ggcaggattc atcaagaggc agcttgtgga
 4500
 aaccaggcag atcactaagc acgttgaca gatcctcgat tctaggatga acaccaagta
 4560
 cgatgagaac gataagttga tcaggaagt gaaggttatc accctcaagt caaagctcgt
 4620
 gtctgatttc agaaaggatt tccaattcta caaggtgagg gaaatcaaca actaccacca
 4680
 cgctcacgat gcttaccta acgctggtgt tggaaccgct ctcatcaaga agtatcctaa
 4740
 gctcgagtca gagttcgtgt acggtgatta caaggtgtac gatgtgagga agatgatcgc
 4800
 taagtctgag caagagatcg gaaaggctac cgctaagtat ttcttctact ctaacatcat
 4860
 gaatttcttc aagaccgaga ttaccctcgc taacggtgag atcagaaaga ggccactcat
 4920
 cgagacaaac ggtgaaacag gtgagatcgt gtgggataag ggaagggtt tcgctaccgt
 4980
 tagaaagggtg ctctctatgc cacaggtgaa catcgttaag aaaaccgagg tgcagaccgg

5040
 tggattctct aaagagtcta tcctccctaa gaggaactct gataagctca ttgctaggaa
 5100
 gaaggattgg gaccctaaga aatacgggtg tttcgattct cctaccgtgg cttactctgt
 5160
 tctcgtttgtg gctaaggttg agaagggaaa gagtaagaag ctcaagtctg ttaaggaact
 5220
 tctcggaaatc actatcatgg aaaggtcatc tttcgagaag aacccaatcg atttcctcga
 5280
 ggctaaggga tacaagagg ttaagaagga tctcatcatc aagctcccaa agtactcact
 5340
 cttcgaactc gagaacggta gaaagaggat gctcgttct gctggtgagc ttcaaaaggg
 5400
 aaacgagctt gctctcccat ctaagtacgt taactttctt tacctcgctt ctcaactacga
 5460
 gaagttgaag ggatctccag aagataacga gcagaagcaa cttttcgttg agcagcacia
 5520
 gcactacttg gatgagatca tcgagcagat ctctgagttc tctaaaaggg tgatcctcgc
 5580
 tgatgcaaac ctcgataagg tgttgtctgc ttacaacaag cacagagata agcctatcag
 5640
 ggaacaggca gagaacatca tccatctctt cacccttacc aacctcgggtg ctctgctgc
 5700
 tttcaagtac ttcgatacaa ccatcgatag gaagagatac acctctacca aagaagtgct
 5760
 cgatgctacc ctcattccatc agtctatcac tggactctac gagactagga tcgatctctc
 5820
 acagctcggg ggtgattcaa gggctgatcc taagaagaag aggaaggttt gagcttgtca
 5880
 agcagatcgt tcaaacattt ggcaataaag tttcttaaga ttgaatcctg ttgccggtct
 5940
 tgcgatgatt atcatataat ttctgttgaa ttacgttaag catgtaataa ttaacatgta
 6000
 atgcatgacg ttatttatga gatgggtttt tatgattaga gtcccgaat tatacattta
 6060
 atacgcgata gaaaacaaaa tatagcgcgc aaactaggat aaattatcgc gcgcggtgtc
 6120
 atctatgtta ctagatc
 6137

<210>10

<211>2311

<212>DNA

<213>Artificial Sequence

<220>

<223>Синтезированная

<400>10

aatgtcgcgg aacaattttt aaaactaaat cctaaatttt tctaattttg ttgccaatag
 60
 tggatatgtg ggccgtatag aaggaatcta ttgaaggccc aaaccatac tgacgagccc
 120
 aaaggttcgt tttgcgtttt atgtttcggg tcgatgccaa cgccacattc tgagctaggc
 180
 aaaaaacaaa cgtgtctttg aatagactcc tctcgttaac acatgcagcg gctgcatggt
 240
 gacgccatta acacgtggcc tacaattgca tgatgtctcc attgacacgt gacttctcgt
 300
 ctcttttctt aatatatcta acaaacactc ctacctcttc caaatatat acacatcttt
 360
 ttgatcaatc tctcattcaa aatctcattc tctctagtaa acaagaacaa aaaaatggcg

420
 gatacagcta gaggaaccca tcacgatatc atcggcagag atcagtaccc gatgatgggc
 480
 cgagatcgtg accagtacca gatgtccgga cgaggatctg actactccaa gtctaggcag
 540
 attgctaaag ctgcaactgc tgtcacagct ggtggttccc tccttgttct ctccagcctt
 600
 acccttggtg gaactgtcat agctttgact gttgcaacac ctctgctcgt tatcttcagc
 660
 ccaatccttg tcccggctct catcacagtt gcactcctca tcaccggttt tctttcctct
 720
 ggagggtttg gcattgccgc tataaccgtt ttctcttgga ttacaagta agcacacatt
 780
 tatcatctta cttcataatt ttgtgcaata tgtgcatgca tgtggtgagc cagtagcttt
 840
 ggatcaattt ttttggtcga ataacaaatg taacaataag aaattgcaaa ttctagggaa
 900
 catttggtta actaaatacg aaatttgacc tagctagctt gaatgtgtct gtgtatatca
 960
 tctatatag taaaatgctt ggtatgatac ctattgattg tgaataggta cgcaacggga
 1020
 gagcaccac agggatcaga caagttggac agtgcaagga tgaagttggg aagcaaagct
 1080
 caggatctga aagacagagc tcagtactac ggacagcaac atactgggtg ggaacatgac
 1140
 cgtgaccgta ctcggtggg ccagcacact actatgagcg agctgattaa ggagaacatg
 1200
 cacatgaagc tgtacatgga gggcaccgtg aacaaccacc acttcaagtg cacatccgag
 1260
 ggcgaaggca agccctacga gggcaccag accatgagaa tcaaggtggg cgagggcggc
 1320
 cctctcccct tgccttcga catcctggct accagcttca tgtacggcag cagaacctc
 1380
 atcaaccaca cccagggcat ccccgacttc tttaaagcagt cttccctga gggcttcaca
 1440
 tgggagagag tcaccacata cgaagatggg ggcgtgctga ccgctacca ggacaccagc
 1500
 ctccaggacg gctgcctcat ctacaacgtc aagatcagag gggatgaactt cccatccaac
 1560
 ggcctgtgga tgcagaagaa aacactcggc tgggaggcca acaccgagat gctgtacccc
 1620
 gctgacggcg gcctggaagg cagaagcgac atggccctga agctcgtggg cgggggccac
 1680
 ctgatctgca acttcaagac cacatacaga tccaagaaac ccgctaagaa cctcaagatg
 1740
 cccggcgtct actatgtgga ccacagactg gaaagaatca aggaggccga caaagaaacc
 1800
 tacgtcgagc agcacgaggt ggctgtggc agatactgcg acctccctag caaactggg
 1860
 cacaagtgag cttacccac tgatgtcatc gtcatagtcc aataactcca atgtcgggga
 1920
 gttagtttat gaggaataaa gtgttttagaa tttgatcagg gggagataat aaaagccgag
 1980
 tttgaatctt tttgttataa gtaatgttta tgtgtgtttc tatatgttgt caaatggtac
 2040
 catgtttttt ttcctctctt tttgtaactt gcaagtgttg tgttgtactt tatttggtt
 2100
 ctttgtaagt tggtaacggg ggtctatata tggaaaaggc cttgttttgt taaacttatg
 2160
 ttagttaact ggattcgtct ttaaccacaa aaagttttca ataagctaca aatttagaca

2220
 cgcaagccga tgcagtcatt agtacatata tttattgcaa gtgattacat ggcaacccaa
 2280
 acttcaaaaa cagtagggtg ctccatttag t
 2311

<210>11

<211>5247

<212>DNA

<213>Artificial Sequence

<220>

<223>Синтезированная

<400>11

gtcaacatgg tggagcacga cactctggtc tactccaaaa atgtcaaaga tacagtctca
 60
 gaagatcaaa gggctattga gacttttcaa caaaggataa ttcgggaaa cctcctcgga
 120
 ttccattgcc cagctatctg tcacttcatc gaaaggacag tagaaaagga aggtggctcc
 180
 tacaatgcc atcattgcca taaaggaaag gctatcattc aagatctctc tgccgacagt
 240
 ggtcccaaaag atggaccccc acccacgagg agcatcgtgg aaaaagaaga ggtccaacc
 300
 acgtctacaa agcaagtgga ttgatgtgat aacatgggtg agcacgacac tctggtctac
 360
 tccaaaaatg tcaaagatac agtctcagaa gatcaaaggg ctattgagac ttttcaacaa
 420
 aggataatth cgggaaacct cctcggattc cattgcccag ctatctgtca cttcatcgaa
 480
 aggacagtag aaaaggaagg tggctcctac aaatgccatc attgcgataa aggaaaggct
 540
 atcattcaag atctctctgc cgacagtggc cccaaagatg gacccccacc cacgaggagc
 600
 atcgtggaaa aagaagaggc tccaaccacg tctacaaagc aagtggattg atgtgacatc
 660
 tccactgacg taagggatga cgcacaatcc cactatcctt cgcaagaccc ttcctctata
 720
 taaggaagtt catttcattt ggagaggaca cgctcgagta taagagctca tttttacaac
 780
 aattaccaac aacaacaac aacaacaac attacaatta catttacaat tatcgatata
 840
 atggacaaga agtactccat tgggctcgat atcggcacia acagcgtcgg ctgggccgtc
 900
 attacggacg agtacaaggc gccgagcaaa aaattcaaag ttctgggcaa taccgatcgc
 960
 cacagcataa agaagaacct cattggcgcc ctctgttctg actccgggga gacggccgaa
 1020
 gccacgcggc tcaaaagaac agcacggcgc agatataccc gcagaaagaa tcggatctgc
 1080
 tacctgcagg agatcttttag taatgagatg gctaagggtg atgactcttt cttccatagg
 1140
 ctggaggagt cttttttggt ggaggaggat aaaaagcacg agcggccacc aatctttggc
 1200
 aatatacgtg acgaggtggc gtaccatgaa aagtacccaa ccatatatca tctgaggaag
 1260
 aagctttag acagactga taaggctgac ttgcggttga tctatctcgc gctggcgcat
 1320
 atgatcaaat ttcggggaca cttcctcatc gagggggacc tgaaccaga caacagcgat
 1380
 gtcgacaaac tctttatcca actggttcag acttacaatc agcttttcga agagaaccgg

1440
 atcaacgcat cgggagttga cgccaaagca atcctgagcg ctaggctgtc caaatcccgg
 1500
 cggctcgaaa acctcatcgc acagctccct ggggagaaga agaacggcct gtttgtaat
 1560
 cttatcgccc tgtcactcgg gctgaccccc aactttaaat ctaacttcga cctggccgaa
 1620
 gatgccaaagc ttcaactgag caaagacacc tacgatgatg atctcgacaa tctgctggcc
 1680
 cagatcggcg accagtacgc agaccttttt ttggcggcaa agaacctgtc agacgccatt
 1740
 ctgctgagtg atattctgcg agtgaacacg gagatcacca aagctccgct gagcgctagt
 1800
 atgatcaagc gctatgatga gcaccaccaa gacttgactt tgctgaaggc ccttgtcaga
 1860
 cagcaactgc ctgagaagta caaggaaatt ttcttcgatc agtctaaaaa tggctacgcc
 1920
 ggatacattg acggcggagc aagccaggag gaatthttaca aatttattaa gcccatcttg
 1980
 gaaaaaatgg acggcaccga ggagctgctg gtaaagctta acagagaaga tctggtgcg
 2040
 aaacagcgca ctttcgacaa tgggaagcatc ccccaccaga ttcacctggg cgaactgcac
 2100
 gctatcctca ggcggcaaga ggatttctac ccctttttga aagataacag ggaaaagatt
 2160
 gagaaaatcc tcacatttcg gataccctac tatgtaggcc ccctcgcccg gggaaattcc
 2220
 agattcgcgt ggatgactcg caaatcagaa gagactatca ctccctggaa cttcgaggaa
 2280
 gtcgtggata agggggcctc tgcccagtc ttcacgaaa ggatgactaa ctttgataaa
 2340
 aatctgccta acgaaaaggt gcttcctaaa cactctctgc tgtacgagta cttcacagtt
 2400
 tataacgagc tcaccaaggt caaatacgtc acagaaggga tgagaaagcc agcattcctg
 2460
 tctggagagc agaagaaagc tatcgtggac ctctcttca agacgaaccg gaaagttacc
 2520
 gtgaaacagc tcaaagaaga ttatthtcaa aagattgaat gtttcgactc tgttgaaatc
 2580
 agcggagtgg aggatcgctt caacgcatcc ctgggaacgt atcacgatct cctgaaaatc
 2640
 attaaagaca aggacttcct ggacaatgag gagaacgagg acattcttga ggacattgtc
 2700
 ctaccctta cgttgthttga agataggag atgattgaag aacgcttgaa aacttacgct
 2760
 catctcttcg acgacaaagt catgaaacag ctcaagaggc gccgatatac aggatggggg
 2820
 cggctgtcaa gaaaactgat caatgggatc cgagacaagc agagtggaaa gacaatcctg
 2880
 gattthtcta agtccgatgg atthtccaac cggaacttca tgcagttgat ccatgatgac
 2940
 tctctcacct ttaaggagga catccagaaa gcacaagtht ctggccaggg ggacagtctc
 3000
 cacgagcaca tcgctaattc tgcaggtagc ccagctatca aaaagggaat actgcagacc
 3060
 gttaaagtcg tggatgaact cgtcaaagta atgggaaggc ataagcccga gaatatcgtt
 3120
 atcgagatgg cccgagagaa ccaaactacc cagaaggagc agaagaacag tagggaaagg
 3180
 atgaagagga ttgaagaggg tataaaagaa ctgggggtccc aaatccttaa ggaacacca

3240
 gttgaaaaca cccagcttca gaatgagaag ctctacctgt actacctgca gaacggcagg
 3300
 gacatgtacg tggatcagga actggacatc aatcggctct ccgactacga cgtggatcat
 3360
 atcgtgcccc agtcttttct caaagatgat tctattgata ataaagtgtt gacaagatcc
 3420
 gataaaaata gaggaagag tgataacgtc ccctcagaag aagttgtcaa gaaaatgaaa
 3480
 aattattggc ggcagctgct gaacgcaaaa ctgatcacac aacggaagtt cgataatctg
 3540
 actaaggctg aacgaggtgg cctgtctgag ttggataaag cgggcttcat caaaaggcag
 3600
 cttgttgaga cacgccagat caccaagcac gtggcccaaaa ttctcgattc acgcatgaac
 3660
 accaagtacg atgaaaatga caaactgatt cgagaggtga aagttattac tctgaagtct
 3720
 aagctggttt cagatttcag aaaggacttt cagttttata aggtgagaga gatcaacaat
 3780
 taccaccatg cgcgatgatgc ctacctgaat gcagtggtag gcaactgcact tatcaaaaaa
 3840
 tatcccaagc ttgaatctga atttgtttac ggagactata aagtgtacga tgttaggaaa
 3900
 atgatcgcaa agtctgagca ggaaataggc aaggccaccg ctaagtactt cttttacagc
 3960
 aatattatga attttttcaa gaccgagatt acaactggcca atggagagat tcggaagcga
 4020
 ccacttatcg aaacaaacgg agaaacagga gaaatcgtgt gggacaaggg tagggatttc
 4080
 ggcacagtcc ggaaggtcct gtccatgccg caggtgaaca tcgttaaaaa gaccgaagta
 4140
 cagaccggag gcttctccaa ggaaagtatc ctcccgaaaa ggaacagcga caagctgatc
 4200
 gcacgcaaaa aagattggga cccaagaaa tacggcggat tcgattctcc tacagtcgct
 4260
 tacagtgtac tggttgtggc caaagtggag aaaggggaagt ctaaaaaact caaaagcgtc
 4320
 aaggaactgc tgggcatcac aatcatggag cgatcaagct tcgaaaaaaaa ccccatcgac
 4380
 tttctcgagg cgaaaggata taaagaggtc aaaaaagacc tcatcattaa gcttccaag
 4440
 tactctctct ttgagcttga aaacggccgg aaacgaatgc tcgctagtgc gggcgagctg
 4500
 cagaaaggta acgagctggc actgccctct aaatacgtta atttcttgta tctggccagc
 4560
 cactatgaaa agctcaaagg atctcccga gataatgagc agaagcagct gttcgtggaa
 4620
 caacacaaac actaccttga tgagatcatc gagcaaataa gcgaattctc caaaagagtg
 4680
 atcctcgccg acgctaacct cgataagggtg ctttctgctt acaataagca cagggataag
 4740
 cccatcaggg agcaggcaga aaacattatc cacttgttta ctctgaccaa cttgggcgcg
 4800
 cctgcagcct tcaagtactt cgacaccacc atagacagaa agcggtacac ctctacaaag
 4860
 gaggtcctgg acgccacact gattcatcag tcaattacgg ggctctatga aacaagaatc
 4920
 gacctctctc agctcgggtg agacagcagg gctgacccca agaagaagag gaaggtgtga
 4980
 gcttgtcaag cagatcgttc aaacatttgg caataaagtt tcttaagatt gaatcctggt

5040
gccggtccttg cgatgattat catataattt ctggtgaatt acgttaagca tgtaataatt
5100
aacatgtaat gcatgacggt atttatgaga tgggttttta tgattagagt cccgcaatta
5160
tacatttaat acgcataga aaacaaaata tagcgcgcaa actaggataa attatcgcgc
5220
gcggtgtcat ctatgttact agatcga
5247

<210>12

<211>1867

<212>DNA

<213>Artificial Sequence

<220>

<223>Синтезированная

<400>12

cgagaatta agggagtcac gttatgacc cgcgatga cgcgggaca gccgttttac
60
gtttggaact gacagaaccg caacggtgaa ggagccactg agccgcgggt ttctggagtt
120
taatgagcta agcacatacg tcagaaacca ttattgcgcg ttcaaaagtc gcctaaggtc
180
actatcagct agcaaatatt tcttgtcaaa aatgctccac tgacgttcca taaattcccc
240
tcggtatcca attagagtct catattcact ctctatattt tacaacaatt accaacaaca
300
acaacaaca aacaacatta caattacatt tacaattacc atggttgaac aagatggatt
360
gcacgcaggt tctccggcgg cttgggtgga gaggtattc ggctatgact gggcacaaca
420
gacaatcggc tgctctgatg ccgccgtggt ccggctgtca gcgcaggggc gcccggttct
480
ttttgtcaag accgacctgt ccggtgccct gaatgaactg caggacgagg cagcgcggct
540
atcgtggctg gccacgacgg gcgttccttg cgcagctgtg ctcgacgttg tcaactgaagc
600
gggaaggac tggctgctat tgggcgaagt gccggggcag gatctcctgt catctcacct
660
tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg cggcggctgc atacgcttga
720
tccggctacc tgcccattcg accaccaagc gaaacatcg atcgagcgag cacgtactcg
780
gatggaagcc ggtcttgctg atcaggatga tctggacgaa gagcatcagg ggctcgcgcc
840
agccgaactg ttcgccaggc tcaaggcgcg catgcccgcg ggcgaggatc tcgtcgtgac
900
tcatggcgat gcctgcttgc cgaatatcat ggtggaaaat ggccgctttt ctggattcat
960
cgactgtggc cggctgggtg tggcggaccg ctatcaggac atagcgttgg ctaccctgga
1020
tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc ctctgctttt acggtatcgc
1080
cgctcccgat tcgcagcgca tcgccttcta tcgccttctt gacgagttct tctgagcggg
1140
actctggggt tcgctagagt cctgctttaa tgagatatgc gagacgccta tgatcgcgat
1200
atatttgctt tcaattctgt tgtgcacggt gtaaaaaacc tgagcatgtg tagctcagat
1260
ccttaccgcc ggtttcgggt cattctaag aatatatcac ccgttactat cgtattttta

1320
 tgaataatat tctccgttca atttactgat tgtaccctac tacttatatg tacaatatta
 1380
 aatgaaaac aatatattgt gctgaatagg tttatagcga catctatgat agagcgccac
 1440
 aataacaaac aattgcgttt tattattaca aatccaattt taaaaaagc ggcagaaccg
 1500
 gtcaaaccta aaagactgat tacataaatc ttattcaaat ttcaaaagtg ccccaggggc
 1560
 tagtatctac gacacaccga gcggcgaact aataacgctc actgaaggga actccggttc
 1620
 cccgccggcg cgcattgggtg agattccttg aagttgagta ttggccgtcc gctctaccga
 1680
 aagttacggg caccattcaa cccggtccag cacggcggcc gggtaaccga cttgctgccc
 1740
 cgagaattat gcagcatttt tttggtgat gtgggccccca aatgaagtgc aggtcaaacc
 1800
 ttgacagtga cgacaaatcg ttgggcgggt ccagggcgaa tttgcgaca acatgtcgag
 1860
 gctcagc
 1867

<210>13
 <211>16
 <212>DNA
 <213>Artificial Sequence
 <220>
 <223>Синтезированная
 <400>13
 gtttacacgg cgactc
 16

<210>14
 <211>401
 <212>DNA
 <213>Artificial Sequence
 <220>
 <223>Синтезированная
 <400>14
 cgtcggaaatc agagtttaca cggcgactcc gccgcaaaaa ccatcaccat caccaccttc
 60
 tcgttcacca aaaccgtctc taatctcttc attgccttcc ctcccgtcag gagccgccgc
 120
 tggaggagga agaggtcga aacgtcgcac ggtggcgcaa ggagttcaaa aaacggtttc
 180
 gaagacatca atgctcgtca acttccttcc gacaggaaca ctcttgatgt tcgaaatggt
 240
 tcttccatca atataccgtg acggagactg taacggaatc aacacactca tgattcatct
 300
 cctcttgctt ctttgcgcaa tgtcttgttt cttcttccat ttaccgaca gtttcaaagc
 360
 atccgatggg aagatctact acggtttcgt gacgccacgt g
 401

<210>15
 <211>20

<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>15
gagaaaacag aggaagcgt
20

<210>16
<211>20
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>16
aagaggtcga aaacgtcgca
20

<210>17
<211>20
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>17
tcaagagtgt tcctgtcgga
20

<210>18
<211>20
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>18
atgaacaccg cgagtccacg
20

<210>19
<211>23
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>19
tgcgaaatga gattggtttt ggg
23

<210>20
<211>20
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>20

aaacaccctg tgactctccg
20

<210>21
<211>20
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>21
ataaccgtca ataaccgссg
20

<210>22
<211>20
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>22
ссаgtcatgc аассаасс
20

<210>23
<211>15
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>23
ggggcgtcgg аатса
15

<210>24
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>24
gactссасgt
10

<210>25
<211>17
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>25
гаааgасgtc gгаатса
17

<210>26
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная

<400>26
gactccacgt
10

<210>27
<211>15
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>27
gggtcgtcgg aatca
15

<210>28
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>28
gactccacgt
10

<210>29
<211>11
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>29
cgtcggaatc a
11

<210>30
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>30
gactccacgt
10

<210>31
<211>17
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>31
gaaagacgtc ggaatca
17

<210>32
<211>10
<212>DNA
<213>Artificial Sequence

<220>
<223>Синтезированная
<400>32
gactccacgt
10

<210>33
<211>17
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>33
gaaagcgctc ggaatca
17

<210>34
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>34
gactccacgt
10

<210>35
<211>17
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>35
gaaagtcgctc ggaatca
17

<210>36
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>36
gactccacgt
10

<210>37
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>37
gactccacgt
10

<210>38
<211>15
<212>DNA

<213>Artificial Sequence
<220>
<223>Синтезированная
<400>38
ggggcgctcgg aatca
15

<210>39
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>39
gactccacgt
10

<210>40
<211>17
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>40
gaaagcgctc ggaatca
17

<210>41
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>41
gactccacgt
10

<210>42
<211>15
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>42
ggggcgctcgg aatca
15

<210>43
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>43
gactccacgt
10

<210>44
<211>17

<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>44
gaaagcgtc ggaatca
17

<210>45
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>45
gactccacgt
10

<210>46
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>46
gtcggaatca
10

<210>47
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>47
gactccacgt
10

<210>48
<211>17
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>48
gaaagcgtc ggaatca
17

<210>49
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>49
gactccacgt
10

<210>50

<211>15
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>50
ggggcgtcgg aatca
15

<210>51
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>51
gactccacgt
10

<210>52
<211>11
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>52
cgtcggaatc a
11

<210>53
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>53
gactccacgt
10

<210>54
<211>11
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>54
cgtcggaatc a
11

<210>55
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>55
gactccacgt
10

<210>56
<211>15
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>56
gggtcgtcgg aatca
15

<210>57
<211>15
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>57
gggtcgtcgg aatca
15

<210>58
<211>11
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>58
cgtcggaatc a
11

<210>59
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>59
gactccacgt
10

<210>60
<211>11
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>60
cgtcggaatc a
11

<210>61
<211>11
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>61
cgtcggaatc a
11

<210>62
<211>11
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>62
cgtcggaatc a
11

<210>63
<211>17
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>63
gaaagcgtc ggaatca
17

<210>64
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>64
gactccacgt
10

<210>65
<211>17
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>65
gaaagacgtc ggaatca
17

<210>66
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>66
gactccacgt
10

<210>67
<211>15
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>67

gggtcgtcgg aatca
15

<210>68
<211>15
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>68
gggtcgtcgg aatca
15

<210>69
<211>17
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>69
gaaagacgtc ggaatca
17

<210>70
<211>10
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>70
gactccacgt
10

<210>71
<211>15
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>71
ggggcgtcgg aatca
15

<210>72
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>72
gactccacgt
10

<210>73
<211>17
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>73

gaaagtcgtc ggaatca
17

<210>74
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>74
gactccacgt
10

<210>75
<211>11
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>75
ccacgtggac t
11

<210>76
<211>17
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>76
gaaagtcgtc ggaatca
17

<210>77
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>77

gactccacgt
10

<210>78
<211>13
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>78
ggaagacgt cgg
13

<210>79
<211>13
<212>DNA
<213>Artificial Sequence

<220>
<223>Синтезированная
<400>79
ggaagacgt cgg
13

<210>80
<211>13
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>80
ggaagacgt cgg
13

<210>81
<211>13
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>81
ggaagcgt cgg
13

<210>82
<211>13
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>82
ggaagcgt cgg
13

<210>83
<211>13
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>83
ggaagacgt cgg
13

<210>84
<211>13
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>84
ggaagcgt cgg
13

<210>85
<211>11

<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>85
tcgcggtggt c
11

<210>86
<211>13
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>86
ggaagacgt cgg
13

<210>87
<211>13
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>87
ggaagacgt cgg
13

<210>88
<211>13
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>88
ggaagacgt cgg
13

<210>89
<211>13
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>89
ggaagacgt cgg
13

<210>90
<211>13
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>90
ggaagacgt cgg
13

<210>91

<211>12
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>91
ggaaagcgtc gg
12

<210>92
<211>13
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>92
ggaaagtcgt cgg
13

<210>93
<211>14
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>93
gacgссacgt ggga
14

<210>94
<211>13
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>94
ggaaagtcgt cgg
13

<210>95
<211>14
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>95
gacgссacgt ggga
14

<210>96
<211>202
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>96
gaaattgacg agcattgatg tcttcgaaac cgttttttga actcctttcg ccaccatgсg
60

acgttttcta ccttttcctc ctcccgcggc ggctcctgcc ggaagcatag gcagtgaaga
120
gagagggaca ggtttgggcg accgagacga tgttgggtgac ggattttgcg tcgttgtcgt
180
cgtgtaaact ctgattccga cg
202

<210>97
<211>24
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>97
cagctgagat gaacgagttg tctt
24

<210>98
<211>24
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>98
tcttttgagt cactccgtat gtcc
24

<210>99
<211>28
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>99
ataataacgc tgcggacatc tacatttt
28

<210>100
<211>20
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>100
tataccta atttaccaca
20

<210>101
<211>20
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>101
tctcctttac caaatactga
20

<210>102
<211>24
<212>DNA

<213>Artificial Sequence
<220>
<223>Синтезированная
<400>102
actgcttagg atattaactg accc
24

<210>103
<211>21
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>103
ttttggсаса tcгасассаа g
21

<210>104
<211>16
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>104
tатсстacta atttac
16

<210>105
<211>11
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>105
ааggggаааа а
11

<210>106
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>106
tатсстacta
10

<210>107
<211>11
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>107
tttggtaaag g
11

<210>108

<211>12
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>108
tactaattta cc
12

<210>109
<211>10
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>109
tatttggtaa
10

<210>110
<211>11
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>110
actaatttac c
11

<210>111
<211>11
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>111
gtatttggt a
11

<210>112
<211>12
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>112
tactaattta cc
12

<210>113
<211>11
<212>DNA
<213>Artificial Sequence
<220>
<223>Синтезированная
<400>113
gtatttggt a
11

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ получения линии-индуктора гаплоидного растения, включающий нокаутирование гена DMP в геноме растения с получением линии-индуктора гаплоидного растения;

где данное растение является двудольным;

ген DMP представляет собой ген, показанный в следующих из B1), или B2), или B3), или B4):

B1) молекула кДНК или молекула геномной ДНК, показанная в SEQ ID NO: 1, или SEQ ID NO: 3, или SEQ ID NO: 5;

B2) молекула кДНК или молекула геномной ДНК, имеющая идентичность 75% или более с нуклеотидной последовательностью, определенной в B1);

B3) молекула кДНК или молекула геномной ДНК, происходящая из двудольного растения и имеющая идентичность 75% или более с нуклеотидной последовательностью, определенной в B1);

B4) молекула кДНК или молекула геномной ДНК, которая гибридизуется с нуклеотидной последовательностью, определенной в B1), или B2), или B3), в жестких условиях;

ген DMP имеет следующие функции: двудольное растение становится линией-индуктором гаплоидного растения, когда ген DMP подвергается нокаутированию в нем.

2. Способ по п.1, в котором двудольное растение представляет собой *Arabidopsis thaliana* или томат.

3. Способ по п.1, в котором указанное нокаутирование гена DMP осуществляют посредством CRISPR/Cas9, и

в котором, когда двудольное растение представляет собой *Arabidopsis thaliana*, целевой сайт CRISPR/Cas9 находится в положениях 98-117 SEQ ID NO: 1, положениях 290-309 SEQ ID NO: 3, положениях 368-387 SEQ ID NO: 3 и положениях 509-528 SEQ ID NO: 1, и

в котором, когда двудольное растение представляет собой томат, целевой сайт CRISPR/Cas9 находится в положениях 76-95 SEQ ID NO: 5, положениях 247-266 SEQ ID NO: 5.

4. Способ по любому из пп.1-3, дополнительно включающий стадии самоопыления полученной линии-индуктора гаплоидного растения.

5. Линия-индуктор гаплоидного растения, полученная любым способом по пп.1-4.

6. Линия-индуктор гаплоидного растения по п.5, где линия-индуктор гаплоидного растения, в частности, включает мутантную линию T1-34 *Arabidopsis thaliana*, мутантную линию T1-6 *Arabidopsis thaliana*, мутантную линию T1-11 *Arabidopsis thaliana*, мутантную линию T1-19 *Arabidopsis thaliana*, мутантную линию T1-24 *Arabidopsis thaliana*, мутантную линию T1-25 *Arabidopsis thaliana*, мутантную линию T1-28 *Arabidopsis thaliana*, мутантную линию T1-32 *Arabidopsis thaliana*, мутантную линию T2-33 *Arabidopsis thaliana*, мутантную линию T2-38 *Arabidopsis thaliana*, мутантную линию *sldmp-1* томата или мутантную линию *sldmp-2* томата;

различие между геномной ДНК мутантной линии T1-34 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, имеющем последовательность SEQ ID NO: 1 и кодирующем белок AtDMP8, имеет место делеция фрагмента в одной хромосоме в положениях 115-512 SEQ ID NO: 1 и имеет место вставка основания (T) в другой хромосоме в положениях 114-115 SEQ ID NO: 1;

различие между геномной ДНК мутантной линии T1-6 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, имеющем последовательность SEQ ID NO: 1 и кодирующем белок AtDMP8, имеет место делеция фрагмента в одной хромосоме в положениях 115-512 SEQ ID NO: 1 и имеет место делеция фрагмента в другой хромосоме в положениях 113-114 SEQ ID NO: 1; и в гене, имеющем последовательность SEQ ID NO: 3 и кодирующем белок AtDMP9, имеет место вставка фрагмента в обеих хромосомах в положениях 160-161 SEQ ID NO: 3, где нуклеотидная последовательность вставленного фрагмента показана как SEQ ID NO: 13;

различие между геномной ДНК мутантной линии T1-11 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, имеющем последовательность SEQ ID NO: 1 и кодирующем белок AtDMP8, имеет место вставка основания (T) в положениях 114-115 SEQ ID NO: 1; и в гене, имеющем последовательность SEQ ID NO: 3 и кодирующем белок AtDMP9, имеет место вставка основания (T) в одной хромосоме в положениях 160-161 SEQ ID NO: 3 и имеет место вставка фрагмента в другой хромосоме в положениях 160-161 SEQ ID NO: 3, где нуклеотидная последовательность вставленного фрагмента показана как SEQ ID NO: 13;

различие между геномом мутантной линии T1-19 *Arabidopsis thaliana* и геномом *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, имеющем последовательность SEQ ID NO: 1 и кодирующем белок AtDMP8, имеет место делеция основания (T) в одной хромосоме в положении 114 SEQ ID NO: 1 и имеет место делеция фрагмента в другой хромосоме в положениях 115-511 SEQ ID NO: 1; и в гене, имеющем последовательность SEQ ID NO: 3 и кодирующем белок AtDMP9, имеет место делеция фрагмента в одной хромосоме в положениях 161-560 SEQ ID NO: 3 и имеет место делеция фрагмента в другой хромосоме в положениях 161-564 SEQ ID NO: 3;

различие между геномной ДНК мутантной линии T1-24 *Arabidopsis thaliana* и геномной ДНК *Arabi-*

dopsis thaliana дикого типа Col-0 заключается только в том, что в гене, имеющем последовательность SEQ ID NO: 1 и кодирующем белок AtDMP8, имеет место делеция фрагмента в одной хромосоме в положениях 115-512 SEQ ID NO: 1 и имеет место вставка основания (T) в другой хромосоме в положениях 114-115 SEQ ID NO: 1; и в гене, имеющем последовательность SEQ ID NO: 3 и кодирующем белок AtDMP9, имеет место делеция фрагмента в одной хромосоме в положениях 161-560 SEQ ID NO: 3 и имеет место делеция фрагмента в другой хромосоме в положениях 159-160 SEQ ID NO: 3;

различие между геномной ДНК мутантной линии T1-25 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, имеющем последовательность SEQ ID NO: 1 и кодирующем белок AtDMP8, имеет место делеция основания (T) в одной хромосоме в положении 114 SEQ ID NO: 1 и имеет место вставка фрагмента CGT в другой хромосоме в положениях 114-115 SEQ ID NO: 1; и в гене, имеющем последовательность SEQ ID NO: 3 и кодирующем белок AtDMP9, имеет место делеция фрагмента в одной хромосоме в положениях 161-162 SEQ ID NO: 3 и имеет место вставка основания (A) в другой хромосоме в положениях 160-161 SEQ ID NO: 3;

различие между геномной ДНК мутантной линии T1-28 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, имеющем последовательность SEQ ID NO: 1 и кодирующем белок AtDMP8, имеет место делеция фрагмента в одной хромосоме в положениях 115-512 SEQ ID NO: 1 и имеет место делеция основания (T) в другой хромосоме в положении 114 SEQ ID NO: 1, и в гене, имеющем последовательность SEQ ID NO: 3 и кодирующем белок AtDMP9, имеет место вставка основания (A) в обеих хромосомах в положениях 160-161 SEQ ID NO: 3;

различие между геномной ДНК мутантной линии T1-32 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, имеющем последовательность SEQ ID NO: 1 и кодирующем белок AtDMP8, имеет место делеция фрагмента в одной хромосоме в положениях 115-511 SEQ ID NO: 1 и имеет место вставка основания (T) в другой хромосоме в положениях 114-115 SEQ ID NO: 1; и в гене, имеющем последовательность SEQ ID NO: 3 и кодирующем белок AtDMP9, имеет место делеция основания (C) в обеих хромосомах в положении 161 SEQ ID NO: 3;

различие между геномной ДНК мутантной линии T2-33 *Arabidopsis thaliana* и *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, имеющем последовательность SEQ ID NO: 3 и кодирующем белок AtDMP9, имеет место вставка основания (T) в обеих хромосомах в положениях 160-161 SEQ ID NO: 3 и имеет место вставка фрагмента в обеих хромосомах в положениях 561-562 SEQ ID NO: 3, где нуклеотидная последовательность вставленного фрагмента показана как SEQ ID NO: 14;

различие между геномной ДНК мутантной линии T2-38 *Arabidopsis thaliana* и геномной ДНК *Arabidopsis thaliana* дикого типа Col-0 заключается только в том, что в гене, имеющем последовательность SEQ ID NO: 1 и кодирующем белок AtDMP8, имеет место делеция фрагмента в обеих хромосомах в положениях 115-127 SEQ ID NO: 1; и в гене, имеющем последовательность SEQ ID NO: 3 и кодирующем белок AtDMP9, имеет место вставка основания (T) и основания (G) в обеих хромосомах в положениях 160-161 SEQ ID NO: 3 и в положениях 562-563 SEQ ID NO: 3, соответственно;

различие между геномной ДНК мутантной линии томата *sldmp-1* и геномной ДНК томата дикого типа AC заключается только в том, что в гене, имеющем последовательность SEQ ID NO: 5 и кодирующем белок SIDMP, имеет место вставка основания (C) в обеих хромосомах в положениях 92-93 SEQ ID NO: 5;

различие между геномной ДНК мутантной линии томата *sldmp-2* и геномной ДНК томата дикого типа AC заключается только в том, что в гене, имеющем последовательность SEQ ID NO: 5 и кодирующем белок SIDMP, имеет место делеция фрагмента в обеих хромосомах в положениях 93-249 SEQ ID NO: 5.

7. Линия-индуктор гаплоидного растения по п.5 или 6, включающая клетки, ткани и органы, происходящие от линии-индуктора гаплоидного растения; где данные органы включают семена, листья, плоды, цветки, стебли и корни.

8. Линия-индуктор гаплоидного растения по п.5 или 6, включающая материалы для размножения, которые включают группу, состоящую из пыльцы, завязей, семязачатков, зародышей, эндоспермов, яйцеклеток, эмбрионов на стадии дробления, корней, кончиков корней, гипокотилей, семядолей, стеблей, листьев, цветков, пыльников, семян, меристематических клеток, протопластов и культур ткани.

9. Способ получения гаплоидного растения, включающий следующие стадии: самоопыление линии-индуктора гаплоидного растения или его потомства или скрещивание линии-индуктора гаплоидного растения или его потомства с другими растительными материалами в качестве родителя мужского пола с получением самоопыленного потомства или межвидового потомства, а именно гаплоидного растения; где данное растение представляет собой двудольное растение; где указанная линия-индуктор гаплоидного растения получена способом по любому из пп.1-4.

10. Гаплоидное растение, полученное способом по п.9.

11. Гаплоидное растение по п.10, включающее клетки, ткани и органы, происходящие от линии-индуктора гаплоидного растения; где данные органы включают семена, листья, плоды, цветки, стебли и корни.

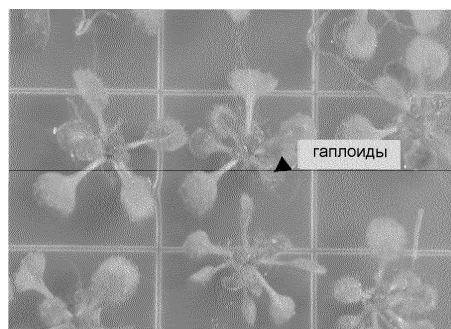
12. Гаплоидное растение по п.10, включающее материалы для размножения, которые включают

группу, состоящую из пыльцы, завязей, семязачатков, зародышей, эндоспермов, яйцеклеток, эмбрионов на стадии дробления, корней, кончиков корней, гипокотилей, семядолей, стеблей, листьев, цветков, пыльников, семян, меристематических клеток, протопластов и культур ткани.

13. Ген AtDMP8 в мутантной линии *Arabidopsis thaliana*, последовательность которого представляет собой SEQ ID NO: 1, в которой имеют место делеции и/или вставки, выбранные из тех, которые указаны в п.6.

14. Ген AtDMP9 в мутантной линии *Arabidopsis thaliana*, последовательность которого представляет собой SEQ ID NO: 3, в которой имеют место делеции и/или вставки, выбранные из тех, которые указаны в п.6.

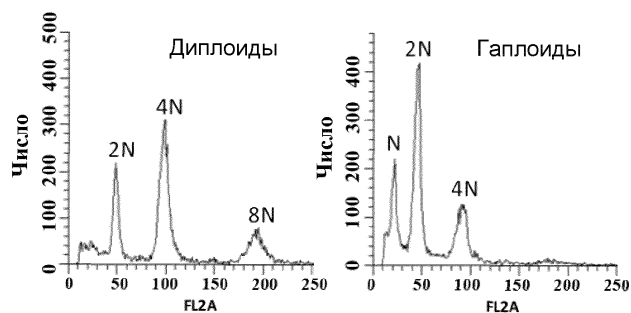
15. Ген SIDMP в мутантной линии томата, последовательность которого представляет собой SEQ ID NO: 5, в которой имеют место делеции и/или вставки, выбранные из тех, которые указаны в п.6.



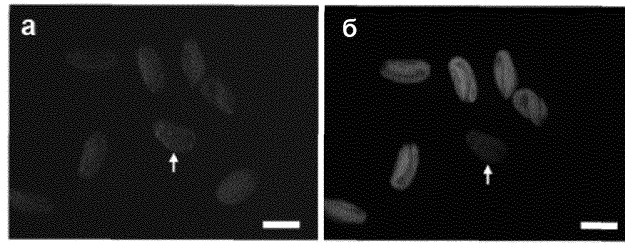
Фиг. 1



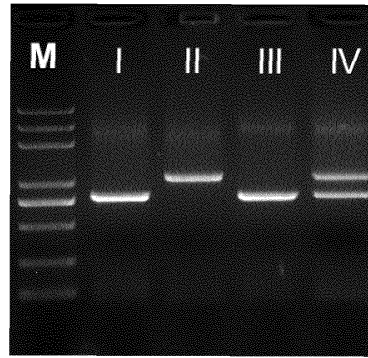
Фиг. 2



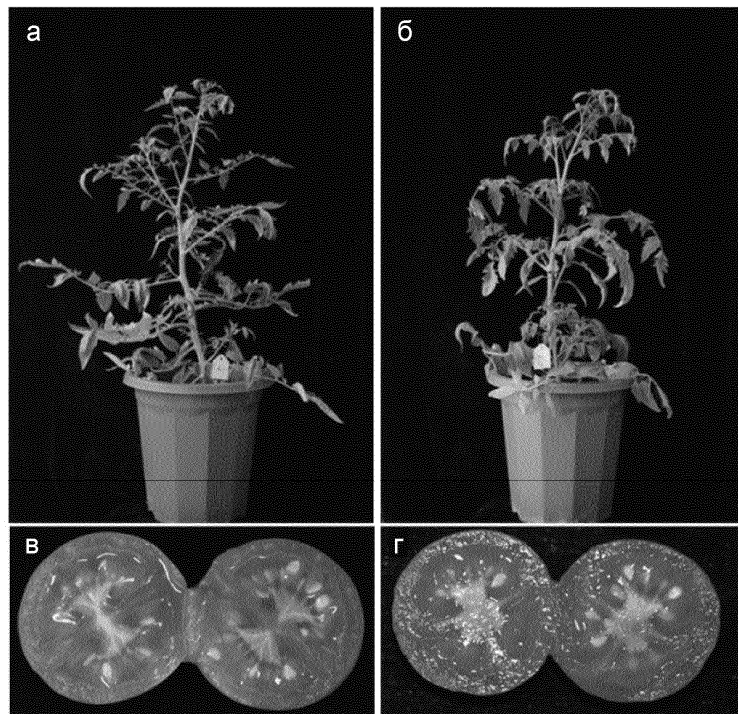
Фиг. 3



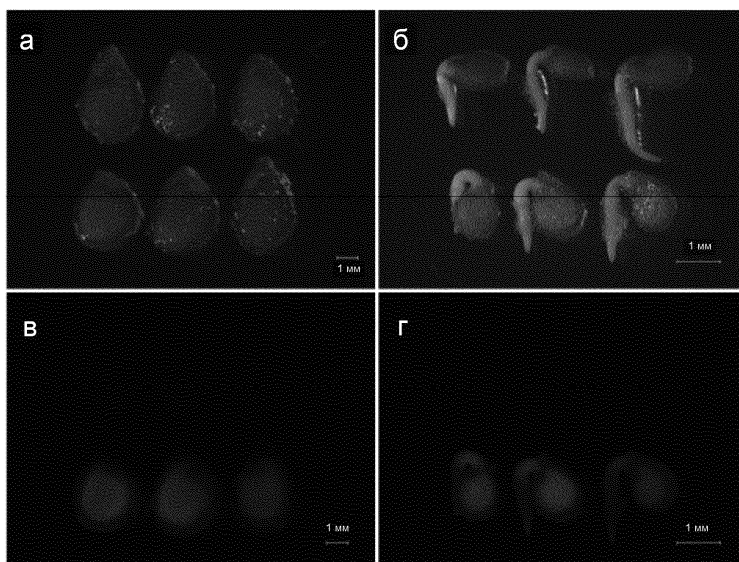
Фиг. 4



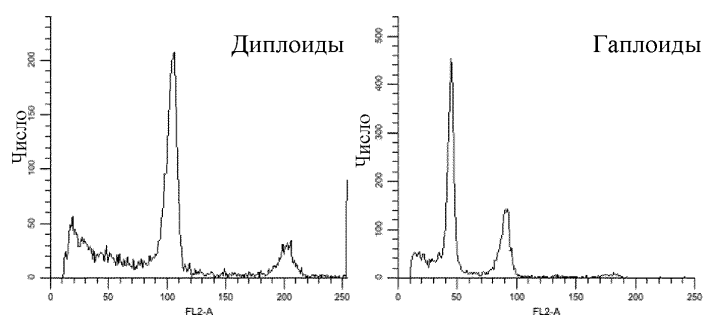
Фиг. 5



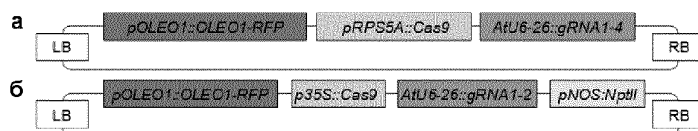
Фиг. 6



Фиг. 7



Фиг. 8



Фиг. 9

