

(19)



Евразийское
патентное
ведомство

(21) 202190980 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2021.08.11

(51) Int. Cl. *C12N 15/63* (2006.01)
A01H 1/00 (2006.01)
A01H 5/00 (2018.01)

(22) Дата подачи заявки
2019.10.10

(54) НОВЫЕ АЛЛЕЛИ CENH3 ПШЕНИЦЫ

(31) PCT/CN2018/110063

(32) 2018.10.12

(33) CN

(86) PCT/CN2019/110404

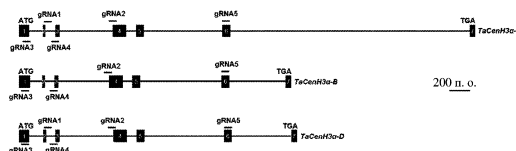
(87) WO 2020/073963 2020.04.16

(71) Заявитель:
СИНГЕНТА КРОП ПРОТЕКШН АГ
(CH)

(72) Изобретатель:
Люй Цзянь, Юй Кунь, Вэй Цзюань,
Лю Чунься, Чжоу Хунцзюй (CN),
Келлихер Тимоти Джозеф (US)

(74) Представитель:
Веселицкий М.Б., Веселицкая И.А.,
Кузенкова Н.В., Каксис Р.А., Белоусов
Ю.В., Куликов А.В., Кузнецова Е.В.,
Соколов Р.А., Кузнецова Т.В. (RU)

(57) Предусмотрены растения пшеницы, содержащие мутацию, вызывающую изменение аминокислотной последовательности центромерного гистона H3 (CENH3), которые обладают биологической активностью индукторов гаплоидии. Кроме того, предусмотрены способы получения растений пшеницы и гаплоидных и удвоенных гаплоидных растений пшеницы, получаемых путем скрещивания растений пшеницы с растениями пшеницы дикого типа.



A1

202190980

202190980

A1

НОВЫЕ АЛЛЕЛИ CENH3 ПШЕНИЦЫ

ОБЛАСТЬ ТЕХНИКИ, К КОТОРОЙ ОТНОСИТСЯ ИЗОБРЕТЕНИЕ

- 5 Настоящее изобретение относится к области сельского хозяйства. В частности, настоящее изобретение относится к белкам CenH3 и полинуклеотидам, их кодирующим, способам получения гаплоидных, а также последующих удвоенных гаплоидных растений и получаемым из них растениям и семенам, в частности, у видов пшеницы.

10 ИСПРАШИВАНИЕ ПРИОРИТЕТА

Настоящая заявка испрашивает приоритет согласно Парижской конвенции по РСТ/CN2018/110063, поданной 12 октября 2018 г., которая включена в данный документ во всей своей полноте.

ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

- 15 К настоящей заявке прилагается перечень последовательностей под названием 81696WOPCT_ST25.txt, созданный 30 сентября 2019 г., который имеет размер примерно 112 килобайтов. Этот перечень последовательностей включен в данный документ посредством ссылки во всей своей полноте. Данный перечень последовательностей подается вместе с настоящей заявкой посредством EFS-Web и
- 20 соответствует §1.824(a)(2)–(6) и (b) раздела 37 C.F.R.

ПРЕДПОСЫЛКИ ИЗОБРЕТЕНИЯ

- Высокая степень гетерозиготности селекционного материала может сделать селекцию растений и отбор полезных признаков очень времязатратным процессом. Широкомасштабный скрининг популяции, даже с использованием самых современных
- 25 инструментов молекулярной селекции, является одновременно и трудоемким, и дорогостоящим. Создание гаплоидных растений с последующим химическим или спонтанным удвоением генома доказало свою эффективность в качестве пути решения проблемы высокой гетерозиготности и ускорения процесса селекции. Такая технология также называется системой получения удвоенных гаплоидов. Применение системы
- 30 получения удвоенных гаплоидов позволило селекционерам достичь гомозиготности во всех локусах за одно поколение с помощью полногеномного удвоения. Это позволяет

фактически исключить необходимость самоопыления или возвратного скрещивания, при которых обычно потребовалось бы по меньшей мере 7 поколений самоопыления или возвратного скрещивания для снижения гетерозиготности до приемлемого уровня.

Гаплоидные растения могут быть получены в соответствии с разными методиками.

- 5 Например, гаплоидные растения в некоторых сельскохозяйственных культурах могут быть получены путем применения способа, называемого культивированием микроспор. Однако данный способ является дорогостоящим, времязатратным и работает не для всех сельскохозяйственных культур. У некоторых видов сельскохозяйственных культур (удвоенные) гаплоидные растения могут быть получены путем партеногенеза
- 10 яйцеклетки или путем элиминации одного из родительских геномов. Однако такие способы не являются оптимальными, так как они работают только у некоторых отдельных видов сельскохозяйственных культур и характеризуются довольно низкими показателями выхода (удвоенных) гаплоидных растений.

- В WO2011/044132 раскрыт способ получения гаплоидных растений, заключающийся в
- 15 инактивации, или изменении, или осуществлении нокаута центромероспецифического белка H3 (CenH3) у растения. На первой стадии способ заключается в элиминации или осуществлении нокаута эндогенного гена CenH3 у растения. На второй стадии в растение вводится кассета экспрессии, кодирующая мутантный или измененный белок CenH3. Мутантный или измененный белок CenH3 образуется путем слияния N-
- 20 концевого домена H3.3, необязательно меченного GFP, с эндогенным доменом гистоновой складки CenH3. Такая методика также известна как "обмен хвостами с GFP" или "обмен хвостами" (обзор которой также приведен в Britt and Kurpu, Front Plant Sci. 2016; 7: 357). Скрещивание растения, несущего такой "обмененный хвост", с растением дикого типа (т. е. имеющим функциональный эндогенный белок CenH3 без
- 25 "обмененного хвоста"), вызывает однородительскую элиминацию генома, что, в свою очередь, приводит к получению гаплоидного растения. Некоторая индукция гаплоидии, хоть и реже встречающаяся, также была обнаружена при добавлении GFP к N-концу эндогенного CenH3 (без "обмененного хвоста"). Однако данная методика не идеальна, так как она является трудоемкой, времязатратной и требует получения трансгенного
- 30 растения.

В WO2014/110274 описан способ получения гаплоидных растений, заключающийся в скрещивании первого растения, экспрессирующего эндогенный ген CenH3, со вторым

- растением, называемым растением, являющимся индуктором гаплоидии, имеющим геном из по меньшей мере двух видов, где большая часть генома получена из первого вида, и геном содержит гетерологичную геномную область из второго вида, где гетерологичная геномная область кодирует полипептид CenH3, отличный от CenH3 первого вида (также описано в Maheshwari et al, PLoS Genet. 2015 Jan 26; 11(1):e1004970). Однако данная методика не является оптимальной, так как она страдает от тех же недостатков, что и вышеупомянутые — она является трудоемкой, времязатратной и требует получения трансгенного растения. Кроме того, данный способ сопряжен с низким выходом гаплоидных растений.
- Другие способы заключаются во введении одной или нескольких точечных мутаций, приводящих к замене одной аминокислоты в С-концевом домене гистоновой складки белка CenH3 или в гене CenH3, кодирующем белок CenH3. О примерах таких мутаций в С-концевом домене гистоновой складки белка CenH3 сообщалось в Karimi-Ashtiyani et al (2015) Proc Natl Acad Sci U S A. 2015 Sep 8;112(36):11211–16; Kuppu, et al. (2015) PLoS Genet. 2015 Sep 9;11(9):e1005494. Однако успех таких способов затмевается тем, что было обнаружено, что некоторые из таких мутаций, но не все из них, являются достаточными для того, чтобы индуцировать однородительскую элиминацию генома после скрещивания с растением дикого типа для получения гаплоидного растения.
- Пшеница (*Triticum aestivum*) является особенно сложным организмом для редактирования или осуществления мутаций ее генов, так как она представляет собой гексаплоидный организм. Эволюционировавшая на протяжении нескольких тысяч лет и подвергшаяся нескольким межвидовым скрещиваниям с предковыми видами пшеницы, *Triticum aestivum* содержит три генома: А (вероятно, от *T. monococcum* или пшеницы-однозернянки), В (вероятно, от *T. searsii*) и D (вероятно, от *T. tauschii*). Каждый геном имеет по 7 хромосом. *Triticum aestivum* имеет две копии каждого из геномов, т. е., AA BB DD; таким образом, она имеет всего 42 хромосомы (6 полных геномов по 7 хромосом в каждом). См. в целом The Evolution of Wheat по адресу www.cerealsdb.uk.net/cerealgenomics/WheatBP/Documents/DOC_Evolution.php, последний доступ 10 июля 2019 г. Кроме того, редактирование или мутация в одной из копий одного гена у *Triticum aestivum* может не проявиться в виде наблюдаемых эффектов, так как дополнительные 5 его копий, вероятно, будут компенсировать изменения в мутантной копии. Для того чтобы на самом деле наблюдать эффект нокаут-мутации, необходимо было бы подвергнуть мутации все 6 копий.

Следовательно, остается весьма труднодостижимым определение того, какая(какие) мутация(мутации) или модификация(модификации) в белке CenH3 или гене *CenH3*, кодирующем белок CenH3, способны индуцировать или являются достаточными для того, чтобы индуцировать однородительскую элиминацию генома для получения гаплоидных растений. Таким образом, в данной области техники остается необходимость в альтернативных или улучшенных способах, обеспечивающих возможность эффективного получения гаплоидных растений (например, менее трудоемких, менее времязатратных, менее дорогостоящих и/или не обязательно требующих получения трансгенного растения), которые впоследствии можно подвергнуть удвоению с получением удвоенных гаплоидных растений. При использовании систем получения удвоенных гаплоидов гомозиготность может быть достигнута за одно поколение.

КРАТКОЕ ОПИСАНИЕ

Для удовлетворения этой потребности один вариант осуществления настоящего изобретения представляет собой растение пшеницы, содержащее по меньшей мере геном А, геном В и геном D, где геном В содержит нокаут-мутацию в гене CENH3, и где необязательно геном D содержит нокаут-мутацию в гене CENH3, и где дополнительно геном А содержит мутантный ген CENH3, содержащий по меньшей мере одну нокдаун-мутацию в 5'-сайте сплайсинга интрона. В одном аспекте нокдаун-мутация представляет собой мутацию по типу восстановления сдвинутой рамки считывания или мутацию по типу большой делеции. В другом варианте осуществления растение пшеницы является гомозиготным по нокаут-мутации в гене CENH3 в геноме В. В альтернативном варианте осуществления растение пшеницы является биаллельным по нокаут-мутации в гене CENH3 в геноме В. В другом варианте осуществления растение пшеницы является гомозиготным по нокаут-мутации в гене CENH3 в геноме D. В альтернативном варианте осуществления растение пшеницы является биаллельным по нокаут-мутации в гене CENH3 в геноме D. В еще одном варианте осуществления растение пшеницы является гомозиготным, биаллельным или характеризуется комбинацией этих свойств по нокаут-мутации в гене CENH3 в геноме В и геноме D. В другом варианте осуществления растение пшеницы является гомозиготным по мутации по типу восстановления сдвинутой рамки считывания в CENH3; или оно является гетерозиготным по мутации по типу восстановления

сдвинутой рамки считывания в CENH3; или оно является биаллельным по мутации по типу восстановления сдвинутой рамки считывания в CENH3.

Другим аспектом настоящего изобретения является способ получения растения пшеницы, индуцирующего гаплоидию, при этом способ включает: (a) получение по меньшей мере клетки растения пшеницы, содержащей по меньшей мере три генома; (b) осуществление мутаций в двух из трех геномов с получением гомозиготных нокаут-мутаций в гене CENH3; (c) осуществление мутаций в третьем геноме с получением гомозиготной нокдаун-мутации в гене CENH3 и (d) получение в результате этого растения пшеницы, содержащего гомозиготные нокаут-мутации в гене CENH3 в двух из трех геномов и дополнительно содержащего гомозиготную нокдаун-мутацию в гене CENH3 в третьем геноме; вследствие чего растение пшеницы, полученное на стадии (d), производит гаплоидное потомство при скрещивании с растением пшеницы дикого типа. В одном варианте осуществления три генома включают в себя геном А, геном В и геном D. В другом варианте осуществления нокаут-мутации в гене CENH3 имеют место в геномах В и D. В еще одном варианте осуществления нокдаун-мутация в гене CENH3 имеет место в геноме А. В одном аспекте нокдаун-мутации в гене CENH3 в геноме А представляют собой мутации по типу восстановления сдвинутой рамки считывания. В другом аспекте мутации по типу восстановления сдвинутой рамки считывания выбраны из группы, состоящей из SEQ ID NO: 56, последовательности нуклеиновой кислоты, на 70% идентичной SEQ ID NO: 56, SEQ ID NO: 63, последовательности нуклеиновой кислоты, на 70% идентичной SEQ ID NO: 63, SEQ ID NO: 69 и последовательности нуклеиновой кислоты, на 70% идентичной SEQ ID NO: 69.

Другим аспектом настоящего изобретения является растение пшеницы, содержащее мутантный ген CENH3, содержащий по меньшей мере одну делеционную мутацию в N-концевом домене, которая приводит к сдвигу рамки считывания, восстановлению сдвинутой рамки считывания или большой делеции. Еще одним аспектом является растение пшеницы, содержащее мутантный ген CENH3, содержащий по меньшей мере одну инсерционную мутацию в N-концевом домене, которая приводит к сдвигу рамки считывания, восстановлению сдвинутой рамки считывания или большой делеции.

Другим аспектом настоящего изобретения является способ получения сконструированного восстановления сдвинутой рамки считывания в гене в клетке,

включающий: (а) приведение генома в контакт с сайт-направленной нуклеазой ("SDN") и по меньшей мере двумя направляющими нуклеиновыми кислотами, где по меньшей мере две направляющие нуклеиновые кислоты нацеливаются на по меньшей мере две последовательности-мишени в гене; (b) предоставление для SDN возможности разрезания гена в по меньшей мере двух последовательностях-мишенях с утратой таким образом вставочной последовательности между по меньшей мере двумя последовательностями-мишенями; и обеспечение возможности прохождения эндогенной репарации ДНК; вследствие чего эндогенная репарация ДНК приводит к получению гена со сконструированным восстановлением сдвинутой рамки считывания.

5 В одном варианте осуществления утрачиваемая вставочная последовательность из стадии (b) содержит (N) пар оснований, где (N) является числом, кратным 3.

Еще одним аспектом настоящего изобретения является способ получения гаплоидного растения пшеницы, включающий: (а) получение растения пшеницы; (b) скрещивание растения пшеницы с растением пшеницы, содержащим мутантный ген CENH3; и (с) отбор потомка, полученного в результате стадии скрещивания; где потомок представляет собой гаплоидное растение пшеницы. В одном варианте осуществления растение пшеницы из стадии (а) представляет собой отцовское родительское растение. В другом варианте осуществления растение пшеницы из стадии (а) представляет собой материнское родительское растение. В другом варианте осуществления способ включает дополнительную стадию превращения растения пшеницы, являющегося потомком, в удвоенное гаплоидное растение пшеницы.

Другим аспектом настоящего изобретения является получение растения пшеницы, содержащего мутантный аллель CENH3, содержащий последовательность нуклеиновой кислоты, на по меньшей мере 70% идентичную последовательности, выбранной из группы, состоящей из SEQ ID NO: 53–73, где мутация представляет собой мутацию по типу восстановления сдвинутой рамки считывания, и где растение пшеницы образует гаплоидное потомство при скрещивании с диплоидным растением пшеницы дикого типа. В одном варианте осуществления растение пшеницы содержит по меньшей мере одну копию мутантного аллеля CENH3; в другом варианте осуществления растение пшеницы содержит по меньшей мере две копии мутантного аллеля CENH3; в еще одном варианте осуществления растение пшеницы содержит по меньшей мере три копии мутантного аллеля CENH3. В одном варианте осуществления мутантный аллель

CENH3 содержит последовательность нуклеиновой кислоты, на 80, 90, 95 или 100% идентичную SEQ ID NO: 53–73.

ТАБЛИЦА 1. КРАТКОЕ ОПИСАНИЕ ПОСЛЕДОВАТЕЛЬНОСТЕЙ В ПЕРЕЧНЕ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

SEQ-ID NO:	ID праймера (или ID зонда)	Местоположение или функция	Краткое описание
1	S2- α 1A	5'-UTR в TaCenH3 α -A	Клонирование геномной последовательности TaCenH3 α -A в сорте Fielder
2	As- α 1A/B/D	Экзон 5 в TaCenH3 α -A/B/D	
3	S- α 1A/B/D	Экзон 5 в TaCenH3 α -A/B/D	
4	As- α 1A	Экзон 7 в TaCenH3 α -A	
5	S- α 1B	5'-UTR в TaCenH3 α -B	Клонирование геномной последовательности TaCenH3 α -B в сорте Fielder
6	As- α 1A/B	Экзон 5 в TaCenH3 α -A/B	
7	S- α 1A/B/D	Экзон 5 в TaCenH3 α -A/B/D	
8	As- α 1B/D	Экзон 7 в TaCenH3 α -B/D	
9	S2- α 1D	5'-UTR в TaCenH3 α -D	Клонирование геномной последовательности TaCenH3 α -D в сорте Fielder
10	As- α 1A/B/D	Экзон 5 в TaCenH3 α -A/B/D	
11	S- α 1A/B/D	Экзон 5 в TaCenH3 α -A/B/D	
12	As- α 1B/D	Экзон 7 в TaCenH3 α -B/D	
13	qRT(A)-S	Смысловый праймер	Экспрессия TaCenH3 α -A
14	qRT(A)-As	Антисмысловый праймер	
15	qRT(A)-зонд	Зонд	
16	qRT(B)-S	Смысловый праймер	Экспрессия TaCenH3 α -B
17	qRT(B)-As	Антисмысловый праймер	
18	qRT(B)-зонд	Зонд	

SEQ ID NO:	ID праймера (или ID зонда)	Местоположение или функция	Краткое описание
19	qRT(B)-S	Смысловой праймер	Экспрессия <i>TaCenH3α-D</i>
20	qRT(B)-As	Антисмысловой праймер	
21	qRT(B)-зонд	Зонд	
22	TQ1115-S	Смысловой праймер	Контрольный анализ для qRT-PCR, нацеленный на фактор АДФ-рибозилирования
23	TQ1115-As	Антисмысловой праймер	
24	TQ1115-зонд	Зонд	
25	gRNA1	sgRNA, нацеливающаяся на границу экзона 2 и интрона 2	sgRNA для <i>TaCenH3α</i>
26	gRNA2	sgRNA, нацеливающаяся на границу интрона 3 и экзона 4	
27	gRNA3	sgRNA, нацеливающаяся на экзон 1	
28	gRNA4	sgRNA, нацеливающаяся на границу интрона 2 и экзона 3	
29	KW2917R	Антисмысловой праймер	Анализ по методу KASP - 2917
30	KW2917F1	Смысловой праймер	
31	KW2917F2	Смысловой праймер	
32	KW11728R	Антисмысловой праймер	Анализ по методу KASP - 11728
33	KW11728F1	Смысловой праймер	
34	KW11728F2	Смысловой праймер	
35	KW11091R	Антисмысловой праймер	Анализ по методу KASP - 11091

SEQ ID NO:	ID праймера (или ID зонда)	Местоположение или функция	Краткое описание
36	KW11091F1	Смысловой праймер	Анализ по методу KASP - 11511
37	KW11091F2	Смысловой праймер	
38	KW11511R	Антисмысловой праймер	
39	KW11511F1	Смысловой праймер	
40	KW11511F2	Смысловой праймер	Анализ по методу KASP - 11129
41	KW11129R	Антисмысловой праймер	
42	KW11129F1	Смысловой праймер	
43	KW11129F2	Смысловой праймер	
44	e35S -S	Смысловой праймер	Проверка числа копий трансгена
45	e35S -As	Антисмысловой праймер	
46	e35S-зонд	Зонд	
47	PMI - S	Смысловой праймер	Проверка числа копий трансгена
48	PMI - As	Антисмысловой праймер	
49	PMI - зонд	Зонд	
50	FA	Смысловой праймер	Праймеры для секвенирования отредактированного <i>TaCENH3α</i> -A
51	R3	Антисмысловой праймер	
52	M13R	Антисмысловой праймер	Праймер для секвенирования клонов
53	M13F	Смысловой праймер	
54	F1	Смысловой праймер	Праймер для RT-PCR <i>TaCENH3α</i>

SEQ ID NO.	ID праймера (или ID зонда)	Местоположение или функция	Краткое описание
55	R1	Антисмысловый праймер	
56	Геномная последовательность A* в A004A	gRNA1, вставка A; gRNA2, вставка A	Геномная последовательность в TaCENH3 α -A
57	CDS-последовательность A* в A004A	Восстановление сдвинутой рамки считывания на N-конце	Восстановление сдвинутой рамки считывания на N-конце в TaCENH3 α -A
58	CDS-последовательность a в A004A	Преждевременный стоп-кодон	Потеря функции в TaCENH3 α -A
59	Последовательность белка A* в A004A	Восстановление сдвинутой рамки считывания на N-конце	Восстановление сдвинутой рамки считывания на N-конце в TaCENH3 α -A
60	Последовательность белка a в A004A	Преждевременный стоп-кодон	Потеря функции в TaCENH3 α -A
61	Геномная последовательность B в A004A	gRNA1, WT; gRNA2, WT	TaCENH3 α -B WT
62	Геномная последовательность d в A004A	gRNA1, WT; gRNA2, ^A	Потеря функции в TaCENH3 α -D

SEQ ID NO:	ID праймера (или ID зонда)	Местоположение или функция	Краткое описание
63	Геномная последовательность A* в C003A	gRNA1, ΔG ; gRNA2, ^A	Проверка числа копий трансгена
64	CDS-последовательность A* в C003A	gRNA1, ΔG ; gRNA2, ^A	Восстановление сдвинутой рамки считывания на N-конце в TaCENH3 α -A
65	Последовательность белка A* в C003A	gRNA1, ΔG ; gRNA2, ^A	Восстановление сдвинутой рамки считывания на N-конце в TaCENH3 α -A
66	Геномная последовательность b в C003A	gRNA1, WT; gRNA2, ^A	Потеря функции в TaCENH3 α -B
67	Последовательность белка b в C003A	gRNA1, WT; gRNA2, ^A	Потеря функции в TaCENH3 α -B
68	Геномная последовательность d в C003A	gRNA1, WT; gRNA2, ^A	Потеря функции в TaCENH3 α -D
69	Геномная последовательность A* в A073A	gRNA3, ^A; gRNA4, ΔG	Восстановление сдвинутой рамки считывания на N-конце в TaCENH3 α -A

SEQ ID NO:	ID праймера (или ID зонда)	Местоположение или функция	Краткое описание
70	CDS-последовательность A* в A073A	gRNA3, ^A; gRNA4, ΔG	Восстановление сдвинутой рамки считывания на N-конце в TaCENH3α-A
71	Последовательность белка A* в A073A	gRNA3, ^A; gRNA4, ΔG	Восстановление сдвинутой рамки считывания на N-конце в TaCENH3α-A
72	Геномная последовательность b в A073A	gRNA3, WT; gRNA4, ^A	Потеря функции в TaCENH3α-B
73	Геномная последовательность d в A073A	gRNA3, ΔGTC; gRNA4, ^A	Потеря функции в TaCENH3α-D
74			Конструкция 24195
75			Конструкция 24194
76			Аминокислотная последовательность, утрачиваемая при RFS
77			Аминокислотная последовательность, добавляемая при RFS

SEQ ID NO:	ID праймера (или ID зонда)	Местоположение или функция	Краткое описание
78			Аминокислотная последовательность, добавляемая при RFS
79	Праймер SQ-1		

КРАТКОЕ ОПИСАНИЕ ФИГУР

На фигуре 1 показана структура гена *TaCenH3a* и относительные местоположения gRNA. Экзоны пронумерованы и представлены закрашенными прямоугольниками. Интроны представлены тонкими линиями. Длина и тех, и других представлена их шириной.

ОПРЕДЕЛЕНИЯ

Настоящее изобретение не ограничено конкретной методикой, протоколами, линиями клеток, видами или родами растений, конструкциями и реагентами, описанными в данном документе. Терминология, используемая в данном документе, представлена исключительно с целью описания конкретных вариантов осуществления, а не предназначена для ограничения объема настоящего изобретения, который будет ограничиваться лишь прилагаемой формулой изобретения. Необходимо отметить, что, как используется в данном документе и в прилагаемой формуле изобретения, формы единственного числа включают ссылки на множественное число, если в контексте явно не указано иное. Таким образом, ссылка на "растение" является ссылкой на одно или несколько растений и включает их эквиваленты, известные специалистам в данной области, и т. д. Используемое в данном документе слово "или" означает любой элемент из конкретного перечня, а также включает любую комбинацию элементов из такого перечня (то есть включает также "и").

Используемый в данном документе термин "приблизительно" означает примерно, ориентировочно, около или в районе. Если термин "приблизительно" используется в сочетании с числовым диапазоном, он модифицирует этот диапазон, расширяя границы выше и ниже указанных числовых значений. В целом, используемый в данном документе термин "приблизительно" модифицирует числовое значение путем его варьирования выше и ниже указанного значения на 20 процентов, предпочтительно на 10 процентов выше или ниже (больше или меньше). Что касается температуры, термин "приблизительно" означает $\pm 1^{\circ}\text{C}$, предпочтительно $\pm 0,5^{\circ}\text{C}$. Если термин "приблизительно" используется в контексте настоящего изобретения (например, в комбинациях со значениями температуры или молекулярной массы), то предпочтительным является точное значение (т. е. без "приблизительно").

Используемый в данном документе термин "амплифицированный" означает конструирование нескольких копий молекулы нуклеиновой кислоты или нескольких

копий, комплементарных молекуле нуклеиновой кислоты, с использованием по меньшей мере одной из молекул нуклеиновых кислот в качестве матрицы. Системы амплификации включают систему полимеразной цепной реакции (ПЦР), систему лигазной цепной реакции (LCR), систему амплификации, основанной на последовательности нуклеиновых кислот (NASBA, Cangene, Миссиссога, Онтарио), систему на основе Q-бета-репликазы, систему амплификации, основанной на транскрипции (TAS), и амплификацию с замещением нитей (SDA). См., например, Diagnostic Molecular Microbiology: Principles and Applications, PERSING et al., Ed., American Society for Microbiology, Washington, D.C. (1993). Продукт амплификации называется "ампликоном".

Термин "биаллельный" относится к паре генов, которая не является ни гомозиготной (AA или aa), ни гетерозиготной (Aa). Вернее говоря, оба гена в паре были отредактированы, но не идентичным образом. Например, пара генов *CenH3* в хромосоме A в настоящем изобретении может содержать одну мутацию по типу RFS в одном аллеле, приводящую к нокауту гена при экспрессии, тогда как другой аллель содержит нокаут-мутацию. Это может быть указано символами как "A*a" и свидетельствует о биаллельной мутации.

Термин "специфичная последовательность ДНК" указывает на полинуклеотидную последовательность, обладающую гомологией нуклеотидной последовательности, составляющей более 80%, предпочтительно более 85%, более предпочтительно более 90%, еще более предпочтительно более 95%, и еще более предпочтительно более 97%, наиболее предпочтительно более 99%, с другой названной последовательностью.

"кДНК" относится к одонитевой или двухнитевой ДНК, которая комплементарна мРНК и получена из нее. Термин "центромероспецифический вариант гистонового белка H3" ("белок CenH3" или просто "CENH3"), используемый в данном документе, относится к белку, который является членом комплекса кинетохора. Белок CenH3 также известен как белок CENP-A. Комплекс кинетохора располагается на хроматидах там, где нити веретена деления прикрепляются в процессе клеточного деления для отделения сестринских хроматид друг от друга. Белки CenH3 принадлежат к хорошо изученному классу белков, которые представляют собой варианты гистоновых белков H3. Такие белки являются существенно важными для правильного образования и функционирования кинетохора и содействуют связыванию кинетохора с ДНК. В

клетках, которые являются дефицитными по CenH3, наблюдается неспособность к локализации белков кинетохора на хроматидах и демонстрируются серьезные дефекты хромосомного расщепления (т. е. все хромосомы растения, экспрессирующего дефицитный белок CenH3, элиминируются или утрачиваются, что приводит к изменению ploидности соматических клеток (например, к редукции числа хромосомных наборов, как, например, от диплоидного к гаплоидному состоянию)). По этой причине белки CenH3 были предметом интенсивного изучения в отношении их потенциального применения в системе получения удвоенных гаплоидов. Белки CenH3 характеризуются наличием варибельного хвостового домена (также называемого "N-концевым доменом" или "N-концевым хвостовым доменом") и консервативного домена гистоновой складки (также называемого "C-концевым доменом"), состоящего из трех альфа-спиральных областей, соединенных петлевыми участками. Домен гистоновой складки CenH3 является относительно высококонсервативным среди белков CenH3 у разных видов. Домен гистоновой складки располагается на карбоксильном конце эндогенного белка CenH3. В отличие от домена гистоновой складки, N-концевой хвостовой домен CenH3 является высоковарибельным даже у близкородственных видов.

Термин "полинуклеотид, кодирующий CenH3, имеющий одну или несколько активных мутаций" относится к неэндогенному или эндогенному мутантному полинуклеотиду, кодирующему CenH3, который кодирует белок CenH3, имеющему одну или несколько активных мутаций, которые в случае наличия в растении при отсутствии его эндогенного полинуклеотида, кодирующего CenH3, и/или эндогенного белка CenH3 позволяют растению быть жизнеспособным и позволяют образовывать гаплоидное потомство или потомство с абберантной ploидностью при скрещивании растения с растением дикого типа. Растение, содержащее полинуклеотид, кодирующий CenH3, имеющий одну или несколько активных мутаций, может называться "модифицированным растением". Процентная доля гаплоидного потомства или потомства с абберантной ploидностью, которое образуется при скрещивании с растением дикого типа, может, например, составлять по меньшей мере 0,1, 0,5, 1, 5, 10, 20 процентов или больше. Мутация, которая вызывает переход из эндогенного полинуклеотида, кодирующего CenH3, в полинуклеотид, кодирующий CenH3, имеющий одну или несколько активных мутаций, в данном документе называется активной мутацией. Активная мутация применительно к белку CenH3 может

приводить, помимо прочего, к сниженной загрузке в центромеры, менее функциональному белку CenH3 и/или сниженной функциональности при разделении хромосом в процессе клеточного деления. Одна или несколько активных мутаций могут быть введены в полинуклеотид, кодирующий CenH3, с помощью любого из
 5 нескольких способов, хорошо известных специалисту в данной области, например, с помощью случайного мутагенеза, как, например, индуцированного обработкой семян или растительных клеток химическими соединениями или радиоактивным излучением, целенаправленного мутагенеза, применения эндонуклеаз, путем получения частичных или полных делеций доменов белков или путем слияния с гетерологичными
 10 последовательностями.

Посредством нокаута или инактивации эндогенного полинуклеотида, кодирующего CenH3, можно добиться отсутствия эндогенного полинуклеотида, кодирующего CenH3, у растения. В качестве альтернативы, эндогенный полинуклеотид, кодирующий CenH3, может быть модифицирован таким образом, чтобы он кодировал неактивный или
 15 нефункциональный белок CenH3.

Модифицированное растение, содержащее полинуклеотид, кодирующий CenH3, имеющий одну или несколько активных мутаций, указанное в данном документе, может быть скрещено с растением дикого типа в качестве родительского растения, образующего пыльцу, либо в качестве родительского растения, образующего
 20 семязачатки. В одном варианте осуществления белок CenH3, имеющий одну или несколько активных мутаций, может содержать 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 15, 20 или больше аминокислотных замен относительно эндогенного белка CenH3. В одном варианте осуществления полинуклеотид, кодирующий CenH3, имеющий одну или несколько активных мутаций, обладает 70-, 75-, 80-, 85-, 90-, 95-, 96-, 97-, 98-, 99-, 99,5-
 25 процентной идентичностью последовательности с эндогенным полинуклеотидом, кодирующим CenH3, предпочтительно по всей длине. Специалист в данной области будет в состоянии легко удостовериться, содержит или не содержит модифицированное растение, указанное в данном документе, одну или несколько активных мутаций. Например, специалист в данной области может воспользоваться прогностическими
 30 инструментами, такими как SIFT (Kumar P, Henikoff S, Ng PC. (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc; 4(7): 1073-81. doi: 10.1038/nprot.2009.86), чтобы сделать предположение о такой активной мутации. Затем в растении могут быть произведены одна или несколько

активных мутаций, и должен произойти нокаут экспрессии эндогенного белка CenH3 в растении. Растение может рассматриваться как содержащее одну или несколько активных мутаций, если процентная доля гаплоидного потомства или потомства с aberrантной ploидностью, которое образуется при скрещивании с растением дикого типа, составляет по меньшей мере 0,1, 0,5, 1, 5, 10, 20 процентов или больше.

Скрещивание растения, у которого отсутствует эндогенный полинуклеотид, кодирующий CenH3, или у которого отсутствует экспрессия эндогенного белка CenH3 и которое экспрессирует белок CenH3, имеющий одну или несколько активных мутаций, в качестве родительского растения, образующего пыльцу, либо родительского растения, образующего семязачатки, с растением дикого типа (т. е. экспрессирующим эндогенный белок CenH3) приводит к получению потомства, которое является гаплоидным или демонстрирует aberrантную ploидность. Такое растение содержит только хромосомы родительского растения, которое экспрессирует эндогенный белок CenH3, и не содержит хромосомы растения, экспрессирующего белок CenH3, имеющий одну или несколько активных мутаций.

Термин "aberrантная ploидность", используемый в данном документе, относится к ситуации, в которой клетка содержит aberrантное или аномальное число наборов хромосом. Например, клетка, имеющая один или три набора хромосом на клетку, тогда как обычное число равно двум, является клеткой с aberrантной ploидностью. В настоящем изобретении активные мутантные белки CenH3 и способы их применения могут применяться для получения мутантных растений с aberrантной ploидностью, например, для получения гаплоидных растений, тогда как немутантное растение является диплоидным. Гаплоидные растения могут использоваться для ускорения селекционных программ по созданию гомозиготных линий и для исключения необходимости в инбридинге.

Термины "химерная конструкция", "химерный ген", "химерный полинуклеотид" или "химерная нуклеиновая кислота" (и аналогичные термины), используемые в данном документе, относятся к конструкции или молекуле, содержащей два или более полинуклеотида разного происхождения, собранных в виде одной молекулы нуклеиновой кислоты. Термины "химерная конструкция", "химерный ген", "химерный полинуклеотид" или "химерная нуклеиновая кислота" относятся к любой конструкции или молекуле, которая содержит (1) полинуклеотиды (например, ДНК), в том числе

регуляторные и кодирующие полинуклеотиды, которые вместе не встречаются в природе (т. е. по меньшей мере один из полинуклеотидов является гетерологичным по отношению к по меньшей мере одному из других ее полинуклеотидов), или (2) полинуклеотиды, кодирующие части белков, которые в естественных условиях не соединены, или (3) части промоторов, которые в естественных условиях не соединены. Кроме того, химерная конструкция, химерный ген, химерный полинуклеотид или химерная нуклеиновая кислота могут содержать регуляторные полинуклеотиды и кодирующие полинуклеотиды, полученные из разных источников, или могут содержать регуляторные полинуклеотиды и кодирующие полинуклеотиды, полученные из одного и того же источника, но расположенные иным способом, чем встречающийся в природе. В предпочтительном аспекте настоящего изобретения химерная конструкция, химерный ген, химерный полинуклеотид или химерная нуклеиновая кислота содержат каскету экспрессии, содержащую полинуклеотиды согласно настоящему изобретению, под контролем регуляторных полинуклеотидов, в частности, под контролем регуляторных полинуклеотидов, функционирующих в растениях.

Термин "хромосома" используется в данном документе как общепризнанный в данной области техники, означающий самовоспроизводящуюся генетическую структуру в клеточном ядре, содержащую клеточную ДНК и несущую линейный набор генов.

"Кодирующий полинуклеотид" представляет собой полинуклеотид, который транскрибируется в РНК, такую как мРНК, рРНК, тРНК, snRNA, смысловая РНК или антисмысловая РНК. РНК затем предпочтительно транслируется в организме с продуцированием белка. Она может представлять собой "непрерывающийся кодирующий полинуклеотид", т. е. не иметь интронов, как, например, в кДНК, или она может содержать один или несколько интронов, ограниченных соответствующими границами сплайсинга. "Инtron" представляет собой поли(рибо)нуклеотид, который содержится в первичном транскрипте, но который удаляется посредством расщепления и повторного лигирования РНК в клетке с получением зрелой мРНК, которая может транслироваться в белок.

Термин "удвоенное гаплоидное растение", используемый в данном документе, относится к генотипу, образуемому в случае, когда гаплоидные клетки подвергаются удвоению числа хромосом. Искусственное получение удвоенных гаплоидов является важным в селекции растений. Удвоенные гаплоиды могут быть получены *in vivo* или *in*

in vitro. Гаплоидные зародыши образуются in vivo в результате партеногенеза, псевдогамии или элиминации хромосом. Известен целый ряд способов получения in vitro удвоенных гаплоидных организмов из гаплоидных организмов. Неограничивающий пример способа получения удвоенных гаплоидов in vitro заключается в обработке соматических гаплоидных клеток, гаплоидных зародышей, гаплоидных семян или гаплоидных растений, полученных из гаплоидных семян, средством для удвоения числа хромосом, таким как колхицин. В настоящем изобретении гомозиготные двойные гаплоидные растения могут быть регенерированы из гаплоидных клеток путем приведения гаплоидных клеток в контакт со средствами для удвоения числа хромосом, такими как колхицин, гербициды, являющиеся ингибиторами образования микротрубочек, или оксид одновалентного азота, для создания гомозиготных удвоенных гаплоидных клеток. Способы удвоения числа хромосом раскрыты, например, в патентах США №№ 5770788, 7135615 и публикациях заявок на патент США №№ 2004/0210959 и 2005/0289673; Antoine-Michard, S. et al., Plant Cell, Tissue Organ Cult., Dordrecht, the Netherlands, Kluwer Academic Publishers 48(3):203-207 (1997); Kato, A., Maize Genetics Cooperation Newsletter 1997, 36-37; а также Wan, Y. et al., Trends Genetics 77: 889-892 (1989), Wan, Y. et al., Trends Genetics 81: 205-211 (1991), раскрытия которых включены в данный документ посредством ссылки. Двойные гаплоидные растения могут быть дополнительно скрещены с другими растениями с получением F1, F2 или последующих поколений растений с желаемыми признаками. Традиционные процедуры инбридинга занимают семь поколений до достижения практически полной гомозиготности, тогда как при использовании удвоенных гаплоидов это достигается за одно поколение.

Термин "E0" относится к первому экземпляру растения, подвергнутого редактированию. Иными словами, растительная клетка, которая подвергается редактированию с помощью, например, CRISPR и которой затем предоставляется возможность развиться до растения, становится растением E0. Растение E1 является потомком E0 (обычно, но не обязательно, полученным путем самооплодотворения), содержащим редактирование. Аналогичным образом, растение E2 является потомком E1 (обычно, но не обязательно, полученным путем самооплодотворения), содержащим редактирование. Растения E3, E4, E5 и т. д. аналогичным образом отдалены в поколениях от растения E0.

Термины "редактирование гена", "редактирование", "редактирование генома", "GE" и т. п. относятся к сайт-специфическим мутациям, произведенным в последовательности-мишени. Это также может называться "целенаправленным мутагенезом". Используемый в данном документе термин "целенаправленный мутагенез" или "стратегия мутагенеза" относится к любому способу мутагенеза, который приводит в результате к преднамеренному мутагенезу выбранного гена. Целенаправленный мутагенез включает способы с использованием CRISPR, TILLING, TALEN и другие способы, которые еще не разработаны, но которые можно применять для достижения такого же результата. Мутагенез можно осуществлять в соответствии с любыми методиками, известными из уровня техники, такими как без ограничения синтез олигонуклеотида, имеющего одну или несколько мутаций в пределах последовательности в конкретной регуляторной последовательности. В частности, сайт-специфический мутагенез представляет собой методику, применимую при получении мутантных промоторов посредством специфического мутагенеза ДНК, лежащей в их основе. Также можно использовать РНК-направляемые эндонуклеазы ("RGEN", например, CRISPR/Cas9). Данная методика дополнительно обеспечивает легкодоступную возможность получения и тестирования вариантов последовательностей, например, с учетом одного или нескольких из вышеуказанных соображений, посредством введения одного или нескольких изменений нуклеотидной последовательности в ДНК. Сайт-специфический мутагенез обеспечивает получение мутантных форм посредством использования специфических олигонуклеотидных последовательностей, которые кодируют последовательность ДНК с требуемой мутацией, а также достаточного количества смежных нуклеотидов для обеспечения получения последовательности праймера, имеющей достаточные размер и сложность последовательности для образования стабильного дуплекса по обе стороны от пересекаемой границы делеции. Как правило, предпочтительным является праймер, длина которого составляет от приблизительно 17 до приблизительно 75 нуклеотидов или больше, содержащий от приблизительно 10 до приблизительно 25 или больше остатков по обе стороны от границы в изменяемой последовательности. См. в целом патент США № 10285348, включенный в данный документ посредством ссылки во всей своей полноте.

Термины "отредактированный N-концевой хвост" или "отредактированный N-концевой домен" используются взаимозаменяемо во всем в данном документе.

Термин "эндогенный", используемый в контексте настоящего изобретения в комбинации с белком или геном, означает, что указанный белок или ген происходит из растения, в котором он по-прежнему содержится. Зачастую эндогенный ген будет присутствовать в растении в его нормальном генетическом окружении. В другом контексте термин "эндогенный" может относиться к нормальным функциям клетки. Для примера и не в качестве ограничения, "эндогенная репарация ДНК" относится к нормальным клеточным механизмам, ферментам или процессам репарации ДНК.

Термин "экспрессия" при использовании в отношении полинуклеотида, такого как ген, ORF или их часть, или трансгена в растениях относится к процессу преобразования генетической информации, закодированной в гене, в РНК (например, мРНК, рРНК, тРНК или snRNA) посредством "транскрипции" гена (т. е. благодаря ферментативному действию РНК-полимеразы) и в белок в соответствующих случаях (например, если ген кодирует белок) посредством "трансляции" мРНК. Экспрессия гена может регулироваться на многих стадиях в ходе этого процесса. Например, в случае антисмысловых конструкций или dsRNA-конструкций соответственно экспрессия может относиться к транскрипции только антисмысловой РНК или только dsRNA. В вариантах осуществления "экспрессия" относится к транскрипции и стабильному накоплению смысловой (мРНК) или функциональной РНК. "Экспрессия" может также относиться к продуцированию белка.

Термины "гомология", "сходство последовательностей" или "идентичность последовательностей" для нуклеотидных или аминокислотных последовательностей означают степень идентичности или сходства двух или более последовательностей и могут быть определены традиционными способами с использованием известного программного обеспечения или компьютерных программ, таких как программы для попарного сравнения BestFit или Gap (GCG Wisconsin Package, Genetics Computer Group, 575 Science Drive, Мэдисон, Висконсин 53711). В BestFit используется алгоритм поиска локальной гомологии из Smith and Waterman, *Advances in Applied Mathematics* 2:482-489 (1981), для обнаружения сегмента с наилучшими идентичностью или сходством между двумя последовательностями. Сравнение последовательностей двух или более полинуклеотидов или полипептидов обычно осуществляют путем сравнения частей двух последовательностей в окне сравнения для выявления и сравнения локальных областей сходства последовательностей. Размер окна сравнения обычно составляет от приблизительно 20 до 200 смежных нуклеотидов. Gap осуществляет

глобальные выравнивания: выравнивание одной полной последовательности с другой полной сходной последовательностью с использованием способа из Needleman and Wunsch, J. Mol. Biol. 48:443-453 (1970). При использовании программы для выравнивания последовательностей, такой как BestFit, для определения степени гомологии, сходства или идентичности последовательностей ДНК можно использовать 5 настройки по умолчанию, или для оптимизации баллов идентичности, сходства или гомологии можно выбрать соответствующую матрицу замен. Аналогично, при использовании программы, такой как BestFit, для определения степени идентичности, сходства или гомологии последовательностей между двумя отдельными 10 аминокислотными последовательностями можно использовать настройки по умолчанию, или для оптимизации баллов идентичности, сходства или гомологии можно выбрать соответствующую матрицу замен, такую как blosum45 или blosum80.

Термин "локус" относится к положению (например, гена, генетического маркера или т. п.) в хромосоме данного вида.

15 Термин "праймер", используемый в данном документе, относится к олигонуклеотиду, который способен отжигаться с мишенью для амплификации, что позволяет прикрепиться ДНК-полимеразе, и тем самым служить в качестве точки начала синтеза ДНК при помещении в условия, в которых индуцируется синтез продукта удлинения праймера, например, в присутствии нуклеотидов и средства для полимеризации, такого 20 как ДНК-полимераза, и при подходящих температуре и pH. Праймер (для амплификации) предпочтительно является односторонним для максимальной эффективности амплификации. Праймер предпочтительно представляет собой олигодезоксирибонуклеотид. Праймер обычно является достаточно длинным, чтобы служить затравкой для синтеза продуктов удлинения в присутствии средства для 25 полимеризации. Точные значения длины праймера будут зависеть от многих факторов, в том числе от температуры и состава (содержания A/T и G/C) праймера. Пара двунаправленных праймеров состоит из одного прямого и одного обратного праймера, обычно используемых в области амплификации ДНК, как, например, при ПЦР-амплификации. Будет понятно, что используемый в данном документе термин 30 "праймер" может относиться более чем к одному праймеру, в частности, в случае, когда существует некоторая неясность в информации о концевой(концевых) последовательности(последовательностях) области-мишени, подлежащей амплификации. Следовательно, "праймер" включает совокупность праймерных

олигонуклеотидов, содержащих последовательности, представляющие возможные вариации в последовательности, или содержит нуклеотиды, обеспечивающие возможность обычного спаривания оснований. Олигонуклеотидные праймеры можно получить любым подходящим способом. Способы получения олигонуклеотидов с конкретной последовательностью известны из уровня техники и включают, например, клонирование и рестрикцию соответствующих последовательностей, а также прямой химический синтез. Способы химического синтеза могут включать, например, фосфодиэфирный или фосфотриэфирный способ, диэтилфосфорамидатный способ и способ с использованием твердых подложек, например, раскрытые в US 4458066. При необходимости праймеры можно пометить путем включения в их состав средств, выявляемых, например, с помощью спектроскопических, флуоресцентных, фотохимических, биохимических, иммунохимических или химических способов. Матрично-зависимое удлинение олигонуклеотидного(олигонуклеотидных) праймера(праймеров) катализируется средством для полимеризации в присутствии достаточных количеств четырех дезоксирибонуклеотидтрифосфатов (dATP, dGTP, dCTP и dTTP; т. е. dNTP) или их аналогов в реакционной среде, которая состоит из подходящих солей, катионов металлов и буферной системы для поддержания pH. Подходящие средства для полимеризации представляют собой ферменты, которые, как известно, катализируют праймер- и матрично-зависимый синтез ДНК. Известные ДНК-полимеразы включают, например, ДНК-полимеразу I из *E. coli* или ее фрагмент Кленова, ДНК-полимеразу T4 и ДНК-полимеразу Taq. Условия реакции для катализа синтеза ДНК с помощью этих ДНК-полимераз известны из уровня техники. Продуктами синтеза являются дуплексные молекулы, состоящие из нитей-матриц и нитей удлинения праймеров, которые содержат последовательность-мишень. Эти продукты, в свою очередь, служат в качестве матрицы для следующего раунда репликации. Во втором раунде репликации нить удлинения праймера из первого цикла отжигается с комплементарным ей праймером; при этом синтез дает "короткий" продукт, который ограничен на обоих из 5'- и 3'-конца праймерными последовательностями или комплементарными им последовательностями. В результате повторяющихся циклов денатурации, отжига и удлинения праймеров происходит экспоненциальное накопление области-мишени, определяемой праймерами. Для достижения требуемого количества полинуклеотида, содержащего область-мишень нуклеиновой кислоты, проводят достаточное число циклов. Требуемое количество может варьировать, и оно определяется функцией, для которой служит

полинуклеотидный продукт. Способ ПЦР хорошо описан в руководствах и известен специалисту в данной области. После амплификации путем ПЦР полинуклеотиды-мишени можно выявлять посредством гибридизации с полинуклеотидным зондом, который образует стабильный гибрид с полинуклеотидом с последовательностью-мишенью в условиях гибридизации и отмывки низкой, умеренной или даже высокой жесткости. Если ожидается, что зонды будут фактически полностью комплементарными (т. е. на приблизительно 99% или больше) последовательности-мишени, то можно применять условия высокой жесткости. Если ожидается некоторое несовпадение, например, если ожидается наличие вариантов штаммов, вследствие чего зонд не будет полностью комплементарен, то жесткость гибридизации может быть снижена. Однако, как правило, подбираются такие условия, которые исключают неспецифическое/случайное связывание. Условия, которые влияют на гибридизацию и которые выбирают для предупреждения неспецифического связывания, известны из уровня техники и описаны, например, в Sambrook and Russell, 2001. В целом, более низкая концентрация солей и более высокая температура повышают жесткость условий гибридизации. "Праймер для ПЦР", как предпочтительно понимают в рамках настоящего изобретения, относится к относительно коротким фрагментам односторонней ДНК, используемым в ПЦР-амплификации специфических областей ДНК.

Термины "белок", "пептид" и "полипептид" используются в данном документе взаимозаменяемо.

Термин "промотор" относится к полинуклеотиду, как правило, расположенному выше (5') от его кодирующего полинуклеотида, который осуществляет контроль экспрессии кодирующего полинуклеотида, обеспечивая узнавание РНК-полимеразой и другими факторами, необходимыми для правильной транскрипции.

Термин "сайт-направленная нуклеаза" относится к любому ферменту, направляемому нуклеотидной последовательностью к последовательности-мишени в цепи ДНК. Сайт-направленная нуклеаза предпочтительно является используемой в CRISPR, но также может представлять собой мегануклеазу, эффекторную нуклеазу, подобную активатору транскрипции (TALEN), или нуклеазу с "цинковыми пальцами". Сайт-направленная(сайт-направленные) нуклеаза(нуклеазы) могут быть обозначены аббревиатурой "SDN". SDN включают без ограничения мегануклеазы (MN), нуклеазы с "цинковыми пальцами" (ZFN), эффекторные нуклеазы, подобные активаторам

транскрипции (TALEN), нуклеазу Cas9, нуклеазу Cpf1 (Cas12a), dCas9-FokI, dCpf1-FokI, химерную Cas9-цитозиндезаминазу, химерную Cas9-адениндезаминазу, химерную FEN1-FokI и Mega-TAL, нуклеазу Cas9 (nC9), химерную нуклеазу dCas9, не содержащую FokI, и нуклеазу Cpf1, не содержащую FokI; и где дополнительно направляющая нуклеиновая кислота представляет собой направляющую РНК.

Термины "жесткие условия" или "жесткие условия гибридизации" включают ссылку на условия, при которых полинуклеотид будет гибридизироваться со своей последовательностью-мишенью в заметно большей степени, чем с другими последовательностями (например, на уровне, в по меньшей мере 2 раза превышающем фоновый). Жесткие условия зависят от последовательности и будут отличаться при разных обстоятельствах. Путем контроля жесткости условий гибридизации и/или отмывки можно идентифицировать полинуклеотиды-мишени, которые на 100% комплементарны зонду (гомологичные зонды). В качестве альтернативы, жесткость условий можно корректировать для обеспечения возможности некоторого несовпадения в последовательностях, чтобы выявлять более низкие степени их подобия (гетерологичные зонды). Как правило, жесткие условия будут представлять собой условия, при которых концентрация солей составляет менее чем примерно 1,5 М ионов Na, как правило, приблизительно 0,01-1,0 М ионов Na (или других солей) при pH 7,0-8,3, а температура составляет по меньшей мере приблизительно 30°C для коротких зондов (например, 10-50 нуклеотидов) и по меньшей мере приблизительно 60°C для длинных зондов (например, более 50 нуклеотидов). Жесткие условия также могут достигаться при добавлении дестабилизирующих средств, таких как формамид. Типичные условия низкой жесткости включают гибридизацию с буферным раствором, содержащим 30-35% формамид, 1 М NaCl, 1% SDS (вес/об.; додецилсульфат натрия) при 37°C и отмывку в 1-2×SSC (20×SSC=3,0 М NaCl/0,3М цитрата натрия) при 50-55°C. Типичные условия умеренной жесткости включают гибридизацию в 40-45% формамиде, 1 М NaCl, 1% SDS при 37°C и отмывку в 0,5-1×SSC при 55-60°C. Типичные условия высокой жесткости включают гибридизацию в 50% формамиде, 1 М NaCl, 1% SDS при 37°C и отмывку в 0,1×SSC при 60-65°C. Специфичность, как правило, зависит от отмывок после гибридизации, при которых ключевыми факторами являются ионная сила и температура конечного раствора для отмывки. Для гибридных молекул ДНК—ДНК Tm можно примерно вычислить из уравнения согласно Meinkoth and Wahl (Anal. Biochem., 138:267-284, 1984): $T_m = 81,5^{\circ}\text{C} + 16,6 (\log M) + 0,41 (\% \text{ GC})$

- 0,61 (% форм.) - 500/L; где М представляет собой молярность одновалентных катионов, % GC представляет собой процентную долю гуанозиновых и цитозиновых нуклеотидов в ДНК, % форм. представляет собой процентную долю формамида в растворе для гибридизации, а L представляет собой длину гибрида в парах оснований.

5 Т_m представляет собой температуру (при заданных ионной силе и рН), при которой 50% комплементарной последовательности-мишени гибридизируется с абсолютно совпадающим зондом. Т_m снижается на приблизительно 1°C на каждый 1% несовпадения, таким образом, Т_m, условия гибридизации и/или отмывки можно регулировать, чтобы обеспечить гибридизацию с последовательностями с требуемой

10 идентичностью. Например, если осуществляют поиск последовательностей с примерно 90% идентичностью, Т_m можно снизить на 10°C. В целом жесткие условия выбирают таким образом, чтобы температура была на приблизительно 5°C ниже температуры точки плавления (Т_m) для специфической последовательности и комплементарной ей последовательности при заданных ионной силе и рН. Однако при условиях сильной

15 жесткости можно использовать гибридизацию и/или отмывку при температуре на 1, 2, 3 или 4°C ниже, чем температура точки плавления (Т_m), при условиях умеренной жесткости можно использовать гибридизацию и/или отмывку при температуре на 6, 7, 8, 9 или 10°C ниже, чем температура точки плавления (Т_m); при условиях низкой жесткости можно использовать гибридизацию и/или отмывку при температуре на 11,

20 12, 13, 14, 15 или 20°C ниже, чем температура точки плавления (Т_m). Рядовые специалисты в данной области будут понимать, что вариации жесткости растворов для гибридизации и/или отмывки по существу описаны с помощью уравнения, составов для гибридизации и отмывки и требуемой Т_m. Если требуемая степень несовпадения приводит к Т_m, составляющей менее 45°C (водный раствор) или 32°C (раствор

25 формамида), предпочтительно повысить концентрацию SSC для того, чтобы можно было использовать более высокую температуру. Подробное руководство по гибридизации нуклеиновых кислот можно найти в Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology — Hybridization with Nucleic Acid Probes, Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe*

30 *assays"*, Elsevier, N.Y. (1993); и в *Current Protocols in Molecular Biology, Chapter 2*, Ausubel, et al., eds., Greene Publishing and Wiley-Interscience, New York (1995). Из уровня техники известны способы жесткой гибридизации, условия которых могут быть рассчитаны с помощью средств, известных из уровня техники. Они раскрыты в Sambrook et al., *Molecular Cloning: A Laboratory Manual, 2nd Ed.*, Cold Spring Harbor

Laboratory Press, 1989, Cold Spring Harbor, N.Y., и в Current Protocols in Molecular Biology, Ausubel et al, eds., John Wiley and Sons, Inc., 2000. Из уровня техники известны способы определения процентной идентичности последовательностей, примером которых является программное обеспечение для компьютерного анализа последовательностей GCG (GCG, Inc, Мэдисон, Висконсин).

Используемый в данном документе термин "восстановление сдвинутой рамки считывания" ("RFS") относится к мутации или серии мутаций в гене, которые, по отдельности или в комбинации, прерывают кодирующую последовательность гена, но не изменяют рамку считывания кодирующей последовательности. Это также может называться "синхронизацией, восстанавливающей рамку считывания". Например, кодирующая последовательность ДНК содержит серию кодонов. Каждый кодон содержит три нуклеотида, и каждый кодон при его транскрипции в РНК кодирует одну аминокислоту после трансляции. Мутация по типу вставки/делеции ("вставка/делеция") одного или двух нуклеотидов в кодирующей последовательности вызывает сдвиг кодирующей рамки считывания ("сдвиг рамки считывания"). Однако вставки или делеции, будь то по отдельности или в комбинации, которые в совокупности имеют место в количестве, кратном трем, восстанавливают кодоны до их первоначальной рамки считывания, даже если кодирующая последовательность сама по себе является измененной. См., например, B.N. Ames and H.J. Whitfield, Jr., *Frameshift Mutagenesis in Salmonella*, COLD SPRING HARB. SYMP. QUANT. BIOL. 31:221–225 (1966). Например, и в пределах объема данного определения, последовательность, которая содержит по меньшей мере две мутации по типу вставки/делеции, представляющие собой делеции, будь то последовательные или нет, и в которой мутации по типу вставки/делеции вызывают восстановление рамки считывания до первоначальной рамки считывания, представляет собой последовательность, содержащую мутацию по типу восстановления сдвинутой рамки считывания. Термин "сконструированное восстановление сдвинутой рамки считывания" также может использоваться для описания мутации по типу RFS, которая была создана с помощью редактирования генома или модификации генома.

Используемый в данном документе термин "большая делеция" ("LD") относится к мутации, которая вызывает утрату нескольких последовательных нуклеотидов. В частности, большая делеция относится к утрате 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 50, 60, 70, 80, 90, или 100 или больше

нуклеотидов. В некоторых вариантах осуществления последовательность, утрачиваемая в результате LD, имеет размер, кратный 3 (т. е. 3, 6, 9, 12, 15, 18, 21, 24, 27, 30 и т. д.) В других вариантах осуществления мутация по типу LD может также иметь место в сочетании с мутацией по типу вставки/делеции где-либо еще в той же последовательности, что тем самым вызывает мутацию по типу восстановления сдвинутой рамки считывания.

В контексте настоящего изобретения использование терминов "дикий тип" или "растение дикого типа" относится к растению, которое не несет мутантный белок или ген CenH3 (т. е. не содержит одну или несколько активных мутаций, указанных в данном документе) и которое эндогенно экспрессирует или продуцирует функциональные гены и белки CenH3.

ПОДРОБНОЕ ОПИСАНИЕ

В данной ситуации авторы настоящего изобретения индуцировали альтернативный сплайсинг у пшеницы (*Triticum aestivum*) путем применения CRISPR-Cas9 для редактирования последовательностей цис-сплайсинга, содержащих 5'- и 3'-сайты сплайсинга. Авторы настоящего изобретения выбрали пшеницу в качестве модельного организма-мишени, поскольку пшеница является гексаплоидной, что обеспечивает функциональную избыточность генома пшеницы. В качестве гена-мишени авторы настоящего изобретения выбрали ген центромерного гистона 3 ("*CenH3*"), кодирующий центромерный белок, поскольку модификации данного гена должны приводить к получению растений, обладающих ценностью для селекции сельскохозяйственных культур. *CenH3* отвечает за точное расщепление хромосом в процессе клеточного деления. В отличие от H3 и других традиционных гистонов, *CENH3* имеет длинный, гипервариабельный N-концевой "хвост". См. J. Monen, et al. *Separase Cleaves the N-Tail of the CENP-A Related Protein CPAR-1 at the Meiosis I Metaphase-Anaphase Transition in C. elegans*, PLOS ONE 10:e0125382 (2015). Направленная или естественная модификация "хвоста" запускает компенсаторные изменения в кинетохоре, которые могут обеспечить для *CENH3* возможность управления видообразованием посредством нарушений мейоза или ингибирования зиготического расщепления хромосом. См. I. Lermontova, et al., *Knockdown of CENH3 in Arabidopsis reduces mitotic divisions and causes sterility by disturbed meiotic chromosome segregation*, PLANT J 68:40-50 (2011), и M. Ravi and R. Bondada, *Genome Elimination by Tailswap CenH3: In Vivo Haploid Production in Arabidopsis thaliana*, METHODS MOL BIOL 1469:77-99 (2016). Обмен N-концевого

"хвоста" и "хвоста" H3 приводил к индукции гаплоидии у *Arabidopsis* (M. Ravi and S. Chan, *Haploid plants produced by centromere-mediated genome elimination*, NATURE 464:615–618 (2010)) и кукурузы (T. Kelliher, et al., *Maternal Haploids are Preferentially Induced by CENH3-tailswap Transgenic Complementation in Maize*, FRONT. PLANT SCI., doi.org/10.3389/fpls.2016.00414, 31(7):414 (2016)). Индукция гаплоидии представляет собой aberrantный репродуктивный процесс, который приводит к уменьшению ploидности от одного поколения к следующему. Гаплоиды можно подвергнуть удвоению для получения инбредных линий, исключая необходимость в шести поколениях самоопыления, обычно требуемых для получения новых чистосортных линий. Реализация подхода с "обменом хвостами" у сельскохозяйственных культур требует многих поколений для осуществления сборки нокаутного нативного аллеля и стабильной вставки трансгенов. Авторы настоящего изобретения смогли индуцировать AS путем непосредственного редактирования N-концевых последовательностей *CenH3* пшеницы. Такие новые последовательности *CenH3* были изучены для того, чтобы определить, могут ли мутантные белки CENH3 вызывать индукцию гаплоидии у пшеницы и в какой комбинации они это делают. В описанных авторами настоящего изобретения обстоятельствах это действительно происходит.

Следовательно, один вариант осуществления настоящего изобретения представляет собой растение пшеницы, содержащее по меньшей мере геном А, геном В и геном D, где геном В содержит нокаут-мутацию в гене CENH3, и где необязательно геном D содержит нокаут-мутацию в гене CENH3, и где дополнительно геном А содержит мутантный ген CENH3, содержащий по меньшей мере одну нокдаун-мутацию в 5'-сайте сплайсинга интрона. В одном аспекте нокдаун-мутация представляет собой мутацию по типу восстановления сдвинутой рамки считывания или мутацию по типу большой делеции. В другом варианте осуществления растение пшеницы является гомозиготным по нокаут-мутации в гене CENH3 в геноме В. В альтернативном варианте осуществления растение пшеницы является биаллельным по нокаут-мутации в гене CENH3 в геноме В. В другом варианте осуществления растение пшеницы является гомозиготным по нокаут-мутации в гене CENH3 в геноме D. В альтернативном варианте осуществления растение пшеницы является биаллельным по нокаут-мутации в гене CENH3 в геноме D. В еще одном варианте осуществления растение пшеницы является гомозиготным, биаллельным или характеризуется комбинацией этих свойств по нокаут-мутации в гене CENH3 в геноме В и геноме D. В

другом варианте осуществления растение пшеницы является гомозиготным по мутации по типу восстановления сдвинутой рамки считывания в CENH3; или оно является гетерозиготным по мутации по типу восстановления сдвинутой рамки считывания в CENH3; или оно является биаллельным по мутации по типу восстановления сдвинутой рамки считывания в CENH3.

Другим аспектом настоящего изобретения является способ получения растения пшеницы, индуцирующего гаплоидию, при этом способ включает: (a) получение по меньшей мере клетки растения пшеницы, содержащей по меньшей мере три генома; (b) осуществление мутаций в двух из трех геномов с получением гомозиготных нокаут-мутаций в гене CENH3; (c) осуществление мутаций в третьем геноме с получением гомозиготной нокдаун-мутации в гене CENH3 и (d) получение в результате этого растения пшеницы, содержащего гомозиготные нокаут-мутации в гене CENH3 в двух из трех геномов и дополнительно содержащего гомозиготную нокдаун-мутацию в гене CENH3 в третьем геноме; вследствие чего растение пшеницы, полученное на стадии (d), производит гаплоидное потомство при скрещивании с растением пшеницы дикого типа. В одном варианте осуществления три генома включают в себя геном A, геном B и геном D. В другом варианте осуществления нокаут-мутации в гене CENH3 имеют место в геномах B и D. В еще одном варианте осуществления нокдаун-мутация в гене CENH3 имеет место в геноме A. В одном аспекте нокдаун-мутации в гене CENH3 в геноме A представляют собой мутации по типу восстановления сдвинутой рамки считывания. В другом аспекте мутации по типу восстановления сдвинутой рамки считывания выбраны из группы, состоящей из SEQ ID NO: 56, последовательности нуклеиновой кислоты, на 70% идентичной SEQ ID NO: 56, SEQ ID NO: 63, последовательности нуклеиновой кислоты, на 70% идентичной SEQ ID NO: 63, SEQ ID NO: 69 и последовательности нуклеиновой кислоты, на 70% идентичной SEQ ID NO: 69.

Другим аспектом настоящего изобретения является растение пшеницы, содержащее мутантный ген CENH3, содержащий по меньшей мере одну делеционную мутацию в N-концевом домене, которая приводит к сдвигу рамки считывания, восстановлению сдвинутой рамки считывания или большой делеции. Еще одним аспектом является растение пшеницы, содержащее мутантный ген CENH3, содержащий по меньшей мере одну инсерционную мутацию в N-концевом домене, которая приводит к сдвигу рамки считывания, восстановлению сдвинутой рамки считывания или большой делеции.

Другим аспектом настоящего изобретения является способ получения сконструированного восстановления сдвинутой рамки считывания в гене в клетке, включающий: (а) приведение генома в контакт с сайт-направленной нуклеазой ("SDN") и по меньшей мере двумя направляющими нуклеиновыми кислотами, где по меньшей мере две направляющие нуклеиновые кислоты нацеливаются на по меньшей мере две последовательности-мишени в гене; (b) предоставление для SDN возможности разрезания гена в по меньшей мере двух последовательностях-мишенях с утратой таким образом вставочной последовательности между по меньшей мере двумя последовательностями-мишенями; и обеспечение возможности прохождения эндогенной репарации ДНК; вследствие чего эндогенная репарация ДНК приводит к получению гена со сконструированным восстановлением сдвинутой рамки считывания. В одном варианте осуществления утрачиваемая вставочная последовательность из стадии (b) содержит (N) пар оснований, где (N) является числом, кратным 3.

Еще одним аспектом настоящего изобретения является способ получения гаплоидного растения пшеницы, включающий: (а) получение растения пшеницы; (b) скрещивание растения пшеницы с растением пшеницы, содержащим мутантный ген CENH3; и (с) отбор потомка, полученного в результате стадии скрещивания; где потомок представляет собой гаплоидное растение пшеницы. В одном варианте осуществления растение пшеницы из стадии (а) представляет собой отцовское родительское растение. В другом варианте осуществления растение пшеницы из стадии (а) представляет собой материнское родительское растение. В другом варианте осуществления способ включает дополнительную стадию превращения растения пшеницы, являющегося потомком, в удвоенное гаплоидное растение пшеницы.

Другим аспектом настоящего изобретения является получение растения пшеницы, содержащего мутантный аллель CENH3, содержащий последовательность нуклеиновой кислоты, на по меньшей мере 70% идентичную последовательности, выбранной из группы, состоящей из SEQ ID NO: 56–73, где мутация представляет собой мутацию по типу восстановления сдвинутой рамки считывания, и где растение пшеницы образует гаплоидное потомство при скрещивании с диплоидным растением пшеницы дикого типа. В одном варианте осуществления растение пшеницы содержит по меньшей мере одну копию мутантного аллеля CENH3; в другом варианте осуществления растение пшеницы содержит по меньшей мере две копии мутантного аллеля CENH3; в еще одном варианте осуществления растение пшеницы содержит по меньшей мере три

копии мутантного аллеля *CENH3*. В одном варианте осуществления мутантный аллель *CENH3* содержит последовательность нуклеиновой кислоты, на 80, 90, 95 или 100% идентичную SEQ ID NO: 56–73.

ПРИМЕРЫ

- 5 Пример 1. Теория, лежащая в основе применения двух направляющих РНК для N-концов

Трансгены *CENH3* с "обменными хвостами" при гетерологичной экспрессии в линии, в которой нативные гены *CENH3* являются нокаутными, приводят к индукции гаплоидии. См., например, публикацию заявки на патент США № 2019/0136250, включенную в данный документ посредством ссылки. Это называется подходом с "обменом хвостами". Важно, что у линий, являющихся индукторами гаплоидии, с "обменными хвостами" отсутствуют аллели дикого типа. Предполагается, что трансгены обладают частичной функциональностью и способны образовывать центромеры, которые являются достаточно стабильными для получения нормально развивающихся растений в гомозиготном состоянии. Однако, если трансгены с "обменными хвостами" являются гетерологичными по отношению к *CENH3* дикого типа в клетке, трансгены с "обменными хвостами" являются нестабильными и приводят к успешной индукции гаплоидов в процессе ауткроссинга. При такой разработке растений с "обменными хвостами" чрезвычайно важно, чтобы нативные гены *CENH3* были нокаутными и чтобы трансгены с "обменными хвостами" имели значительные изменения в N-концевом домене в сочетании с лишь незначительными либо предпочтительно отсутствующими изменениями в С-концевом домене. Индукции гаплоидии не произойдет, даже если мутантные гены *CENH3* будут кодировать белки *CENH3*, которые сохраняют нормальную или близкую к нормальной функциональность.

Для того чтобы достичь индукции гаплоидии у пшеницы, авторы настоящего изобретения непосредственно отредактировали шесть генов *CENH3a* для нокаута нескольких копий и создания модификаций в N-концевом домене (оставляя интактным С-концевой домен) в еще нескольких других копиях. Исходя из экспериментов авторов настоящего изобретения по измерению экспрессии генов для генов *CENH3a* из геномов А, В и D, авторы настоящего изобретения сосредоточились на создании N-концевых модификаций в геноме А и нокаут-мутаций в геномах В и D. В случае, если

редактирования, произведенные авторами настоящего изобретения, были успешными, то не останется ни одной нормальной (интактной) копии *CENH3α*: все гены будут отредактированы, но результаты редактирования будут различаться. Важно, что план редактирования, разработанный авторами настоящего изобретения, не включал какие-либо трансгены *CENH3* — авторы настоящего изобретения всего лишь хотели путем непосредственного редактирования создать частично функциональный вариант *CENH3α* с N-концевой модификацией в геноме A.

Как правило, значительные изменения в областях, кодирующих белки, могут быть достигнуты путем редактирования генома с помощью SDN II CRISPR (также называемого заменой аллеля ("AR") или гомологичной рекомбинацией ("HR")), но эффективность этой технологии у растений является крайне низкой и у пшеницы достигается очень редко. Следовательно, авторы настоящего изобретения разработали стратегию редактирования с использованием двух направляющих РНК ("gRNA"), которые обладали способностью модифицировать *CENH3α* для создания аллеля с частичной потерей функции, который имел значительное изменение (изменение более чем 5 аминокислот) в N-концевом домене и нативную (неизмененную) последовательность С-концевого домена. Это будет требовать специфичного разрезания в обоих сайтах для направляющей РНК. Авторы настоящего изобретения знали, что выбранные направляющие РНК также будут редактировать копии *CENH3α* в геномах В и D, и это было преднамеренным. На самом деле из-за непредсказуемости каждого конкретного результата редактирования для двух направляющих РНК авторы настоящего изобретения ожидали, что большая часть отредактированных аллелей будет аллелями с полной потерей функции (в копиях в геномах A, В и D) вследствие сдвигов рамок считывания в кодирующих последовательностях, которые приводили бы к появлению преждевременных стоп-кодонов, обуславливающих усечение белкового продукта *CENH3α*, и таким образом, были бы полностью нокаутными формами нативного гена. Однако авторы настоящего изобретения также знали, что если бы все копии гена *CENH3α* были нокаутными, растения погибли бы, поскольку частичное функционирование *CENH3α* необходимо для развития растения. Если бы те же направляющие последовательности, которые приводили к созданию модифицированного аллеля, являющегося индуктором гаплоидии, в геноме A, могли одновременно (в одних и тех же растениях) осуществлять нокаут аллелей в геномах В и D, это помогло бы авторам настоящего изобретения: результатом была бы идеально

сконструированная линия, являющаяся индуктором гаплоидии. Таким образом, целью авторов настоящего изобретения было использование двух gRNA для имитации трансгенной системы с "обменными хвостами" путем непосредственного редактирования, но ключевой стадией настоящего изобретения было то, что вместо выполнения какой-либо замены аллеля или вставки трансгена *CENH3α* авторы настоящего изобретения получили новые модифицированные варианты посредством небольших мутаций по типу вставки/делеции, создаваемых с помощью разрезания нуклеазой сайтов для двух направляющих РНК в N-концевом домене.

Однако не каждая пара возможных направляющих последовательностей обладала способностью совмещать в себе создание измененного N-концевого домена вместе с функциональным С-концевым доменом: не каждая пара направляющих последовательностей могла производить эти измененные копии "отредактированного N-концевого хвоста" с частичной функциональностью и способностью к индукции гаплоидии. После проверки того, к чему могут приводить редактирования в отношении аминокислотной последовательности, выяснилось, что многие из направляющих последовательностей будут приводить к появлению преждевременных стоп-кодонов. Другими словами, авторы настоящего изобретения поняли, что они должны специально отбирать направляющие РНК, для которых авторы настоящего изобретения предсказали, что они могут производить такую комбинацию редактирований в сайтах-мишенях, которая приводила бы к образованию аминокислотных последовательностей в зрелом белковом продукте *CENH3α*, содержащих существенные изменения в N-концевом домене, но остающихся нетронутыми в С-концевом домене. В частности, авторы настоящего изобретения планировали провести скрининг растений и идентифицировать те из них, которые содержали такие продуктивные аллели с N-концевыми модификациями для копий *CENH3α* в геноме А и которые также имели нокаутные аллели для копий *CENH3α* в геномах В и D. Зная, что редактирование, опосредованное сайт-направленной нуклеазой, не всегда происходит непосредственно в процессе трансформации, авторы настоящего изобретения заключили, что в поколении E0 они могут создавать нокаутные аллели (с полной потерей функции) для некоторых копий, а также некоторое количество аллелей с частичной потерей функции (индукторов гаплоидии) для еще нескольких других копий гена *CENH3α*, и что такие материалы могут быть способны производить гаплоиды путем самоопыления—при условии, что мужские и женские гаметы (спермии и яйцеклетки) имеют разные

редактирования или разные комбинации аллелей и, следовательно, потенциально разные центромеры. Другими словами, авторы настоящего изобретения полагали, что они могут обнаружить некоторое количество гаплоидов среди растений поколения E1. Авторы настоящего изобретения надеялись, что эти наблюдения позволят направить их на отбор определенных растений E1 для того, чтобы генотипировать (т. е. генотипировать диплоидных сибсов из популяций, в которых появилось некоторое количество гаплоидов в результате самоопыления) и идентифицировать растения E1, которые имели аллели с частичной потерей функции в гомозиготном состоянии. Если бы авторы настоящего изобретения могли так сделать, то затем они бы использовали эти конкретные растения для тестирования частоты индукции гаплоидии посредством ауткроссинга. Подводя итог, авторы настоящего изобретения полагали, что после создания ими линий со стабильными мутациями они могли бы протестировать то, являются ли они действительно индукторами гаплоидии, но авторы настоящего изобретения знали, что они не могли сделать это в первом поколении, поскольку редактирование может не быть "полным" на тот момент времени, и что им было необходимо тестировать и повторно тестировать индукцию гаплоидии для различных генотипов в поколениях E1 и E2.

Имелось множество направляющих РНК, которые авторы настоящего изобретения могли выбрать из тех, для которых имелись подходящие РАМ-сайты, но только несколько отборных пар могли создавать желаемые аллели, являющиеся индукторами гаплоидии. Что касается желаемых аллелей, то одним из способов, посредством которого пара направляющих РНК могла бы создавать такие аллели, является редактирование первого сайта-мишени для gRNA с образованием сдвига рамки считывания, восстановление которого затем происходит посредством компенсаторной мутации во втором сайте-мишени, которая приводит к возвращению транскрипта в правильную (нативную) рамку считывания. Это будет приводить к изменению аминокислотной последовательности N-концевого домена по вставочной нити из аминокислот между двумя сайтами-мишенями для направляющей РНК и будет обеспечивать восстановление "нативной" последовательности C-концевого домена. Такие аллели авторы настоящего изобретения решили называть аллелями с восстановлением сдвинутой рамки считывания ("RFS"). В качестве альтернативы, одновременное или практически одновременное разрезание в обоих сайтах-мишенях для направляющей РНК может приводить к делеции вставочной последовательности

нуклеиновой кислоты. Во многих случаях эта делеция будет приводить к сдвигу рамки считывания в нижерасположенной последовательности, но в некоторых случаях может произойти так, что при такой делеции 3'-последовательность транскрипта остается в нормальной рамке считывания, так что значительная часть аминокислотной последовательности N-концевого домена отсутствует в образующемся в результате белковом продукте, но С-концевой домен опять-таки остается интактным. Авторы настоящего изобретения решили называть такие аллели аллелями с большой делецией ("LD"). Наконец, в некоторых случаях авторы настоящего изобретения разработали gRNA для нацеливания на границы сайтов сплайсинга, и редактирование в этих сайтах-мишенях может приводить к формированию паттернов альтернативного сплайсинга (например, оно может приводить к удержанию интрона или пропуску экзона). Такие аллели, подвергнутые альтернативному сплайсингу ("AS"), в большинстве случаев будут приводить к появлению преждевременных стоп-кодонов и генетическим нокаутам, но авторы настоящего изобретения также обнаружили одну направляющую РНК, которая в случае, если происходит правильное редактирование и правильный сплайсинг, не будет приводить к появлению каких-либо преждевременных стоп-кодонов в рамке считывания. Вместо этого она может приводить к большой вставке в зрелый транскрипт, что приводит к значительным изменениям в N-концевом домене путем вставки длинного аминокислотного фрагмента, но затем специфичное редактирование во втором сайте-мишени может приводить к возвращению последовательности в рамку считывания для С-концевого домена. Другими словами, благодаря продуманной разработке gRNA может быть предсказана мутация в сайте сплайсинга, приводящая к образованию вариантов зрелых мРНК-транскриптов, в которых изменен аминокислотный код N-концевого домена, но восстановлены нормальная рамка считывания и последовательность С-концевого домена.

В следующих примерах авторы настоящего изобретения подробно описывают обнаруженные ими конкретные направляющие РНК и редактирования, а также комбинации редактирований в конкретных растениях. Авторы настоящего изобретения подробно демонстрируют то, как они создали линии пшеницы, являющиеся индукторами гаплоидии, путем непосредственного редактирования нативных генов *CENH3* с использованием двух направляющих РНК, нацеливающих на гипервариабельный N-концевой домен. В соответствии с запланированным растением, имеющие мутации в обоих сайтах-мишенях, в некоторых случаях продуцировали

белковые продукты, которые содержали значительные изменения в аминокислотной последовательности N-концевого домена, не затрагивающие аминокислотную последовательность С-концевого домена. Авторы настоящего изобретения отбирали и поддерживали линии, подвергнутые редактированию, которые имели такие типы измененных аллелей *CENH3* с отредактированным N-концевым "хвостом", и удостоверились в том, что другие копии генов *CenH3* (из геномов А, В и D) были подвергнуты нокауту посредством мутаций, произведенных теми же самыми направляющими РНК. Авторы настоящего изобретения обнаружили и протестировали индукцию гаплоидии в линиях, которые имели желаемые мутации, включающие комбинацию RFS-аллелей из генома А с нокаутными аллелями из геномов В и D. В тех же самых линиях правильная комбинация редактирований, которая согласно предсказанию авторов настоящего изобретения приводила к образованию гаплоидов, действительно приводила к индукции гаплоидии.

Пример 2. Определение последовательностей gRNA для редактирования генов *CENH3α* в геноме Fielder

У гексаплоидной пшеницы имеется два гена *CenH3* — *TaCen3α* и *TaCenH3β*. Копии обоих генов из геномов А, В и D были клонированы в сорте пшеницы "Fielder" с праймерами, разработанными для геномной последовательности сорта "Chinese Spring_v2". Последовательности приведены в (SEQ ID NO: 1–12). Предыдущие исследования показали, что вирус-индуцированный сайленсинг гена ("VIGS") *TaCenH3α* приводил к карликовости и недоразвитости корневой системы, тогда как сайленсинг *TaCenH3β* приводил к снижению степени завязывания семян (Yuan et al., New Phytol. 206(2):839-51. 2015). Как и в случае с большинством генов пшеницы, специфические паттерны экспрессии и функции каждого из гомологов из геномов А, В и D являются недостаточно хорошо изученными. Для редактирования генома авторы настоящего изобретения решили модифицировать *TaCenH3α* у яровой пшеницы сорта Fielder, полагая, что мутации в данном гене не должны оказывать такое значительное влияние на завязывание семян, как мутации в *TaCenH3β*. Специфичные для гомологов qPCR-анализы TaqMan применяли для сбора информации об уровне экспрессии *TaCENH3α*-А, -В и -D (SEQ ID NO: 13–21) в репродуктивных тканях (пыльце, завязи и пыльнике), а также ткани молодых листьев. *TaCenH3α*-А и -В экспрессировались на высоких уровнях в пыльниках, пыльце и завязях, тогда как экспрессионный транскрипт *TaCenH3α*-D практически отсутствовал (таблица 2). В листе *TaCenH3α*-А был

преобладающим транскриптом, что может указывать на то, что потеря функции данного гена способствует фенотипу карликовости после сайленсинга *TaCenH3α*.

Таблица 2. Относительная экспрессия *TaCenH3α-A*, *TaCenH3α-B* и *TaCenH3α-D*

Относительная экспрессия (среднее значение $\pm$ SD)			
Ткань	<i>TaCenH3α-A</i>	<i>TaCenH3α-B</i>	<i>TaCenH3α-D</i>
Лист	1,88 $\pm$ 0,36	0,00 $\pm$ 0,00	0,01 $\pm$ 0,00
Пыльник	44,08 $\pm$ 19,62	46,66 $\pm$ 18,83	0,47 $\pm$ 0,12
Пыльца	43,62 $\pm$ 23,69	54,21 $\pm$ 17,28	4,03 $\pm$ 1,38
Завязь	22,39 $\pm$ 11,50	15,54 $\pm$ 2,73	0,19 $\pm$ 0,05

5

Направляющие РНК подбирались с использованием РАМ-сайтов NGG и согласно сделанным авторами настоящего изобретения предсказаниям типов аминокислотных последовательностей, которые будут приводить к получению белкового продукта CENH3, если определенные редактирования будут произведены в этих сайтах-мишенях. Важно, что большая часть направляющих РНК, которые рассматривались авторами настоящего изобретения, не была способна производить комбинации редактирований, которые приводили бы к получению RFS- или AS-аллелей. Авторы настоящего изобретения сосредоточились на отборе немногих пар gRNA, которые предположительно могли бы это сделать. Направляющая РНК 1 (ACGTCGGCGACACCGGTGCG; SEQ ID NO: 25) (подчеркнут примерный сайт разрезания с образованием двухнитевого разрыва, индуцированного комплексом CRISPR-Cas9) локализуется в области границы экзона 2 и интрона 2. Данная gRNA1 находилась под управлением промотора TaU6. Направляющая РНК 2 (CTTGTGGGAGCAGGGGCAAC; SEQ ID NO: 26) нацеливается на область, расположенную сразу после границы интрона 3 и экзона 4, под управлением TaU6. Направляющая РНК 2 в большинстве случаев не будет редактировать 3'-акцепторный сайт сплайсинга интрона 3. Был сделан выбор в пользу использования двух направляющих РНК, так что авторы настоящего изобретения могли производить значительные изменения, например, RFS, LD или AS-аллели, в N-концевом домене, оставляя в то же время С-концевой домен в рамке считывания. Например, в некоторых растениях и отредактированных аллелях обе gRNA будут производить разрезание

10

15

20

25

одновременно, приводя к делеции вставочной последовательности. В некоторых случаях в ходе происходящей в результате репарации образуется сдвиг рамки считывания, который приведет к нокауту белка. В других случаях будет образовываться укороченный LD-транскрипт, у которого отсутствуют интрон 2, весь экзон 3 и часть экзона 4, с удалением примерно аминокислотной последовательности RAGRAAAPGGAQGA (SEQ ID NO: 76) из белкового продукта, что составляет значительное изменение N-концевого домена.

В качестве альтернативы, одновременное разрезание в обоих сайтах может приводить к образованию сдвига рамки считывания в сайте для gRNA1 (например, любой мутации по типу вставки/делеции, которая встречается в кодирующей последовательности и имеет размер, не кратный 3), восстановление которого происходит в сайте для gRNA2 посредством дополняющей мутации по типу вставки/делеции, что таким образом приводит к возвращению кодирующей последовательности в ее нормальную рамку считывания. Например, делеция размером 1 нуклеотид ("нт") в сайте для gRNA1 и вставка размером 1 нт в сайте для gRNA2 восстанавливают рамку считывания, что приводит к образованию RFS-аллеля. Данный аллель, по всей вероятности, не будет аллелем с потерей функции при предположении о том, что во вставочной последовательности со сдвигом рамки считывания не образуются стоп-кодона. На основании оценки авторами настоящего изобретения потенциальных изменений можно видеть, что по меньшей мере одна комбинация редактирований будет приводить к образованию RFS между сайтами для gRNA1 и gRNA2, который не имеет какого-либо стоп-кодона во вставочной последовательности. Таким образом, авторы настоящего изобретения могут предсказать, какие пары небольших мутаций по типу вставки/делеции в сайтах для gRNA1 и gRNA2 будут в комбинации давать функциональный RFS-аллель.

Две направляющие РНК могут приводить к образованию аллелей, подвергнутых альтернативному сплайсингу ("AS"), которые также могут выступать в качестве RFS-аллелей или аллелей с большой вставкой. Направляющая РНК 1 производит разрезание между G и T (SEQ ID NO: 25; подчеркнуто выше). Это 5'-донорный сайт сплайсинга на конце экзона 2. AS-аллели могут образовываться в случае модификации GT таким образом, что сплайсинг интрона 2 происходит ненадлежащим образом, что приводит к удержанию интрона 2 в кодирующей последовательности. При трансляции рибосома будет осуществлять сквозное прочтение этого интрона и создавать новую вставку

из 44–47 аминокислот в зависимости от природы мутаций по типу вставки/делеции в сайтах для gRNA1 и gRNA2. Эта новая вставка может быть предсказана путем прочитывания новой кодирующей рамки считывания после учета фактора мутаций по типу вставки/делеции и трансляции новой зрелой мРНК. Например, если gRNA1 и gRNA2 создают вставки одиночного нуклеотида А в обоих сайтах *TaCENH3α-A*, транскрипт может быть подвергнут альтернативному сплайсингу, что приводит к вставке аминокислотной последовательности "VARDLPGSLPFRFVLFSVFWSDLLVTCSTECRGEPGGRRPQGGLKGQ" (SEQ ID NO: 77) с удалением WT-последовательности "RRAGRAAAPGGAQGA" (SEQ ID NO: 76) из экзона 3 перед восстановлением нормальной последовательности с помощью мутации, опосредованной gRNA. Аналогично, можно подобным образом предсказать, что другая мутация в сайте для gRNA1 (например, делеция GTG) в комбинации с мутацией в сайте для gRNA2 (делецией С) вызывают альтернативный сплайсинг с образованием новой вставки последовательности "GTFPGRFLFVSSCFLFFGLTCSSPVRRNAEASRAGGGPRGGSRG" (SEQ ID NO: 78) и удалением нативной последовательности "RAGRAAAPGGAQGA" (SEQ ID NO: 76). Авторы настоящего изобретения также могут предсказать, что другие комбинации мутаций в сайтах для gRNA1 и gRNA2 будут приводить к образованию сдвигов рамки считывания, которые не подвергаются возвращению в рамку считывания, что приводит к появлению аллелей с потерей функции. Подобным образом, аллели, подвергнутые альтернативному сплайсингу, индуцированному в других сайтах N-концевой области CenH3, будут неспособны производить модифицированные последовательности мРНК с восстановленной до нормального состояния аминокислотной последовательностью С-концевого домена, поскольку там будут присутствовать стоп-кодоны, образовавшиеся в удержанных интронах или после пропуска одного или нескольких экзонов. Следовательно, gRNA могут быть специально отобраны по их способности производить большие изменения в N-концевом домене, оставляя при этом С-концевой домен нетронутым и нормально транслируемым.

Пример 3. Разработка конструкций и трансформация растений

После клонирования специфических последовательностей *TaCENH3α-A*, *-B* и *-D* в сорте Fielder с праймерами, разработанными для геномной последовательности сорта Chinese Spring, авторы настоящего изобретения отобрали gRNA в соответствии с РАМ-сайтами. gRNA1 (ACGTCGGCGACACCGGTGCG; SEQ ID NO: 25) локализуется в

области границы экзона 2 и интрона 2 (фигура 1). gRNA2 (CTTGTGGGAGCAGGGGCAAC; SEQ ID NO: 26) нацеливается на область, расположенную сразу после границы интрона 3 и экзона 4. Ген SpCas9 был кодон-оптимизированным для пшеницы с помощью двух NLS на обоих концах и находился под управлением промотора и двух энхансеров Ubi сахарного тростника. Кассеты gRNA, содержащие промотор U6 пшеницы и каркасные последовательности gRNA, были синтезированы в GenScript (www.genscript.com) и клонированы в бинарный вектор — конструкцию 24194 (SEQ ID NO: 74).

Для трансформации использовали Fielder — инбредный сорт яровой пшеницы. Незрелые зародыши диаметром приблизительно 2,0-2,5 мм собирали, стерилизовали 70 % этанолом в течение 1 мин и 1% гипохлоритом натрия в течение 10 мин. После стерилизации незрелые зародыши выделяли с помощью скальпеля и лабораторного шпателя в небольшую пробирку и центрифугировали при $20000 \times g$ при 4°C в течение 10 мин в среде для инокуляции. Выделенные зародыши инфицировали *Agrobacterium* в течение 5 мин, затем переносили в среду для совместного культивирования при 23°C в темноте в течение 2 дней. Зародышевую ось вырезали из незрелых зародышей перед переносом в среду для покоя, культивировали при 25°C в темноте в течение 5 дней, затем переносили в селективную среду, содержащую маннозу в концентрации 15 г/л. См. Y. Ishida, et al., *Wheat (Triticum aestivum L.) Transformation Using Immature Embryos*, METHODS IN MOLECULAR BIOLOGY 1223: 189-198 (2015). Через 4 недели интенсивно растущие каллюсы переносили в среду для регенерации для образования зеленых растений. Выжившие растения проходили контроль по методу TaqMan, с помощью которого анализировали присутствия или отсутствия сегментов ДНК из трансгенной вставки ДНК; среди них в теплицу были отправлены только растения, положительные согласно анализам по методу TaqMan в отношении 35S и PMI.

Пример 4. Секвенирование у растений E0, подвергнутых редактированию

Геномную ДНК выделяли из молодых листьев растений E0, положительных согласно TaqMan. Секвенирование осуществляли с помощью высокоточной ДНК-полимеразы, а именно KOD-Plus-Neo (поставщик: TOYOBO Life Science). Использовали аллель-специфические праймеры для *TaCenH3a-A* (FA, SEQ ID NO: 50; R3, SEQ ID NO: 51). ПЦР осуществляли следующим образом: 5 мин при 95°C ; 35 циклов по 30 с при 95°C , 30 с при 65°C , 1 мин при 68°C ; 10 мин при 68°C . Реакционная смесь для ПЦР содержит

11,5 мкл дистиллированной воды, 2,5 мкл 10X ПЦР-буфера для KOD-Plus-Neo, 1 мкл 2 mM dNTP, 1 мкл 25 mM MgSO₄, 1 мкл ДНК-полимеразы KOD-Plus-Neo, 1 мкл прямого праймера FA (10 мкМ), 1 мкл обратного праймера R3 (10 мкМ) и 1 мкл геномной ДНК. Продукты ПЦР секвенировали напрямую с помощью праймера SQ-1 (SEQ ID NO: 79) или клонировали в клонирующий вектор pEASY-Blunt Zero (Transgen Biotech). M13R (SEQ ID NO: 52) и M13F (SEQ ID NO: 53) использовали для секвенирования по методу молекулярных колоний.

Пример 5. Трансформант пшеницы A004A

Трансформант пшеницы A004A демонстрировал индукцию гаплоидии. Трансформант пшеницы № A004A представляет собой одно из сотен трансгенных растений, которые получали посредством трансформации конструкцией 24194. Анализ по методу TaqMan с последующим прямым секвенированием указывал на то, что генотип по генам *TaCENH3α* представлял собой AA*BBdd на стадии прорастания E0. В данном случае заглавная буква указывает на аллель *TaCENH3α* дикого типа без редактирования, буква нижнего регистра указывает на аллель с потерей функции, и заглавная буква со звездочкой (*) указывает на аллель с восстановлением сдвинутой рамки считывания (RFS), большой делецией (LD) или аллель, подвергнутый альтернативному сплайсингу (AS), что означает, что аллель является предполагаемым индуктором гаплоидии. Аллель A* растения A004A содержит вставку аденина в сайте-мишени для gRNA1 и другую вставку аденина в сайте-мишени для gRNA2 (SEQ ID NO: 56). Вставка аденина в сайте для gRNA1 фактически находится в интроне на 3 п. о. ниже конца экзона 2 и сразу после 5'-границы сплайсинга. Сама по себе она не нарушает кодирующую последовательность, но она может в некоторых случаях изменять паттерн сплайсинга. Вставка аденина в сайте для gRNA2 находится в экзоне 3 и приводит к сдвигу рамки считывания кодирующей последовательности. Предсказание паттерна сплайсинга, индуцированного вставкой аденина в сайт-мишень для gRNA1, указывает на то, что это может быть AS-аллель, который демонстрирует удержание интрона (IR) для интрона № 2, поскольку наличие дополнительного аденина после 5'-донорного сайта сплайсинга может изменять инициацию удаления интрона, запуская альтернативный сплайсинг. Альтернативный сплайсинг в данном случае будет приводить к вставке многих аминокислот в экзон 3. При наличии альтернативного сплайсинга анализ результатов указывает на то, что последовательности в экзоне 3 будут находиться вне рамки считывания до тех пор, пока редактирование с помощью gRNA2, являющееся еще

одной вставкой аденина, не восстановит нормальную рамку считывания и аминокислотный код для С-концевого домена.

Чтобы удостовериться в том, что аллель A* у A004A подвергнут альтернативному сплайсингу и представляет собой аллель, являющийся предполагаемым индуктором гаплоидии, авторы настоящего изобретения изучили последовательности зрелой мРНК гена *CENH3α-A* в листе растения A004A E0 на ювенильной стадии. Общую РНК из молодых листьев экстрагировали с использованием TRIzol от INVITROGEN, следуя инструкциям производителя. кДНК синтезировали из 1 мг общей РНК с помощью системы для синтеза первой нити SuperScript III (Invitrogen) с праймером олиго-dT. KOD-Plus-Neo (TOYOBO) использовали для амплификации транскриптов *TaCenH3α-A* с праймерами F1 (SEQ ID NO: 54) и R1 (SEQ ID NO: 55). ПЦР осуществляли в соответствии с инструкциями производителя и следующим образом: 5 мин при 95°C; 35 циклов по 30 с при 95°C, 20 с при 62°C, 20 с при 68°C, 35 циклов; 10 мин при 68°C. Реакционная смесь для ПЦР содержит 11,5 мкл дистиллированной воды, 2,5 мкл 10X ПЦР-буфера для KOD-Plus-Neo, 1 мкл 2 mM dNTP, 1 мкл 25 mM MgSO₄, 1 мкл KOD-Plus-Neo, 1 мкл праймера F1 (10 мкМ), 1 мкл праймера R1 (10 мкМ) и 1 мкл кДНК. Продукт ПЦР очищали с помощью набора для очистки продуктов ПЦР GeneJET (Thermo Scientific) и клонировали в клонирующий вектор pEASY-Blunt Zero (Transgen Biotech). Праймеры M13R и M13F использовали для секвенирования по методу молекулярных колоний. Несколько клонов на один продукт ПЦР секвенировали и анализировали с использованием программного обеспечения Vector NTI (Invitrogen). Относительную экспрессию сплайс-вариантов рассчитывали по числу клонов. Анализ ПЦР-секвенирования по методу молекулярных колоний указывал на то, что мРНК *TaCENH3α-A* в A004A имеет два транскрипта, что свидетельствует об альтернативном сплайсинге. Один из транскриптов (SEQ ID NO: 58), обнаруживаемый в 8 из 18 (44%) колоний, был подвергнут сплайсингу с использованием канонического 5'-сайта сплайсинга. Для этих транскриптов нормальный сплайсинг интрона 2 означает, что редактирование с помощью gRNA1 не влияло на трансляцию аминокислот со зрелой мРНК; однако редактирование с помощью gRNA2 вызывало сдвиг рамки считывания. Таким образом, в этом случае зрелые мРНК, подвергнутые конститутивному сплайсингу, фактически представляют собой нокаутные транскрипты. С другой стороны, 10 из 18 (56%) колоний имели зрелые мРНК-транскрипты с удержанным интроном 2 (SEQ ID NO: 57), что приводило к появлению аллеля с N-концевым RFS,

который содержал 47 новых аминокислот — так, 32 аминокислоты были вставлены из интрона 2, а также 15 аминокислот были изменены вследствие сдвига рамки считывания в экзоне 3. Данная вставка 32 а.к. и изменение в целом 47 а.к. существенно изменяет N-концевой домен белков, полученных в результате трансляции мРНК с RFS.

5 Важно, что последовательность зрелой мРНК, которую авторы настоящего изобретения получили в AS-аллеле *TaCENH3α-A** у A004A, является точным результатом прогнозирования авторами настоящего изобретения последовательности зрелой мРНК, который должен был произойти. Растение A004A E0 поддерживали в течение цветения. Авторы настоящего изобретения не наблюдали каких-либо аномальных фенотипов, и
10 семена E1 (после самоопыления) были собраны с колосьев, образованных данным растением.

Проверяли уровень пloidности семян-потомков E1, полученных при самоопылении растения A004A. Семена высаживали, и из проросших всходов отбирали образцы и анализировали в отношении содержания ДНК с помощью проточной цитометрии.

15 Гаплоиды получали в первой партии высеянных растений-потомков: частота индукции гаплоидии ("HIR") составляла 3,8%. Во второй партии всходов также наблюдались гаплоиды, и HIR составляла 4,2%. Гаплоиды пшеницы, подобно гаплоидам риса и кукурузы, имеют меньший размер, чем соответствующие им диплоиды. Важно, что эти растения либо не имели нокаута обеих копий аллеля "B", либо не имели
20 восстановления сдвинутой рамки считывания, индуцированного мутацией в сайте-мишени для gRNA2. Это позволяет предположить, что только правильная комбинация редактирований в обоих сайтах-мишенях для gRNA1 и gRNA2 в гене *CENH3α-A* вкупе с нокаутом обеих копий аллеля "B" является достаточной для запуска индукции гаплоидии.

25 В первой партии авторы настоящего изобретения наблюдали два растения, являющихся проростками-близнецами; оба проростка были гаплоидами согласно проверке с помощью проточной цитометрии. Это указывает на то, что в одном семени имеется два гаплоидных зародыша. Образование зародышей-близнецов может быть вызвано нарушением развития семязачатков, которое отчасти может запускаться
30 редактированиями в *CENH3α*, хотя для подтверждения этого необходимо больше экспериментальной работы.

Хотя авторы настоящего изобретения были немного удивлены, что они наблюдали гаплоидов после самоопыления (поскольку обычно гаплоиды индуцируются только при модификациях *CENH3* в процессе ауткроссинга), непрерывная возможность редактирования в растениях E0 означает, что мужские и женские половые клетки могут унаследовать разные последовательности (редактирования) и, таким образом, обладать разными по сравнению друг с другом свойствами связывания с центромерами и строением кинетохора, что приводит к гаплоидии после самоопыления.

Пример 6. Трансформант пшеницы C003A

Растение C003A подвергнуто редактированию таким образом, что генотип *TaCenH3α* представляет собой A*abdd на стадии растения E0. A* вводится путем делеции гуанина в сайте для gRNA1 и вставки аденина в сайте для gRNA2. На уровне белка наблюдается отличие в одиннадцать аминокислот в N-концевом домене по сравнению с последовательностью дикого типа. Под SEQ ID NO: 63–65 показаны соответственно геномная последовательность *CenH3* A*, CDS-последовательность A* и последовательность белка A* для C003A.

Семена E1 получали путем самоопыления у растения C003A E0. Растения E1 с генотипом A*A*abdd выращивали в теплице для определения их способности к индукции гаплоидии при ауткроссинге. Растение дикого типа (тестерная линия 03S0352-22) было выбрано в качестве донора пыльцы. У C003A E1 вручную удаляли незрелые пестики, и его опыляли пыльцой от растений дикого типа. Гаплоиды выявляли с помощью маркеров SNP (SEQ ID NO: 29-43), по которым можно определить различия между Fielder и 03S0352-22, впоследствии подтверждаемые проверкой с помощью проточной цитометрии. Среди 208 растений F1 авторы настоящего изобретения получили один гаплоид. Это демонстрировало наличие только отцовских генотипов по четырем маркерам и вместе с тем материнского генотипа по одному маркеру (KW11091).

Пример 7. Трансформант пшеницы A073A

Растение A073A имело генотип AA*B*abdd по *TaCENH3α* на стадии прорастания E0. Аллель A* в A073A имел вставку аденина, вызванную gRNA3, и делецию гуанина, вызванную gRNA4, в геномной ДНК. Это приводит к запуску восстановления сдвинутой рамки считывания на уровне белка с отличием в тридцать одну

аминокислоту между вариантом дикого типа и отредактированным вариантом N-концевого домена. Данное растение было отмечено в качестве потенциального предшественника линий, являющихся индукторами гаплоидии, благодаря его способности к образованию потомства, которое имело генотип A^*A^*bbdd —
 5 идеальную генетическую комбинацию для запуска индукции гаплоидии. Под SEQ ID NO: 69-71 показаны соответственно геномная последовательность CenH3 A^* , CDS-последовательность A^* и последовательность белка A^* для A073A.

Семена E1 получали путем обеспечения возможности самоопыления у растения A073A E0. Растения E1 подвергали секвенированию, и тех, которые имели генотипическую комбинацию AA^*bbdd , отбирали для дальнейшего выращивания в теплице для
 10 определения способности к индукции гаплоидии при ауткроссинге. Используя тестерную линию 03S0352-22, выбранную в качестве донора пыльцы, у растений E1, подвергнутых редактированию, вручную удаляли несозревшие пестики, и их опыляли ручным способом. Гаплоиды выявляли с помощью маркеров SNP, по которым можно
 15 провести генотипические различия между Fielder и 03S0352-22. Предполагаемые гаплоиды, идентифицированные по гомозиготности по этим маркерам, затем были подтверждены путем проверки общего содержания ДНК с помощью проточной цитометрии. Среди 57 растений F1 53 были преимущественно гетерозиготными по маркерам SNP и были диплоидными согласно проверке плоидности, что указывает на
 20 то, что они были гибридами. В противоположность этому, четыре растения имели только отцовские геномные маркеры SNP и были гаплоидами согласно данным проточной цитометрии, что было эквивалентно частоте индукции гаплоидии при ауткроссинге, составляющей 7%.

Растения E1 с генотипом AA^*bbdd также давали гаплоиды E2 в процессе самоопыления
 25 (таблица 3). Авторы настоящего изобретения наблюдали один гаплоид среди 13 растений E2 (частота индукции гаплоидии 7,7%). В то же время авторы настоящего изобретения наблюдали несколько растений с частичной элиминацией хромосом.

Таблица 3. Уровень плоидности растений E2, полученных из A073A в результате самоопыления

ID растения F1	Плоидность
001-11	$1n + X$ (анеуплоидия)

ID растения F1	Плоидность
001-13	1n + X (анеуплоидия)
001-14	1n + X (анеуплоидия)
001-17	1n + X (анеуплоидия)
001-18	1n + X (анеуплоидия)
001-19	2n (диплоид)
001-22	1n + X (анеуплоидия)
001-23	1n + X (анеуплоидия)
001-24	2n (диплоид)
001-26	1n + X (анеуплоидия)
001-27	2n (диплоид)
001-28	1n (гаплоид)
001-30	2n (диплоид)

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Растение пшеницы, содержащее по меньшей мере геном А, геном В и геном D, где геном В содержит нокаут-мутацию в гене CENH3, и где необязательно геном D
5 содержит нокаут-мутацию в гене CENH3, и где дополнительно геном А содержит мутантный ген CENH3, содержащий по меньшей мере одну нокдаун-мутацию в 5'-сайте сплайсинга интрона.
2. Растение пшеницы по п. 1, где нокдаун-мутация представляет собой мутацию по типу восстановления сдвинутой рамки считывания или мутацию по типу большой
10 делеции.
3. Растение пшеницы по п. 1, где растение пшеницы является гомозиготным по нокаут-мутации в гене CENH3 в геноме В.
4. Растение пшеницы по п. 1, где растение пшеницы является биаллельным по нокаут-мутации в гене CENH3 в геноме В.
- 15 5. Растение пшеницы по п. 1, где растение пшеницы является гомозиготным по нокаут-мутации в гене CENH3 в геноме D.
6. Растение пшеницы по п. 1, где растение пшеницы является биаллельным по нокаут-мутации в гене CENH3 в геноме D.
7. Растение пшеницы по п. 1, где растение пшеницы является гомозиготным,
20 биаллельным или характеризуется комбинацией этих свойств по нокаут-мутации в гене CENH3 в геноме В и геноме D.
8. Растение пшеницы по п. 2, где растение пшеницы является гомозиготным по мутации по типу восстановления сдвинутой рамки считывания в CENH3.
9. Растение пшеницы по п. 2, где растение пшеницы является гетерозиготным по
25 мутации по типу восстановления сдвинутой рамки считывания в CENH3.
10. Растение пшеницы по пп. 1–9, где растение пшеницы является гомозиготным по нокдаун-мутации в гене CENH3 в геноме А и гомозиготным по нокаут-мутации в гене CENH3 в геноме В и геноме D.

11. Способ получения растения пшеницы, индуцирующего гаплоидию, при этом способ включает:

- a. получение по меньшей мере клетки растения пшеницы, содержащей по меньшей мере три генома;
 - 5 b. осуществление мутаций в двух из трех геномов с получением гомозиготных нокаут-мутаций в гене CENH3;
 - c. осуществление мутаций в третьем геноме с получением гомозиготной нокдаун-мутации в гене CENH3 и
 - 10 d. получение в результате этого растения пшеницы, содержащего гомозиготные нокаут-мутации в гене CENH3 в двух из трех геномов и дополнительно содержащего гомозиготную нокдаун-мутацию в гене CENH3 в третьем геноме;
- вследствие чего растение пшеницы, полученное на стадии (d), производит гаплоидное потомство при скрещивании с растением пшеницы дикого типа.

15 12. Способ по п. 11, где три генома включают в себя геном A, геном B и геном D.

13. Способ по п. 11, где нокаут-мутации в гене CENH3 имеют место в геномах B и D.

14. Способ по п. 11, где нокдаун-мутация в гене CENH3 имеет место в геноме A.

20 15. Способ по п. 14, где нокдаун-мутации в гене CENH3 представляют собой мутации по типу восстановления сдвинутой рамки считывания.

16. Способ по п. 15, где мутации по типу восстановления сдвинутой рамки считывания выбраны из группы, состоящей из SEQ ID NO: 56, последовательности нуклеиновой кислоты, на 70% идентичной SEQ ID NO: 56, SEQ ID NO: 63, последовательности нуклеиновой кислоты, на 70% идентичной SEQ ID NO: 63, SEQ ID NO: 69 и последовательности нуклеиновой кислоты, на 70% идентичной SEQ ID NO: 69.

25

17. Растение пшеницы, содержащее мутантный ген CENH3, содержащий по меньшей мере одну делеционную мутацию в N-концевом домене, которая приводит к сдвигу рамки считывания, восстановлению сдвинутой рамки считывания или большой делеции.

30

18. Растение пшеницы, содержащее мутантный ген CENH3, содержащий по меньшей мере одну инсерционную мутацию в N-концевом домене, которая приводит к сдвигу рамки считывания, восстановлению сдвинутой рамки считывания или большой делеции.

5 19. Способ получения сконструированного восстановления сдвинутой рамки считывания в гене в клетке, включающий:

10 a. приведение генома в контакт с сайт-направленной нуклеазой ("SDN") и по меньшей мере двумя направляющими нуклеиновыми кислотами, где по меньшей мере две направляющие нуклеиновые кислоты нацеливаются на по меньшей мере две последовательности-мишени в гене;

 b. предоставление для SDN возможности разрезания гена в по меньшей мере двух последовательностях-мишенях с утратой таким образом вставочной последовательности между по меньшей мере двумя последовательностями-мишенями и

15 c. обеспечение возможности прохождения эндогенной репарации ДНК; вследствие чего эндогенная репарация ДНК приводит к получению гена со сконструированным восстановлением сдвинутой рамки считывания.

20. Способ по п. 19, где утрачиваемая вставочная последовательность из стадии (b) содержит (N) пар оснований, где (N) является числом, кратным 3.

20 21. Способ получения гаплоидного растения пшеницы, включающий:

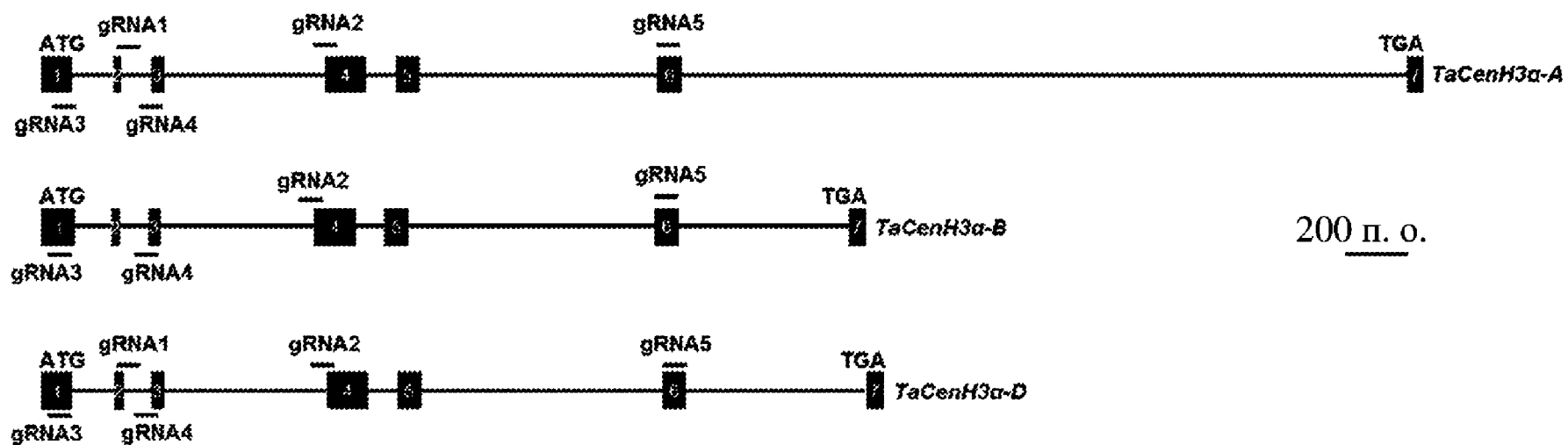
- a. получение растения пшеницы;
- b. скрещивание растения пшеницы с растением пшеницы по пп. 1–10 и пп. 17–18 и
- c. отбор потомка, полученного на стадии скрещивания;

25 где потомок представляет собой гаплоидное растение пшеницы.

22. Способ по п. 20, где растение пшеницы из стадии (a) представляет собой отцовское родительское растение.

23. Способ по п. 20, где растение пшеницы из стадии (a) представляет собой материнское родительское растение.

24. Способ по п. 20, дополнительно включающий превращение растения пшеницы, являющегося потомком, в удвоенное гаплоидное растение пшеницы.
25. Растение пшеницы, содержащее мутантный аллель CENH3, содержащий последовательность нуклеиновой кислоты, на по меньшей мере 70% идентичную
5 последовательности, выбранной из группы, состоящей из SEQ ID NO: 56–73, где мутация представляет собой мутацию по типу восстановления сдвинутой рамки считывания, и где растение пшеницы образует гаплоидное потомство при скрещивании с диплоидным растением пшеницы дикого типа.
26. Растение пшеницы по п. 25, где растение пшеницы содержит по меньшей мере
10 одну копию мутантного аллеля CENH3.
27. Растение пшеницы по п. 25, где растение пшеницы содержит по меньшей мере две копии мутантного аллеля CENH3.
28. Растение пшеницы по п. 25, где растение пшеницы содержит по меньшей мере три копии мутантного аллеля CENH3.
- 15 29. Растение пшеницы по п. 25, где мутантный аллель CENH3 содержит последовательность нуклеиновой кислоты, на 80, 90, 95 или 100% идентичную SEQ ID NO: 56–73.



ФИГ. 1