

(19)



Евразийское  
патентное  
ведомство

(21) 202091671 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки  
2020.11.13

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

(22) Дата подачи заявки  
2019.01.11

(54) ГЕН, ЛЕЖАЩИЙ В ОСНОВЕ QTL ДЛЯ КОЛИЧЕСТВА КОЛОСКОВ НА КОЛОС НА ХРОМОСОМЕ 7A В ПШЕНИЦЕ

(31) 18151432.4; 18151436.5

(32) 2018.01.12

(33) EP

(86) PCT/EP2019/050711

(87) WO 2019/138083 2019.07.18

(71) Заявитель:

БАСФ СЕ (DE); КОММОНВЕЛС  
САЙЕНТИФИК ЭНД ИНДАСТРИАЛ  
РИСЕРЧ ОРГАНИЗЕЙШН (AU)

(72) Изобретатель:

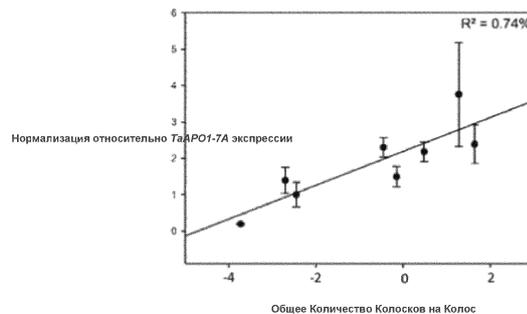
Давей Марк (BE), Каванаг Колин  
Роберт (AU), Ариядаза Рувини (BE),  
Бовилль Виллиэм, Барреро Занчес  
Джозе, Фербила Клара, Шприггс  
Эндру (AU), Ханна Мэсью, Ванг Си  
(BE), Виттерн Лукас, Гарднер Кейт,  
Вебб Элекс Аранделль (GB)

(74) Представитель:

Беляева Е.Н. (BY)

(57) Настоящее изобретение относится к области сельского хозяйства. В частности, изобретением представляется белок, нуклеиновая кислота, рекомбинантный ген, растения, содержащие рекомбинантный ген, а также способы изменения количества колосков на колос растения пшеницы.

ТаАРО1-7А экспрессия против Общего Количества Колосков



202091671  
A1

202091671  
A1

## **ГЕН, ЛЕЖАЩИЙ В ОСНОВЕ QTL ДЛЯ КОЛИЧЕСТВА КОЛОСКОВ НА КОЛОС НА ХРОМОСОМЕ 7А В ПШЕНИЦЕ**

### **ОБЛАСТЬ ИЗОБРЕТЕНИЯ**

[1] Настоящее изобретение относится к области оптимизации характеристик растений с помощью методов молекулярной биологии, маркерной технологии и генной инженерии. Настоящим изобретением предоставляются технические средства, такие как молекулы нуклеиновой кислоты, векторы, способы и их применение для получения и идентификации нетрансгенных и трансгенных растений пшеницы с измененными фенотипами «общего количества колосков на колос» (далее по тексту – «SPS» (от англ. total spikelet number per spike)).

### **ПРЕДПОСЫЛКИ К СОЗДАНИЮ ИЗОБРЕТЕНИЯ**

[2] Урожайность пшеницы определяется преимущественно тремя составляющими: количество продуктивных колосьев на единицу площади, количество зерен на колос и вес зерна. Одним из основных факторов, которые способствуют повышению урожайности пшеницы, является увеличение количества зерен на колос или увеличение количества зерен на колос и количество колосьев на единицу площади. На общее количество зерен могут дополнительно влиять такие характеристики, как количество продуктивных побегов на растение, количество колосков на колосе, количество жизнеспособных соцветий на колоске. Увеличение по любой из составляющих или признаков урожайности теоретически может увеличить потенциальную урожайность пшеницы. Однако, из-за конкуренции за ассимиляты на стадии роста колоса могут возникать компенсационные эффекты, таким образом, увеличение одной из составляющих или признаков может и не приводить к увеличению общего урожая зерна.

[3] Генетика, определяющая архитектуру соцветий пшеницы, остается в значительной степени неизвестной. До сих пор было показано, что только ген чувствительности к фотопериоду *Ppd-1* влияет на число колосков [Shaw, L.M., et al., PLoS One, 2013. 8(11): стр. e79459]. Все это представляет собой огромный источник

неиспользованного генетического потенциала для содействия усилиям по достижению 70%-ного увеличения урожайности сельскохозяйственных культур, необходимого к 2050 году для питания растущего населения мира [United Nations, F.a.A.O.o.t.U. *How to Feed the World in 2050*. в *Rome: High-Level Expert Forum*. 2009]. Соцветие пшеницы (обычно называемое колосом, ушком или головкой) состоит из колосков, прикрепленных в узлах оси колоса. Каждый из колосков, в свою очередь, состоит из двух чешуек, число цветков которых обычно от двух до четырех образуют зерно после осеменения. Конечное количество колосков определяется образованием верхушечного колоска. Это происходит, когда последние инициированные примордии, вместо того чтобы стать колосковыми примордиями, развиваются в чешуйчатые и цветковые примордии [Kirby, E.J.M. и M. Appleyard, F.G.H. Lupton, Editor. 1987, Springer Netherlands: Dordrecht. стр. 287-311].

[4] Развитие постоянных картирующих популяций пшеницы в последние годы, сопровождаемое построением полногеномных карт маркеров на основе большого количества молекулярных маркеров, открыло возможность идентифицировать, анализировать и использовать QTL в отношении агрономических признаков, включая количество колосков на колосе.

[5] В работе Tian et al. (2015 Genetic analyses of wheat and molecular marker-assisted breeding (Генетический анализ пшеницы и селекция с использованием молекулярных маркеров), том 1 Science Press Beijing) обобщается информация для QTL, относящегося к морфологии колоса и (на странице 167, Таблица 1.37), в частности, к количеству колосков. QTL идентифицируются на хромосоме 2D, 2DS, 3AS, 3B, 3DL, 4AL, 4DS, 5A, 5B, 5D, 7A, 7AL и 7D.

[6] В работе Jantasuriyarat et al. (2004, Theor. Appl Genet. 108: 261-273) сообщается о двух QTL, которые связаны с количеством колосков и расположены на хромосоме 7A среди других QTL, с использованием рекомбинантных инбредных линий картирующей популяции Международной инициативы по картированию Triticaceae. Один QTL, ограниченный маркерами *Xfba69-XksuH9* (182,7-213.4 cM – пиковый маркер 196.3 cM – наиболее близкий локус *Xmwg938*), или ограниченный маркерами *Xfba350-Xfbb18* – 188.5.3-201.3 cM – пиковый маркер 196.3 cM- наиболее близкий

локус *Xmwg938*) был значительным в двух локациях за два года, в то время как другой QTL был значительным за один год лишь в одной локации (маркеры *Xfbg354-Xfba350-160.1-174.9 cM* – пиковый маркер 164.9 cM – наиболее близкий локус *Xfba69*). Во всех случаях количество колосков увеличивалось аллелями, полученными от *Opata 85*.

[7] В работе Ma et al. (2007, *Mol. Gen. Genomics* 277: 31-42) сообщается о двух QTL, относящихся к количеству колосков на колосе, на хромосоме 7A в популяции рекомбинантных инбредных линий (далее – «RIL»), полученных с использованием поколения одного семени от скрещивания *Nanda2419* и *Wangshuibai*, или в иммортализованной популяции F2, полученной путем случайного перекрестного скрещивания этих RIL. В популяции RIL интервал QTL был ограничен маркерами *Xbarc154-Xwmc83e*, в то время как в популяции IF2 QTL был ограничен маркерами *Xwmc83-Xwmc17*. Аллели *Wangshuibai* способствовали увеличению количества колосков на колос.

[8] В работе Xu et al. (2014, *Theor. Appl. Genet.* 127: 59-72) сообщается об идентификации QTL, относящегося к характеристике SPS, на хромосоме 7A в популяции RIL от скрещивания между *Xiaoyan 54* и *Jing 411*), который был идентифицирован маркерами *Xgwm276-Xbarc192-Xbarc253*. Благоприятный аллель был получен от родителя *Jing 411*.

[9] В работе Zhai et al. (2016 *Frontiers in Plant Science*, Том 7, статья 1617) описана область на хромосоме 7A, идентифицированная в работе Xu et al., 2014 г. и указывается область, расположенная в интервале 123.50-137.50 cM.

[10] В работе Saarah Noriko Kuzay et al. (P0848 International Plant and Animal Genome XXV, январь 14-18, 2017 Сан-Диего), в реферате сообщается об идентификации QTL на длинном плече хромосомы 7AL, относящегося к характеристике SPS, с использованием полногеномных исследований ассоциаций. Валидация этого QTL в бипарентальной популяции *BerkutxRC875* позволила получить точное генетическое картирование области хромосомы 7AL длиной 2 Мб. В среднем линии, несущие аллель *Berkut* для SPS, имели в 2,4 раза больше колосков на колос по сравнению с линиями, несущими аллель *RAC875* для области пика QTL. В указанной работе также сообщается о развитии большой популяции с высокой плотностью из двух

гетерозиготных инбредных семейств для точного картирования и клонирования гена, лежащего в основе этого QTL, в результате.

[11] В работе Zhang et al. (2015, Scientific Reports DOI 10:1038/srep12211) сообщается, что в развитии колосков пшеницы может участвовать предполагаемый ортолог MOC1 из пшеницы (MOC1 означает *MONCULM1* в рисе). При картировании TaMoc1-A был сопоставлен с областью, к которой прилегают *WMC488*(4.7 cM) и *P2071-180* (11.6cM), на хромосоме 7A в популяции удвоенных гаплоидов от скрещивания Nanxuan 10 и Lumai 14. Гаплотип TaMoc1-7A NapH был связан с небольшим увеличением количества колосков на колос в 10 средах в течение 3 лет и на 2 участках. Однако этот ортолог TaMOC1 не является геном, лежащим в основе описанного в настоящем документе QTL для характеристики SPS на хромосоме 7A. После выравнивания TaMOC1-7A относительно эталонного генома NRgene-HiC пшеницы сорта Chinese Spring (иногда сокращенно – «CS») TaMOC1-7A картируется на 557 480 502 п.о. на хромосоме 7A, что отстоит более чем на 100 Мб от описанного в настоящем документе и проанализированного QTL для SPS на 7A и, следовательно, является другим. Как показано ниже, левый и правый маркеры, идентифицирующие интервал QTL на карте картирующей популяции MAGIC, находятся в положении 671,146,796 и 674,103,435 соответственно, в то время как маркеры, определяющие интервал QTL на карте исследования GWAS, находятся в положении 674,203,435 и 674,203,741 на хромосоме 7A пшеницы (положения относятся к эталонной геномной последовательности NRgene-HiC Chinese Spring).

[12] Таким образом, все еще существует потребность в дальнейшем генетическом анализе SPS QTL, расположенного на хромосомах 7 пшеницы, в частности 7A, для идентификации лежащих в основе генов, чтобы облегчить оптимизацию количества колосков на колос для максимальной урожайности пшеницы.

#### **КРАТКОЕ ИЗЛОЖЕНИЕ СУЩНОСТИ ИЗОБРЕТЕНИЯ**

[13] В одном аспекте изобретением предоставляется белок, участвующий в определении количества колосков на колос пшеницы, который является ортологом белка Aberrant panicle organization 1 (Apo1) из риса. Этот белок содержит аминокислотную последовательность, выбранную из группы, состоящей из

следующих последовательностей: а) аминокислотная последовательность SEQ ID NO: 3, 15 or 17 или ее функциональный вариант, и b) аминокислотная последовательность, имеющая, по меньшей мере, 85% идентичности последовательности с аминокислотной последовательностью SEQ ID NO: 3, 15 или 17, или ее функциональным вариантом.

[14] Другой целью настоящего изобретения является предоставление изолированной нуклеиновой кислоты, кодирующей белок по изобретению, которая может содержать нуклеотидную последовательность, выбранную из группы, состоящей из следующих последовательностей: а) нуклеотидная последовательность SEQ ID NO: 1 или SEQ ID NO: 2, b) нуклеотидная последовательность, имеющая, по меньшей мере, 80% идентичности с нуклеотидной последовательностью SEQ ID NO: 1 или SEQ ID NO: 2, c) нуклеиновая кислота, имеющая последовательность, комплементарную нуклеиновой кислоте а) или b). Нуклеиновая кислота по изобретению может располагаться на хромосоме 7А пшеницы в интервале, включающем нуклеотидную последовательность между нуклеотидом в положении 674,081,462 в эталонной геномной последовательности NRgene-HiC Chinese Spring и нуклеотидом в положении 674,082,918 в эталонной геномной последовательности NRgene-HiC Chinese Spring, фланкированном маркерами SEQ ID NO: 10 и SEQ ID NO: 11 или фланкированном маркерами SEQ ID NO:12 и или SEQ ID NO: 13, или SEQ ID NO: 14, или фланкированном маркерами SEQ ID NO: 23 и SEQ ID NO: 24, или может располагаться на хромосоме 7В пшеницы в интервале, фланкированном маркерами SEQ ID NO: 26 и 27. В одном варианте осуществления изобретения изолированная нуклеиновая кислота, кодирующая белок по настоящему изобретению, может содержать нуклеотидную последовательность, выбранную из группы, состоящей из следующих последовательностей: а) нуклеотидная последовательность SEQ ID NO: 1, 2, 6, 7, 15, 16, 20, 21, 28 или 30, b) нуклеотидная последовательность, имеющая, по меньшей мере, 80% идентичности с нуклеотидной последовательностью SEQ ID NO: 1, 2, 6, 7, 15, 16, 20, 21, 28 или 30, c) нуклеиновая кислота, имеющая последовательность, комплементарную нуклеиновой кислоте а) или b). В одном варианте осуществления изобретения изолированная нуклеиновая кислота, кодирующая белок по настоящему изобретению, может содержать нуклеотидную последовательность, выбранную из

группы, состоящей из следующих последовательностей: а) нуклеотидная последовательность SEQ ID NO: 1, 2, 6, 7 или 28, б) нуклеотидная последовательность, имеющая, по меньшей мере, 80% идентичности с нуклеотидной последовательностью SEQ ID NO: 1, 2, 6, 7 или 28, с) нуклеиновая кислота, имеющая последовательность, комплементарную нуклеиновой кислоте а) или б). В одном варианте осуществления изобретения любая из таких нуклеотидных последовательностей представляет собой изолированную или искусственную нуклеиновую кислоту.

[15] Настоящим изобретением также предоставляется рекомбинантный ген, содержащий экспрессируемый в растении промотор, функционально связанный с нуклеотидной последовательностью, кодирующей белок по настоящему изобретению, и, при необходимости, последовательность терминации транскрипции и полиаденилирования, предпочтительно участок терминации транскрипции и полиаденилирования, которые являются функциональными в растениях. В еще одном варианте осуществления изобретения промотор, экспрессируемый в растении, может быть выбран из конститутивного промотора, индуцируемого промотора или тканеспецифического промотора. Экспрессируемый в растении промотор может представлять собой промотор CaMV35S, убиквитинный промотор или нативный промотор гена APO1 по изобретению, полученный из сорта пшеницы с относительно высоким количеством колосков на колосе.

[16] В соответствии с еще одним аспектом изобретением предоставляется растение пшеницы, часть растения или семя, состоящее из клеток растения пшеницы, содержащие рекомбинантный ген, описанный в настоящем документе.

[17] В альтернативных вариантах осуществления изобретения предоставляются способ получения растений пшеницы с измененным количеством колосков на колос или способ изменения количества колосков на колос растения пшеницы, при этом оба способа включают этап изменения содержания белка по изобретению в растении пшеницы. В соответствии с одним из вариантов осуществления изобретения увеличивается количество белка и увеличивается количество колосков на колос по сравнению с количеством колосков на колос растения пшеницы с неизменным количеством белка, в частности, в случае, если растение пшеницы изначально имеет

(относительно) низкое количество колосков на колосе. Количество белка по изобретению можно увеличить путем получения указанного растения пшеницы с а) рекомбинантным геном по изобретению или б) гетерологичным геном, кодирующим белок по изобретению, при этом степень экспрессии гетерологичного гена выше, чем степень экспрессии соответствующего эндогенного гена. Гетерологичный ген может содержать нуклеотидную последовательность SEQ ID NO: 4, 5, 9, 19 или SEQ ID NO: 22 или нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с любой из указанных последовательностей. В соответствии с одним из вариантов осуществления изобретения гетерологичный ген может содержать нуклеотидную последовательность SEQ ID NO: 4, 5, 9, 19, или нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с любой из указанных последовательностей, при этом указанная последовательность характеризуется делецией приблизительно 115 нуклеотидов (например, 100–130 нуклеотидов или 115 нуклеотидов) в положении примерно на 500 нуклеотидов выше стартового кодона ATG (что соответствует стартовому кодону в эталонной последовательности SEQ ID NO: 1).

[18] В соответствии с еще одним из вариантов осуществления изобретения количество белка уменьшается, и уменьшается количество колосков на колос по сравнению с количеством колосков на колос растения пшеницы с неизменным количеством белка, в частности, в случае, если растение пшеницы изначально имеет (относительно) высокое количество колосков на колосе. Количество белка по изобретению можно уменьшить путем получения растения пшеницы с а) гетерологичным геном, кодирующим белок по изобретению, при этом промотор указанного гетерологичного гена имеет более низкую промоторную активность, чем промотор эндогенного гена; или с б) мутантным аллелем эндогенного гена, кодирующего белок по изобретению. Гетерологичный ген может содержать нуклеотидную последовательность SEQ ID NO: 9 или нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с указанной последовательностью и предпочтительно не содержащую нуклеотидную последовательность от положения 4399 до положения 4513 SEQ ID NO: 5, нуклеотидную последовательность, имеющую, по меньшей мере,

90% идентичности последовательности с указанной последовательностью. Гетерологичный ген может также содержать нуклеотидную последовательность SEQ ID NO: 19 или нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с указанной последовательностью и предпочтительно не содержащую нуклеотидную последовательность от положения 7816 до положения 7930 SEQ ID NO: 19 или нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с указанной последовательностью. Мутантный аллель может представлять собой нокаут-аллель. Мутантный аллель также может представлять собой мутантный аллель с заменой или мутантный аллель с делецией или вставкой, предпочтительно с более низкой активностью.

[19] В еще одном варианте осуществления изобретения этап получения растения в вышеописанных способах включает получение растения путем трансформации, скрещивания, обратного скрещивания, интрогрессии, редактирования генома или мутагенеза.

[20] В соответствии с другими вариантами осуществления изобретения описываются способы идентификации и/или отбора растения пшеницы, содержащего аллель гена, положительно или отрицательно влияющего на количество колосков на колос, которые включают, соответственно, этап идентификации наличия или отсутствия, соответственно, в геноме растения пшеницы нуклеиновой кислоты, имеющей нуклеотиды от положения 4399 до положения 4513 SEQ ID NO: 5, или нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с указанной последовательностью, или нуклеиновой кислоты, имеющей нуклеотидную последовательность от положения 7816 до положения 7930 SEQ ID NO: 19 или нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с указанной последовательностью.

## **КРАТКОЕ ОПИСАНИЕ СХЕМ**

[21] **Фигура 1:** Уровень экспрессии РНК APO1 в различных сортах яровой пшеницы (основатели MAGIC) и контрастных HIF с аллелем, влияющим на SPS, и без него. TS: Terminal Spikelet, DR: Double Ridge. Baxter, Chara, Westonia и Yitpi являются

родителями двойной межлинейной популяции MAGIC. Fam1\_A\_1, Fam1\_B\_1, Fam2\_B\_1, Fam2\_C\_1, Fam2\_H\_1, Fam3\_E\_1, Fam3\_I\_1, Fam4\_A, Fam4\_G, Fam5\_C\_1 и Fam5\_F\_1 – одиннадцать анализируемых НИФ. Звездочкой отмечены линии с большим количеством колосков на колос.

[22] **Фигура 2:** А. Распределение средних фенотипов всех линий из популяции озимой пшеницы 2014 г., фенотипированных по общему количеству колосков на колос (SPS) и указание SPS для сортов-основателей пшеницы. В. Сводные данные по вариации фенотипов SPS и связанной наследственности.

[23] **Фигура 3:** Точное картирование *QTsn.jbl-7*. а) Mrwgaim QTL модель б) выравнивание генетической карты MAGIC с) физическая карта IWGScv1 с моделями аннотированных генов MEGAP д) полиморфизмы последовательностей между Robigus и Claire/Chinese Spring в ортологе *APO1*.

[24] **Фигура 4:** Синтенические отношения *QTsn.jbl-7A* QTL к *QTsn.jbl-7B* QTL и *qPBN6* QTL риса.

[25] **Фигура 5:** а) Экспрессия TaAPO1-7A транскрипта относительно генов «домашнего хозяйства» TaRP15 [Shaw, L.M., A.S. Turner, и D.A. Laurie, Plant J, 2012. 71(1): стр. 71-84] Ta2291 [Paolacci, A.R., et al., BMC Molecular Biology, 2009. 10(1): стр. 11] и приведенных к экспрессии TaAPO1-7A в Brompton. б) Регрессия экспрессии TaAPO1-7A на BLUP общего числа колосков для линий MAGIC Founder в полевых испытаниях 2014 года. Все сорта были отобраны на стадии GS32, за исключением Soissons, который был на стадии GS34 из-за ускоренного цветения, вызванного аллелем Ppd-D1. Таким образом, причины низкой экспрессии TaAPO1-7A в Soissons, вероятно, отличаются от тех, которые связаны с вариацией последовательности, наблюдаемой у Robigus и Brompton.

## ПОДРОБНОЕ ОПИСАНИЕ

[26] Настоящее изобретение основано на той идее, что ортолог пшеницы Apo1 риса участвует в определении количества колосков на колос у сортов пшеницы, включая сорта яровой и озимой пшеницы.

[27] В одном аспекте изобретением предоставляется белок, участвующий в определении количества колосков на колос пшеницы, который является ортологом белка Aberrant panicle organization 1 (Apo1) из риса. Этот белок содержит аминокислотную последовательность, выбранную из группы, состоящей из следующих последовательностей: а) аминокислотная последовательность SEQ ID NO: 3, 15 или 17 или ее функциональный фрагмент, и б) аминокислотная последовательность, имеющая, по меньшей мере, 85% идентичности последовательности с аминокислотной последовательностью SEQ ID NO: 3, 15 или 17 или ее функциональный вариант.

[28] Количество колосков на колос контролируется как генетически, так и экологически. У разных сортов пшеницы имеется разное среднее количество колосков на колос в определенной среде. Наблюдаемое количество колосков на колос первичного стебля колеблется в диапазоне от 17 до 40 в зависимости от наблюдаемой линии пшеницы. Сорта яровой пшеницы, как правило, имеют меньшее количество колосков на колосе (18 - 24), тогда как сорта озимой пшеницы обычно имеют большее количество колосков на колосе. Если линии пшеницы содержат аллель SPS QTL, который влияет на увеличение количества колосков на колосе, количество колосков увеличивается, по меньшей мере, на 1, но иногда на 2 или на 3 по сравнению с аналогичной линией без аллеля, который влияет на увеличение количества колосков на колосе, независимо от оставшегося генетического профиля или среды.

[29] При использовании по тексту настоящего документа термины «белок» и «полипептид» используются взаимозаменяемо и описывают группу молекул, состоящую из более чем 30 аминокислот, тогда как термин «пептид» описывает молекулы, которые состоят менее чем из 30 аминокислот. Белки и пептиды могут дополнительно образовывать димеры, тримеры и высшие олигомеры, то есть соединения, состоящие из более чем одной (поли)пептидной молекулы. Молекулы

белка или пептида, образующие такие димеры, тримеры и т.д., могут быть идентичными или не идентичными. Соответствующие структуры более высокого порядка, следовательно, именуется гомо- или гетеродимерами, гомо- или гетеротримерами и т.д. Термины «белок» и «пептид» также относятся к естественно модифицированным белкам или пептидам, в которых модификация осуществляется, например, путем гликозилирования, ацетилирования, фосфорилирования и другими подобными способами. Такие модификации известны специалистам.

[30] В работе Ikeda et al. 2005 (*Developmental Biology*, 282:349-360) идентифицирован ген ABERRANT PANICLE ORGANIZATION 1 (APO1) в качестве ключевого регулятора цветков риса. Утрата функции Apo1 приводила к преждевременному превращению меристем соцветий в меристемы колосков, в результате чего количество колосков уменьшалось (Ikeda et al 2005, Ikeda et al. 2007, *Plant Journal* 51, 1030-1040). Мутация с усилением функции Apo1 приводила к замедленному превращению меристем соцветий в меристемы колосков, в результате чего количество колосков увеличивалось (Ikeda et al. 2007, Ikeda Kawakatsu et al. 2009, *Plant physiol.* 150:736-747). APO1 был также идентифицирован Terao et al. 2010 (*Theor Appl Genet*, 120:875-893) как ген, ответственный за локус количественного признака, который влияет на увеличение количества первичных осей колоса, количества зерен на метелку и урожая зерна на растение риса.

[31] При использовании по тексту настоящего документа термин «ген, ортологичный APO1» означает ген, который обнаружен у разных видов, но произошел от общего предкового гена при видообразовании и сохранил ту же функцию. APO1 кодирует белок F-box, а известные ортологичные гены включают ген, именуемый UNUSUAL FLORAL ORGANS (UFO), из *Arabidopsis* и ген, именуемый DOUBLE TOP (DOT), из петунии, которые, как было показано, также управляют временем перехода к цветению и архитектурой соцветия.

[32] SEQ ID NO: 3 представляет собой аминокислотную последовательность гена APO1 из сорта пшеницы Chinese Spring. Сорта Baxter и Westonia продуцируют белок APO1 с аминокислотной последовательностью, идентичной последовательности SEQ ID NO: 3. SEQ ID NO: 8 представляет собой аминокислотную последовательность гена

АРО1 из сорта пшеницы Chara. Сорт Yitpi продуцирует белок АРО1 с аминокислотной последовательностью, идентичной последовательности SEQ ID NO: 8. Белок АРО1, имеющий аминокислотную последовательность SEQ ID NO: 8, представляет собой функциональный вариант белка АРО1 с аминокислотной последовательностью SEQ ID NO: 3. Сорт Claire продуцирует белок АРО1 с аминокислотной последовательностью, идентичной последовательности SEQ ID NO: 3. Сорта Robigus, Cadenza и Paragon продуцируют белок АРО1 с аминокислотной последовательностью SEQ ID NO: 3, где фенилаланин в положении 47 замещен цистеином, а аспарагиновая кислота в положении 384 замещена аспарагином. Белок АРО1, имеющий аминокислотную последовательность SEQ ID NO: 3, где фенилаланин в положении 47 замещен цистеином, а аспарагиновая кислота в положении 384 замещена аспарагином, представляет собой функциональный вариант белка АРО1 с аминокислотной последовательностью SEQ ID NO: 3. SEQ ID NO: 29 представляет собой аминокислотную последовательность гена АРО1 на хромосоме 7А из сорта пшеницы Chinese Spring в соответствии с альтернативной генной моделью и не содержит 27 N-концевых аминокислот SEQ ID NO: 3. SEQ ID NO: 17 представляет собой аминокислотную последовательность гена АРО1 на хромосоме 7В из сортов пшеницы Chinese Spring и Claire. В сорте Robigus белок характеризуется заменой Н47R и А173S. SEQ ID NO: 31 представляет собой аминокислотную последовательность гена АРО1 на хромосоме 7В из сорта пшеницы Chinese Spring в соответствии с альтернативной генной моделью и не содержит 71 N-концевых аминокислот SEQ ID NO: 17. SEQ ID NO: 3 имеет 89% идентичности последовательности с SEQ ID NO: 17. SEQ ID NO: 29 и 31 имеют 98% идентичности последовательности.

[33] В соответствии с изобретением могут использоваться белки АРО1 (также именуемые вариантами), которые содержат аминокислотную последовательность, имеющую, по меньшей мере, 85% или, по меньшей мере, 90% или, по меньшей мере, 95% или, по меньшей мере, 98% или, по меньшей мере, 99% идентичности последовательности с последовательностью белка в соответствии с описанием в настоящем документе, или идентичны последовательности белка в соответствии с описанием в настоящем документе. Термин «вариант» в отношении аминокислотных последовательностей SEQ ID NO: 3 или SEQ ID NO: 8 означает существенно схожие

последовательности. В соответствии с одним из вариантов осуществления изобретения вариант белка по изобретению представляет собой искусственный белок в соответствии с определением или вариант белка, который не включает какой-либо природный белок.

[34] При использовании по тексту настоящего документа термин «процент идентичности последовательности» относится к проценту идентичных аминокислот в двух участках оптимально выровненных аминокислотных последовательностей или проценту идентичных нуклеотидов в двух участках оптимально выровненных нуклеотидных последовательностей. Оптимальное выравнивание последовательностей для выравнивания сравнительного участка хорошо известно специалистам и может осуществляться с помощью таких инструментов, как алгоритм локальной гомологии Смита и Ватермана (Waterman, M. S., Chapman & Hall. Лондон, 1995), алгоритм гомологического выравнивания Нидлмана-Вунша (1970), метод поиска сходства Пирсона-Липмана (1988), и предпочтительно с помощью компьютеризированного обеспечения этих алгоритмов, такого как GAP, BESTFIT, FASTA, и TFASTA, доступного как часть программного пакета GCG (зарегистрированный торговый знак), пакета Wisconsin (зарегистрированный торговый знак Accelrys Inc., Сан-Диего, Калифорния). Термин «относительное идентичное количество» для выровненных сегментов исследуемой или эталонной последовательности означает число идентичных компонентов в двух выровненных последовательностях, разделенное на общее число компонентов в сегменте эталонной последовательности, т.е. целой эталонной последовательности или меньшей определенной части эталонной последовательности. «Процент идентичности» – это доля идентичности от 100%. Одна или несколько аминокислотных последовательностей или последовательностей ДНК могут сравниваться с аминокислотной последовательностью или последовательностью ДНК полной длины или с их частью, или с более длинной аминокислотной последовательностью или последовательностью ДНК. Расчет идентичности последовательности осуществляется на основе более короткой нуклеотидной или аминокислотной последовательности.

[35] Кроме того, очевидно, что варианты белков APO1 пшеницы с делецией, заменой или вставкой одного или нескольких аминокислотных остатков также эффективно

могут использоваться в способах по изобретению, при условии, что делеции, замены или вставки нет в домене F-box (SEQ ID NO: 3 от аминокислотного положения 33 до аминокислотного положения 77 (в соответствии с определением в базе данных Pfam).

[36] Примерами замен являются консервативные замены, то есть замены одной аминокислоты другой с аналогичными физико-химическими свойствами. Известно, что эти замены не влияют на структуру белка. Такие замены достигаются путем замены одной аминокислоты другой аминокислотой, принадлежащей той же группе, следующим образом:

Группа 1: цистеин (C);

Группа 2: фенилаланин (F), триптофан (W) и тирозин (Y);

Группа 3: гистидин (H), лизин (K) и аргинин (R);

Группа 4: аспарагиновая кислота (D), глутаминовая кислота (E), аспарагин (N) и глутамин (Q);

Группа 5: изолейцин (I), лейцин (L), метионин (M) и валин (V);

Группа 6: аланин (A), глицин (G), пролин (P), серин (S) и треонин (T).

[37] Другой целью настоящего изобретения является предоставление нуклеиновой кислоты, включая изолированную или искусственную нуклеиновую кислоту, кодирующей белок по изобретению, которая может содержать нуклеотидную последовательность, выбранную из группы, состоящей из следующих последовательностей: а) любая из нуклеотидных последовательностей SEQ ID NO: 1, 2, 6, 7 или 28, б) нуклеотидная последовательность, имеющая, по меньшей мере, 80% идентичности с нуклеотидной последовательностью SEQ ID NO: 1, 2, 6, 7 или 28; и с) нуклеиновая кислота, имеющая последовательность, комплементарную нуклеиновой кислоте а) или б). Нуклеиновая кислота по изобретению может располагаться на хромосоме 7A пшеницы в интервале, включающем нуклеотидную последовательность между нуклеотидом в положении 674,081,462 и нуклеотидом в положении 674,082,918 в эталонной геномной последовательности пшеницы Chinese Spring (NRgene-HiC), фланкированном маркерами SEQ ID NO: 10 и SEQ ID NO: 11, или фланкированном маркерами SEQ ID NO: 12 и либо SEQ ID NO: 13, либо SEQ ID NO: 14, или фланкированном маркерами SEQ ID NO: 23 и SEQ ID NO: 24, или может располагаться

на хромосоме 7В пшеницы в интервале, фланкированном маркерами SEQ ID NO: 26 и 27.

[38] При использовании по тексту настоящего документа термин «изолированная нуклеиновая кислота», используемая взаимозаменяемо с термином «изолированная ДНК», относится к нуклеиновой кислоте, не встречающейся в ее естественном геномном контексте, независимо от ее длины и последовательности. Термин «изолированная ДНК» может, например, относиться к ДНК, которая физически отделена от геномного контекста, например, к фрагменту геномной ДНК. Изолированная ДНК также может представлять собой искусственно полученную ДНК, например, химически синтезированную ДНК или ДНК, полученную с помощью реакций амплификации, таких как полимеразная цепная реакция (ПЦР), известных специалистам. Термин «изолированная ДНК» может также относиться к ДНК, присутствующей в контексте ДНК, в которой она не встречается в природе. Например, термин «изолированная ДНК» может относиться к фрагменту ДНК, присутствующему в плазмиде. Кроме того, термин «изолированная ДНК» может относиться к фрагменту ДНК, присутствующему в другом хромосомном контексте, чем тот, в котором он встречается в природе, например, в таком положении в геноме, которое отличается от естественного положения, в геноме другого вида, чем тот вид, в котором он встречается в естественных условиях, или в искусственной хромосоме. При использовании по тексту настоящего документа термин «искусственная ДНК» или «искусственная нуклеиновая кислота» означает ДНК или нуклеиновую кислоту, которая отличается от природной ДНК или нуклеиновой кислоты (т.е., например, имеет отличия в последовательности, или которая отличается каким-либо другим образом, например, имеет одно или несколько внутренних нуклеотидных делеций (за исключением делеций на обоих концах), которые не встречаются в природе, или нуклеотидных замен или вставок, которые не встречаются в природе), которая имеет другую нуклеотидную последовательность по сравнению с природной последовательностью, которая связана с меткой или молекулой, с которыми такая ДНК или нуклеиновая кислота не связаны по природе (например, связь с гетерологичным или искусственным промотором или 3'-нетранслируемой областью) и т.д.). Аналогичным образом, термин «искусственный белок» по изобретению

представляет собой белок, который отличается от встречающегося в природе белка (белок с отличиями в последовательности или любыми другими отличиями, например, с наличием одной или нескольких делеций аминокислот (в одном варианте осуществления – это внутренние делеции аминокислот (не делеции на одном из концов белка)), не встречающихся в природе, или с наличием аминокислотных замен или вставок, которые не встречаются в белке в природе, с другой аминокислотной последовательностью, отличной от с природной, с наличием связи с меткой или молекулой, с которой белок не связан в природе и т.д.). Последовательность искусственной ДНК или нуклеиновой кислоты была изменена человеком по сравнению с природной формой, например, посредством (химического или другого) мутагенеза, рекомбинации, направленного редактирования генома или редактирования оснований с использованием последовательность-специфических нуклеаз и т.д.

[39] В соответствии с изобретением могут использоваться нуклеиновые кислоты (также именуемые вариантами), кодирующие белок APO1, которые содержат нуклеотидную последовательность, имеющую, по меньшей мере, 40%, по меньшей мере, 50% или, по меньшей мере, 60% или, по меньшей мере, 70% или, по меньшей мере, 80% или, по меньшей мере, 85% или, по меньшей мере, 90% или, по меньшей мере, 95% или, по меньшей мере, 98% идентичности последовательности с геном, описанном в настоящем документе. Термин «вариант» в отношении нуклеотидных последовательностей SEQ ID NO: 1, 2, 6, 7 или 28 по настоящему изобретению означает существенно схожие нуклеотидные последовательности, кодирующие аминокислотные последовательности существенно схожие с любой из аминокислотных последовательностей SEQ ID NO: 3, 8 или 29. Термин «вариант» в отношении нуклеотидных последовательностей SEQ ID NO: 15, 20, 21, 7 или 30 по настоящему изобретению означает существенно схожие нуклеотидные последовательности, кодирующие аминокислотные последовательности существенно схожие с любой из аминокислотных последовательностей SEQ ID NO: 17 или 31. Термин «вариант» в отношении нуклеотидной последовательности SEQ ID NO: 16 по настоящему изобретению означает существенно схожие нуклеотидные последовательности, кодирующие аминокислотные последовательности существенно схожие с любой из аминокислотных последовательностей SEQ ID NO: 18. Природные

аллельные варианты могут быть идентифицированы при помощи известных методов молекулярной биологии таких как, например, метод полимеразной цепной реакции (ПЦР) и метод гибридизации в соответствии с описанием в настоящем документе. Вариантные нуклеотидные последовательности также включают нуклеотидные последовательности, полученные синтетическим методом, такие как последовательности, образованные, например, путем направленного мутагенеза любой из SEQ ID NO: 1, 2, 6, 7, 15, 16, 20, 21, 28 или 30. В целом, варианты нуклеотидной последовательности по изобретению будут иметь, по меньшей мере, по меньшей мере, 40%, 50%, 60%, до 70%, например, предпочтительно 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, до 79%, в общем, по меньшей мере, 80%, например, 81% - 84%, по меньшей мере, 85%, например, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, до 98% и 99% идентичности последовательности по отношению к любой из нуклеотидных последовательностей SEQ ID NO: 1, 2, 6, 7, 15, 16, 20, 21, 28 или 30. Производные описанных в настоящем документе молекул ДНК могут включать, помимо прочего, удаления последовательности, одиночные или множественные точечные мутации, изменения в конкретном сайте рестриктазы, добавление функциональных элементов или другие средства молекулярной модификации. Способы получения таких производных хорошо известны специалистам (смотрите, например, J. F. Sambrook, D. W. Russell, и N. Irwin (2000) *Molecular Cloning: A Laboratory Manual*, 3<sup>rd</sup> издание Тома 1, 2 и 3. Cold Spring Harbor Laboratory Press). Специалистами известны стандартные методические материалы, в которых описаны конкретные условия и процедуры для конструирования, манипулирования и изоляции макромолекул (например, молекул ДНК, плазмид и т.д.), создания рекомбинантных организмов, а также скрининга и изолирования молекул ДНК. В соответствии с одним из вариантов осуществления изобретения вариант ДНК или нуклеиновой кислоты по изобретению представляет собой искусственную ДНК или нуклеиновую кислоту, или вариант ДНК или нуклеиновой кислоты, который не включает какую-либо природную ДНК или нуклеиновую кислоту.

[40] SEQ ID NO: 1 представляет собой нуклеотидную последовательность, кодирующую ДНК Apo1 из сорта пшеницы Chinese Spring. SEQ ID NO: 2 представляет собой соответствующую геномную ДНК Apo1 из сорта Chinese Spring. SEQ ID NO: 28

представляет собой нуклеотидную последовательность, кодирующую ДНК Apo1 на хромосоме 7A из сорта пшеницы Chinese Spring в соответствии с альтернативной генной моделью. Сорта Baxter и Westonia содержат ген APO1 с нуклеотидной последовательностью, идентичной SEQ ID NO:1, в качестве нуклеотидной последовательности, кодирующей ДНК, и нуклеотидной последовательностью, идентичной SEQ ID NO: 2 в качестве соответствующей геномной DNA Apo1. SEQ ID NO: 6 представляет собой нуклеотидную последовательность, кодирующую ДНК Apo1 из сорта пшеницы Chara. SEQ ID NO: 7 представляет собой соответствующую геномную ДНК Apo1 из сорта Chara. Сорт Yitpi содержит ген APO1 с нуклеотидной последовательностью, идентичной SEQ ID NO: 6 в качестве нуклеотидной последовательности, кодирующей ДНК, и нуклеотидной последовательностью, идентичной SEQ ID NO: 7. Сорт Claire содержит ген APO1 с нуклеотидной последовательностью, кодирующей ДНК, и соответствующую геномную ДНК Apo1, идентичной SEQ ID NO: 1 и SEQ ID NO: 2, соответственно. Сорта Robigus, Cadenza и Paragon содержат ген APO1, имеющий в качестве нуклеотидной последовательности, кодирующей ДНК, нуклеотидную последовательность SEQ ID NO: 1, где тимин в положении 140 замещен гуанином, гуанин в положении 1150 замещен аланином и имеет в качестве нуклеотидной последовательности геномной ДНК нуклеотидную последовательность SEQ ID NO: 2, где тимин в положении 140 замещен гуанином, гуанин в положении 1284 замещен аланином. SEQ ID NO: 20 представляет собой нуклеотидную последовательность, кодирующую ДНК Apo1 из сорта пшеницы Chinese Spring на хромосоме 7B. SEQ ID NO: 30 представляет собой нуклеотидную последовательность, кодирующую ДНК Apo1 на хромосоме 7B из сорта пшеницы Chinese Spring в соответствии с альтернативной генной моделью. SEQ ID NO: 21 представляет собой соответствующую геномную ДНК Apo1-7B из сорта Chinese Spring. При рассмотрении ключевых консервативных однонуклеотидных полиморфизмов (SNP) и инсерционно-делеционных полиморфизмов в аллеле APO1 Robigus (2 SNP в кодирующей последовательности (изменение 2 аминокислот) и 1 SNP в интроне), связанных с SPS-фенотипом, в Brompton присутствовали те же консервативные SNP и инсерционно-делеционные полиморфизмы, что и в Robigus.

[41] Ген или аллель SPS Apo1 по изобретению (например, в Robigus или Yitpi) имеет следующие ключевые отличия от эталонной последовательности Apo1 Chinese Spring, причем эти различия характерны для всех SPS-аллелей Apo1, протестированных в разных популяциях для яровой или озимой пшеницы. Эти характерные отличия от эталонной последовательности Apo1-7A Chinese Spring выбраны из группы, состоящей из: а) делеция 115 п.о. приблизительно на 500 нуклеотидов выше стартового кодона ATG, 2 миссенс-SNP (где миссенс-SNP представляет собой однонуклеотидную вариацию, в результате которой кодон кодирует другую аминокислоту) в кодирующей последовательности, делеция приблизительно 5 - 7,5 т.п.о. примерно на 7,5 т.п.о. выше стартового кодона, SNP и инсерционно-делеционные полиморфизмы, присутствующие в промоторе размером приблизительно 5 т.п.о. (такие как SNP и инсерционно-делеционные полиморфизмы, показанные в Таблице 2 ниже, для Yitpi/Chara), и SNP в интроне; б) делеция 115 п.о. примерно на 500 нуклеотидов перед стартовым кодоном ATG, 2 миссенс-SNP в кодирующей последовательности, делеция приблизительно 5 - 7,5 т.п.о. примерно на 7,5 т.п.о. от кодона, SNP и инсерционно-делеционные полиморфизмы, присутствующие в промоторе размером приблизительно 5 т.п.о. (такие как SNP и инсерционно-делеционные полиморфизмы, показанные в Таблице 2 ниже, для Yitpi/Chara); в) делеция 115 п.о. приблизительно на 500 нуклеотидов выше стартового кодона ATG, 2 миссенс-SNP в кодирующей последовательности, делеция приблизительно 5 - 7,5 т.п.о., приблизительно на 7,5 т.п.о. выше стартового кодона; д) делеция 115 п.о. приблизительно на 500 нуклеотидов выше стартового кодона ATG, 2 миссенс-SNP в кодирующей последовательности; или е) делеция 115 п.о. примерно на 500 нуклеотидов выше стартового кодона ATG. Эти различия, сохраняющиеся в тестируемых линиях SPS, могут влиять на наблюдаемый фенотип SPS. Очевидно, что между аллелями SPS-Apo1 в различном генетическом окружении растений пшеницы могут иметь место и другие незначительные различия (такие как SNP/инсерционно-делеционные полиморфизмы), однако они не считаются биологически значимыми.

[42] Нуклеиновая кислота, содержащая нуклеотидную последовательность, имеющую, по меньшей мере, 80% идентичности последовательности с SEQ ID NO: 1 или SEQ ID NO: 2, таким образом может представлять собой нуклеиновую кислоту, содержащую

нуклеотидную последовательность, имеющую, по меньшей мере, 80%, или, по меньшей мере, 85%, или, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% или 100% идентичности последовательности с SEQ ID NO: 1 или SEQ ID NO: 2, соответственно. Нуклеотидная последовательность SEQ ID NO: 6 имеет, по меньшей мере, 99% идентичности последовательности с нуклеотидной последовательностью SEQ ID NO: 1. Нуклеотидная последовательность SEQ ID NO: 7 имеет, по меньшей мере, 99% идентичности последовательности с нуклеотидной последовательностью SEQ ID NO: 2.

[43] Настоящим изобретением также предоставляется рекомбинантный ген, содержащий экспрессируемый в растении промотор, включая гетерологичный или искусственный промотор, экспрессируемый в растении, функционально связанный с нуклеотидной последовательностью Apo1, кодирующей белок APO1 по настоящему изобретению, и, при необходимости, последовательность терминации транскрипции и полиаденилирования, предпочтительно участок терминации транскрипции и полиаденилирования, которые являются функциональными в растениях. В одном варианте осуществления изобретения промотор, экспрессируемый в растении, может представлять собой конститутивный промотор, индуцируемый промотор или тканеспецифический промотор. Экспрессируемый в растении промотор может представлять собой промотор CaMV35S, убиквитиновый промотор или нативный промотор гена Apo1 по изобретению, полученный из сорта пшеницы с высоким количеством колосков на колосе. В еще одном варианте осуществления нуклеиновая кислота Apo1 выбрана из следующих последовательностей: а) нуклеотидная последовательность SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 7, или SEQ ID NO: 28; или б) нуклеотидная последовательность, имеющая, по меньшей мере, 80% идентичности с нуклеотидной последовательностью SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 6, SEQ ID NO: 7, или SEQ ID NO: 28, с) нуклеиновая кислота, имеющая последовательность, комплементарную нуклеиновой кислоте а) или б), такая как искусственная нуклеиновая кислота.

[44] При использовании по тексту настоящего документа термин «рекомбинантный ген» означает искусственный ген, сконструированный путем функционального связывания фрагментов неродственных генов или других нуклеотидных

последовательностей. Другими словами, термин «рекомбинантный ген» обозначает ген, который обычно отсутствует у определенного вида растений, или относится к любому гену, в котором промотор или одна или несколько других регуляторных областей гена не связаны в природе со всей транскрибируемой нуклеиновой кислотой или с ее частью, т.е. являются гетерологичными по отношению к транскрибируемой нуклеиновой кислоте. В частности, рекомбинантный ген представляет собой искусственный, т.е. не встречающийся в природе, ген, полученный посредством функционального связывания экспрессируемого в растении промотора с нуклеотидной последовательностью, кодирующей белок АРО1.

[45] При использовании по тексту настоящего документа термин «промотор, экспрессируемый в растении» означает область последовательности ДНК, которая является значимой для инициации транскрипции в клетке растения. Этот термин включает любой промотор растительного происхождения, а также любой промотор не растительного происхождения, который способен контролировать транскрипцию в клетке растения, то есть определенные промоторы вирусного или бактериального происхождения, такие как CaMV35S, промотор вируса подземного клевера № 4 или № 7 (WO9606932) или промоторы гена Т-ДНК и другие подобные промоторы.

[46] Примеры конститутивных промоторов включают промоторы бактериального происхождения, такие как промоторы октопин-синтазы (OCS) и нопалин-синтазы (NOS) из *Agrobacterium*, а также промоторы вирусного происхождения, такие как транскрипт 35S вируса мозаики цветной капусты (CaMV) (Harster et al., 1988, Mol. Gen. Genet. 212: 182-190) или гены 19S РНК (Odell et al., 1985, Nature. 6;313(6005):810-2; патент США № 5,352,605; WO 84/02913; Benfey et al., 1989, EMBO J. 8:2195-2202), усиленный промотор 2x35S (Kay et al., 1987, Science 236:1299–1302; Datla et al. (1993), Plant Sci 94:139–149) промоторы вируса мозаики прожилок виноградного листа кассавы (CsVMV; WO 97/48819, US 7,053,205), промотор цирковируса 2xCsVMV (WO2004/053135) (AU 689 311), промотор баднавируса сахарного тростника (ScBV) (Samac et al., 2004, Transgenic Res. 13(4):349-61), промотор вируса мозаики норичника (FMV) (Sanger et al., 1990, Plant Mol Biol. 14(3):433-43), промотор вируса подземного клевера № 4 или № 7 (WO 96/06932) и усиленный промотор 35S в соответствии с описанием в US 5,164,316, US 5,196,525, US 5,322,938, US 5,359,142 и US 5,424,200.

Среди промоторов растительного происхождения, необходимо упомянуть о следующих промоторах: промотор малой субъединицы рибулозо-бискарбоксылазы/оксигеназы (Rubisco) растений (US 4,962,028; WO99/25842) из *Zea mays* и подсолнечника, промотор гена гистона H4 *Arabidopsis thaliana* (Chabouté et al. al., 1987), промотор актина 1 риса (Act-1, US 5641876), промоторы гистона в соответствии с описанием в EP 0 507 698 A1, промотор алкогольдегидрогеназы 1 *Zea mays* (Adh-1) (по данным <http://www.patentlens.net/daisy/promoters/242.html>). Также можно использовать промотор малой субъединицы из *Chrysanthemum*, если он используется в сочетании с соответствующим терминатором (Outchkourov et al., *Planta*, 216: 1003–1012, 2003). В частности, упоминаются убиквитиновые промоторы (Holtorf et al., 1995, *Plant Mol. Biol.* 29:637-649, US 5,510,474) из кукурузы, риса и сахарного тростника, например, такие как описаны Christensen and Quail (1996, *Transgenic Research*, том 5, выпуск 3, стр. 213-218).

[47] Примеры индуцибельных промоторов включают промоторы, регулируемые применением химических соединений, включая спирт-регулируемые промоторы (смотрите, например, EP637339), тетрациклин-регулируемые промоторы (смотрите, например, US 5464758), стероид-регулируемые промоторы (смотрите, например, US5512483; US6063985; US6784340; US6379945; WO01/62780), металл-регулируемые промоторы (смотрите, например, US4601978).

[48] Примеры тканеспецифических промоторов включают промоторы, специфичные для меристемы, такие как промотор OSH1 из риса (Sato et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:8117-8122) металлотеиновый промотор из риса (BAD87835.1) промоторы WAK1 и WAK2 (Wagner & Kohorn (2001) *Plant Cell* 13(2): 303-318, промотор D5, специфичный для ткани колоса, из ячменя (US6291666), промотор Lem2, специфичный для нижней/верхней цветковой чешуи, из ячменя (Abebe et al. (2005) *Planta*, 221, 170-183), промотор Pvrn1 из ячменя, специфичный для раннего соцветия (Alonse Peral et al. 2011, *PLoS ONE* 6(12) e29456), промотор Pcrs4/PrA2, из ячменя, специфичный для раннего соцветия (Koppolu et al. 2013, *Proc. Natl. Acad. Sci USA*, 110(32) 13198-13203), промотор rkn1, специфичный для меристемы, с интроном Act1 из риса (Zhang et al., 1998, *Planta* 204: 542-549, Postma-Haarsma et al. 2002, *Plant*

Molecular Biology 48: 423-441) промотор, специфичный для апикальной меристемы побега/соцветия, из *Dendrobium* sp. Pdomads1 (Yu et al. 2002).

[49] Термин «функционально связанный» относится к функциональному пространственному расположению двух или более областей нуклеиновых кислот или нуклеотидных последовательностей. Например, промоторная область может быть расположена относительно нуклеотидной последовательности таким образом, чтобы транскрипция нуклеотидной последовательности управлялась промоторной областью. Таким образом, промоторная область «функционально связана» с нуклеотидной последовательностью. «Является функционально связанной» – эквивалентный термин.

[50] Термин «гетерологичный» относится к отношению между двумя или несколькими нуклеотидными или белковыми последовательностями, которые получены из разных источников. Например, промотор является гетерологичным по отношению к функционально связанной нуклеотидной последовательности, такой как кодирующая последовательность, если такая комбинация, как правило, не встречается в природных условиях. Кроме того, конкретная последовательность может быть «гетерологичной» по отношению к клетке или организму, в которую она вставлена (то есть не встречается в природе в указанной конкретной клетке или организме). Например, рекомбинантный ген, описанный в настоящем документе, представляет собой гетерологичную нуклеиновую кислоту.

[51] Модулирование экспрессии гена APO1 пшеницы, включая увеличение его экспрессии, приводящее к изменению уровня белка Apo1, включая увеличение количества белка APO1, также может быть достигнуто путем получения растения (пшеницы) с факторами транскрипции, которые, например (в частности), распознают промоторную область APO1 и способствуют транскрипции, такие как TALEffectors, dCas, dCpf1 и т.д., и которые связаны с энхансерами транскрипции (смотрите, например, Moore et al. 2014 ACS Synth Biol. 3(10) 708-716; Qi et al. (2013) Cell 152(5) 1173-118, Liu et al. 2017 Nature Communications 8 Article Number 2095).

[52] При использовании по тексту настоящего документа термины «содержать» и «включать» указывают на присутствие заявленных свойств, целых чисел, этапов или компонентов, или их групп, однако не исключают присутствия или дополнительного

наличия одного или нескольких других свойств, целых чисел, этапов, компонентов или их групп. Таким образом, например, нуклеиновая кислота или белок, включающий последовательность нуклеотидов или аминокислот, может включать большее количество нуклеотидов или аминокислот, чем фактически указано, т.е. они могут быть встроены в нуклеиновую кислоту или белок большего размера. Рекомбинантный ген, включающий область ДНК, которая функционально или структурно определена, может включать дополнительные области ДНК и т.д. Тем не менее, в контексте настоящего описания термин «содержащий» также может употребляться в значении «состоящий из».

[53] Рекомбинантные гены, описанные в настоящем документе, при необходимости включают область ДНК, участвующую в терминации транскрипции и полиаденилировании. Специалистам известны различные участки ДНК, участвующие в терминации транскрипции и полиаденилировании у растений, и специалистам также известны терминаторные и энхансерные последовательности, которые могут использоваться в способах по изобретению. Участок полиаденилирования можно получить из природного гена, набора других растительных генов или Т-ДНК генов или даже из геномов вирусов растений. Добавляемая 3'-концевая последовательность может быть получена, например, из генов нопалин-синтазы или октопин-синтазы, или также из другого растительного гена, или из любого другого эукариотического гена.

[54] Термины «ДНК», «последовательность ДНК», «последовательность нуклеиновой кислоты», «молекула нуклеиновой кислоты», «нуклеотидная последовательность» и «нуклеиновая кислота» относятся к физической структуре, содержащей упорядоченно расположенные нуклеотиды. Последовательность ДНК или нуклеотидная последовательность может содержаться в более крупной нуклеотидной молекуле, в векторе или в других подобных структурах. Кроме того, упорядоченное расположение нуклеиновых кислот в этих последовательностях может быть изображено в форме перечня последовательностей, фигуры, таблицы, на электронном носителе и т.д.

[55] В соответствии с еще одним аспектом изобретением предоставляется растение пшеницы, часть растения или семя, состоящее из клеток растения пшеницы, содержащие рекомбинантный ген, описанный в настоящем документе.

[56] При использовании по тексту настоящего документа термин «пшеница» или «растение пшеницы» может относиться к любому сорту, пригодному для выращивания пшеницы. Примеры пшеницы включают, помимо прочего, *Triticum aestivum*, *Triticum aethiopicum*, *Triticum compactum*, *Triticum dicoccoides*, *Triticum dicoccum*, *Triticum durum*, *Triticum monococcum*, *Triticum spelta*, *Triticum turgidum*. Термин «пшеница», кроме того, включает сорта яровой и озимой пшеницы, при этом сорта озимой пшеницы определяются требованием яровизации для начала процесса цветения, тогда как сорта яровой пшеницы не требуют такой яровизации для цветения.

[57] При использовании по тексту настоящего документа термин «части растения» означают части растения, которые могут представлять собой клетки, ткани или органы растения, такие как семена, отдельные части растения, такие как корни, листья, цветы, пыльца, волокна и т.д.

[58] При упоминании по тексту термина «растение» или «растения» согласно данному изобретению подразумевается, что, за исключением случаев, когда прямо указано иное, настоящее изобретение также включает части растения (клетки, ткани или органы, семянки, семена, отделенные части, такие как корни, листья, цветки, пыльца и т.д.), потомство растений, которое сохраняет отличительные характеристики родителей, такое как семя, полученное путем самоопыления или скрещивания, например, гибридное семя (полученное путем скрещивания двух инбредных родительских линий), гибридные растения и части растений, полученный из таких растений.

[59] В соответствии с некоторыми вариантами осуществления изобретения клетки растений по изобретению, а также клетки растений, полученные способами по изобретению, могут представлять собой не размножающиеся клетки.

[60] Полученные растения по настоящему изобретению могут быть использованы в обычной схеме селекции для получения других растений с теми же характеристиками или для введения тех же характеристик в другие сорта того же вида или в растения

родственных видов или в гибридные растения. Полученные растения также могут использоваться для создания материала для размножения. Растения по настоящему изобретению также могут использоваться для получения гамет, семян (включая измельченные семена и жмых из семян), масла из семян, волокон, нитей, зиготных или соматических зародышей, потомства или гибридов растений, полученных способами по настоящему изобретению. Также изобретение включает семена, полученные из растений по настоящему изобретению.

[61] При использовании по тексту настоящего документа выражение «создание материала для размножения» относится к любым известным специалистам средствам для получения других растений, частей или семян растений, такое выражение включает, помимо прочего, способы вегетативного размножения (например, размножение отводками, деление, прививка (почками), микроклональное размножение, столоны или плети, запасающие органы, такие как луковицы, клубнелуковицы, клубни и ризомы, разрезание или разделение клубней или луковиц), половое размножение (скрещивание с другим растением) и бесполое размножение (например, апогамогония, соматическая гибридизация).

[62] В некоторых вариантах осуществления изобретения предоставляются способы получения растений пшеницы с измененным количеством колосков на колос или способ изменения количества колосков на колос растения пшеницы, при этом оба способа включают этап изменения содержания белка по изобретению в растении пшеницы. В соответствии с одним из вариантов осуществления изобретения увеличивается количество белка и увеличивается количество колосков на колос по сравнению с количеством колосков на колос растения пшеницы с неизменным количеством белка, в частности, в случае, если растение пшеницы изначально имеет низкое количество колосков на колосе. Количество белка по изобретению можно увеличить путем получения указанного растения пшеницы с а) рекомбинантным или модифицированным геном по изобретению или с b) гетерологичным геном, кодирующим белок по изобретению, при этом степень экспрессии гетерологичного гена выше, чем степень экспрессии соответствующего эндогенного гена; или с) иным способом, описанным в настоящей заявке, с использованием рекомбинантных эффекторов транскрипции. Гетерологичный ген может содержать нуклеотидную

последовательность SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 9, или SEQ ID NO: 19 или нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с указанной последовательностью.

[63] В соответствии с одним из вариантов осуществления изобретения увеличивается количество белка APO1-7A или увеличивается количество белка APO1-7A и белка APO1-7B, или увеличивается количество белка APO1-7A и белка APO1-7D, или увеличивается количество белков APO1-7A, APO1-7B и APO1-7D.

[64] В соответствии с еще одним из вариантов осуществления изобретения количество белка уменьшается, и уменьшается количество колосков на колос по сравнению с количеством колосков на колос растения пшеницы с неизменным количеством белка, в частности, в случае, если растение пшеницы изначально имеет высокое количество колосков на колосе. Количество белка по изобретению можно уменьшить путем получения растения пшеницы с а) гетерологичным геном, кодирующим белок по изобретению, при этом степень экспрессии указанного гетерологичного гена ниже, чем степень экспрессии соответствующего эндогенного гена; или с б) мутантным аллелем эндогенного гена, кодирующего белок по изобретению. Гетерологичный ген может содержать нуклеотидную последовательность SEQ ID NO: 9 или нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с указанной последовательностью, и предпочтительно не содержит нуклеотидную последовательность от положения 4399 до положения 4513 SEQ ID NO: 4 или SEQ ID NO: 5 или не содержит нуклеотидную последовательность от положения 7816 до положения 7930 в SEQ ID NO: 19 или нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с указанной последовательностью. Мутантный аллель может представлять собой нокаут-аллель или аллель замещения с более низкой активностью, чем аллель дикого типа. В соответствии с одним из вариантов осуществления изобретения уменьшается количество белка APO1-7A или уменьшается количество белка APO1-7A и белка APO1-7B, или уменьшается количество белка APO1-7A и белка APO1-7D, или уменьшается количество белков APO1-7A, APO1-7B и APO1-7D.

[65] «Растение пшеницы с изначально низким количеством колосков на колосе» означает растение пшеницы из сорта, для которого среднее количество колосков на колос составляет приблизительно менее 23, приблизительно менее 22, приблизительно менее 21, приблизительно менее 20, приблизительно менее 19 или приблизительно менее 18 колосков на колос. Среднее количество колосков на колос для указанного сорта может быть в диапазоне приблизительно 17 - 23, приблизительно 17 - 22, приблизительно 17 - 21, приблизительно 17 - 20, приблизительно 17 - 19, приблизительно 17 - 18, приблизительно 18 - 23, приблизительно 18 - 22, приблизительно 18 - 21, приблизительно 18 - 20, приблизительно 18 - 19, приблизительно 19 - 23, приблизительно 19 - 22, приблизительно 19 - 21, приблизительно 19 - 20, приблизительно 20 - 23, приблизительно 20 - 22, приблизительно 20 - 21, приблизительно 21 - 23, приблизительно 21 - 22, приблизительно 22 - 23 колосков на колос.

[66] «Растение пшеницы с изначально высоким количеством колосков на колосе» означает растение пшеницы из сорта, для которого среднее количество колосков на колос составляет, по меньшей мере, приблизительно 23, по меньшей мере, приблизительно 24, по меньшей мере, приблизительно 25 или, по меньшей мере, приблизительно 26, по меньшей мере, приблизительно 27, по меньшей мере, приблизительно 28 или, по меньшей мере, приблизительно 29 или, по меньшей мере, приблизительно 30 колосков на колос. Среднее количество колосков на колос для указанного сорта может быть в диапазоне приблизительно 23 - 30, приблизительно 24 - 30, приблизительно 25 - 30, приблизительно 26 - 30, приблизительно 27 - 30, приблизительно 28 - 30, приблизительно 29 - 30, приблизительно 23 - 29, приблизительно 24 - 29, приблизительно 25 - 29, приблизительно 26 - 29, приблизительно 27 - 29, приблизительно 28 - 29, приблизительно 23 - 28, приблизительно 24 - 28, приблизительно 25 - 28, приблизительно 26 - 28, приблизительно 27 - 28, приблизительно 23 - 27, приблизительно 24 - 27, приблизительно 25 - 27, приблизительно 26 - 27, приблизительно 23 - 26, приблизительно 24 - 26, приблизительно 25 - 26, приблизительно 23 - 25, приблизительно 24 - 25 или приблизительно 23 - 24 колосков на колос.

[67] При использовании по тексту настоящего документа термин «изменение количества колосков на колос» означает значительное увеличение или значительное уменьшение среднего количества колосков на колос растения пшеницы.

[68] Увеличение количества колосков на колос означает увеличение, по меньшей мере, приблизительно на 1, по меньшей мере, приблизительно на 2, по меньшей мере, приблизительно на 3, по меньшей мере, приблизительно на 5 колосков на колос по сравнению с количеством колосков на колосе растения пшеницы, в частности, в случае, если растение пшеницы изначально имеет низкое количество колосков на колосе.

[69] Уменьшение количества колосков на колос означает уменьшение, по меньшей мере, приблизительно на 3, по меньшей мере, приблизительно на 2, по меньшей мере, приблизительно на 1, по меньшей мере, приблизительно на 5 колосков на колос по сравнению с количеством колосков на колосе растения пшеницы, в частности, в случае, если растение пшеницы изначально имеет высокое количество колосков на колосе.

[70] При использовании по тексту настоящего документа термин «изменение количества белка» означает (значительное) увеличение или (значительное) уменьшение количества белка в соответствии с описанием в настоящем документе.

[71] Увеличение означает увеличение на, по меньшей мере, 10%, по меньшей мере, 20%, по меньшей мере, 30%, по меньшей мере, 40%, 50%, по меньшей мере, 60%, по меньшей мере, 70%, по меньшей мере, 80%, по меньшей мере, 90%, по меньшей мере, 100% по сравнению с количеством белка, продуцируемого клеткой растения пшеницы, в частности, растения пшеницы, имеющего изначально низкое количество колосков на колос.

[72] Уменьшение означает уменьшение на, по меньшей мере, 10%, по меньшей мере, 20%, по меньшей мере, 25%, по меньшей мере, 30%, по меньшей мере, 35%, по меньшей мере, 40%, по меньшей мере, 45% или, по меньшей мере, 50% по сравнению с количеством белка, продуцируемого клеткой растения пшеницы, в частности, растения пшеницы, имеющего изначально высокое количество колосков на колос.

[73] В соответствии с одним из вариантов осуществления изобретения уменьшение экспрессии и/или активности гена и/или белка АРО1 может осуществляться путем уменьшения количества продуцируемого функционального белка АРО1. Такое уменьшение может представлять собой уменьшение на, по меньшей мере, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% или 100% (т.е. клетка не продуцирует функциональный белок АРО1) по сравнению с количеством функционального белка АРО1, продуцируемым клеткой с уровнями экспрессии и активностью АРО1 дикого типа. Указанное уменьшение экспрессии и/или активности может относиться к конститутивному уменьшению количества функционального продуцируемого белка АРО1. Указанное уменьшение также может представлять собой временное/индуцируемое уменьшение количества функционального продуцируемого белка АРО1.

[74] Снижение экспрессии и/или активности гена АРО1 по изобретению также может быть достигнуто с использованием молекулы РНК, которая обеспечивает снижение экспрессии и/или активности гена АРО1. Молекула РНК, которая обеспечивает снижение экспрессии и/или активности гена и/или белка АРО1, может представлять собой РНК, кодирующую белок, который подавляет экспрессию и/или активность указанного белка АРО1. Кроме того, указанная молекула РНК, которая обеспечивает снижение экспрессии и/или активности гена и/или белка АРО1, также может представлять собой молекулу РНК, которая подавляет экспрессию гена, который является активатором экспрессии и/или активности указанного белка АРО1. Указанная молекула РНК, которая подавляет экспрессию и/или активность гена и/или белка АРО1, также может представлять собой молекулу РНК, которая непосредственно подавляет экспрессию и/или активность гена и/или белка АРО1, например, РНК, которая медирует сайленсинг указанного гена АРО1.

[75] Экспрессия и/или активность гена и/или белка АРО1 может быть с легкостью снижена или ликвидирована посредством транскрипционного или посттранскрипционного сайленсинга экспрессии эндогенных генов АРО1. С этой целью молекула сайленсинг-РНК может быть введена в клетки растения с нацеливанием на эндогенные гены, кодирующие АРО1. При использовании по тексту настоящего документа термин «сайленсинг-РНК» или «молекула сайленсинг-РНК»

относится к любой молекуле РНК, которая при введении в клетку растения снижает экспрессию целевого гена.

[76] Сайленсинг-РНК также может представлять собой искусственные молекулы микро-РНК, как описано, например, в WO2005/052170, WO2005/047505 или US 2005/0144667, или транс-активирующие малые интерферирующие РНК, как описано в WO2006/074400 (все указанные документы включены в настоящий документ посредством ссылки). В некоторых вариантах осуществления нуклеиновая кислота, экспрессируемая химерным геном по изобретению, представляет собой каталитическую РНК или обладает рибозимной активностью, специфичной для целевой последовательности. Таким образом, полинуклеотид вызывает деградацию эндогенной матричной РНК, транскрибируемой от целевого гена/последовательности, что приводит к снижению экспрессии белка, присутствующего в растении. В соответствии с одним из вариантов осуществления изобретения нуклеиновая кислота, экспрессируемая химерным геном по изобретению, кодирует белок «цинковые пальцы», который связывается с геном, кодирующим указанный белок, что приводит к пониженной экспрессии целевого гена. В соответствии с частными вариантами осуществления изобретения белок «цинковые пальцы» связывается с регуляторной областью указанного гена. В соответствии с другими вариантами осуществления изобретения белок «цинковые пальцы» связывается с матричной РНК, кодирующей указанный белок, и предотвращает ее трансляцию.

[77] В соответствии с альтернативными вариантами осуществления изобретения уменьшение экспрессии и/или активности гена и/или белка APO1 может обеспечиваться путем подавления экспрессии указанного белка APO1 в растении. Подавление экспрессии указанного гена и/или белка APO1 может быть индуцировано в необходимый момент с использованием аэрозоля (системное применение) ингибирующих нуклеиновых кислот, таких как молекулы РНК или ДНК, которые функционируют в РНК-опосредованном геном сайленсинге в соответствии с описанием, например, в документе WO2011/112570 (включен в настоящий документ посредством ссылки).

[78] В одном варианте осуществления изобретения повышение урожайности может быть обеспечено, когда растения пшеницы с меньшим количеством колосков на колос (аллельная форма Apo1-7A SPS-) выращиваются в определенной среде, при этом те же растения при выращивании в другой среде могут демонстрировать увеличение урожайности при наличии большего количества колосков на колос (аллельная форма Apo1-7A SPS+). При том, что эффекты урожайности могут быть различными в разных условиях выращивания, эффекты SPS одинаковы для разных сред. Такое изменение по уровню в разных средах (в данном случае в отношении урожайности) называется взаимодействием «влияние окружающей среды на генотип» (GxE) и является основным ограничением генетического эффекта в сельскохозяйственных культурах. Идентифицировав лежащий в основе ген, можно использовать соответствующий аллель для любой целевой среды.

[79] SEQ ID NO: 4 представляет собой нуклеотидную последовательность не кодирующей ДНК выше 3'-конца Apo1 длиной приблизительно 5 т.п.о. из сорта пшеницы Westonia. SEQ ID NO: 5 представляет собой нуклеотидную последовательность не кодирующей ДНК выше 5'-конца Apo1 длиной приблизительно 5 т.п.о. из сорта пшеницы Baxter. SEQ ID NO: 4 и SEQ ID NO: 5 являются функциональными вариантами и имеют 99% идентичности последовательности. SEQ ID NO: 9 представляет собой нуклеотидную последовательность соответствующей не кодирующей ДНК выше 5'-конца Apo1 из сорта пшеницы Chara. Сорт Yitpi содержит соответствующую не кодирующую ДНК выше 5'-конца Apo1 с нуклеотидной последовательностью, идентичной SEQ ID NO: 9. SEQ ID NO: 19 представляет собой нуклеотидную последовательность не кодирующей ДНК выше 5'-конца Apo1 длиной приблизительно 8 т.п.о. из сорта пшеницы Chinese Spring на хромосоме 7A. Сорт Robigus содержит соответствующую не кодирующую ДНК выше 5'-конца Apo1 с нуклеотидной последовательностью SEQ ID NO: 19, с делецией нуклеотидов от положения 7816 до 7930 SEQ ID NO: 19 и вставкой приблизительно 5 - 7,7 т.п.о нуклеотидов в нуклеотидном положении 901 на SEQ ID NO: 19 (в частности, между нуклеотидным положением 900 и нуклеотидным положением 901 SEQ ID NO: 19 – см. первый доп\_признак в SEQ ID NO: 19). Кроме того, сорт Robigus имеет те же SNP и инсерционно-делеционные полиморфизмы, что и сортов Yitpi/Chara в Таблице 2, а

сорт Claire имеет те же SNP и инсерционно-делеционные полиморфизмы, что и Westonia в Таблице 2.

[80] Нуклеиновая кислота, содержащая нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с SEQ ID NO: 4, 5, 9 или 19, таким образом может представлять собой нуклеиновую кислоту, содержащую нуклеотидную последовательность, имеющую, по меньшей мере, 90%, или, по меньшей мере, 95%, или, по меньшей мере, 98%, или, по меньшей мере, 99% или 100% идентичности последовательности с SEQ ID NO: 4, 5, 9 или 19, соответственно. Нуклеотидная последовательность, имеющая 100% идентичности последовательности с SEQ ID NO: 4, 5 или 9, также именуется нуклеотидной последовательностью, идентичной SEQ ID NO: 4, 5 или 9, соответственно. Нуклеотидная последовательность SEQ ID NO: 9 имеет 97% идентичности с нуклеотидной последовательностью SEQ ID NO: 4 или 5, но не содержит нуклеотидную последовательность от положения 4399 до положения 4513 SEQ ID NO: 5, или нуклеотидную последовательность от положения 4401 до положения 4516 SEQ ID NO: 4.

[81] В еще одном варианте осуществления изобретения этап получения растения в вышеописанных способах может означать получение растения путем трансформации, скрещивания, обратного скрещивания, интрогрессии, редактирования генома или мутагенеза.

[82] Термин «получение растения» может относиться к введению молекулы экзогенной ДНК в клетку растения путем трансформации, при необходимости с последующей регенерацией растения из трансформированной клетки растения. Этот термин может также относиться к введению молекулы рекомбинантной ДНК путем скрещивания трансгенного растения, содержащего молекулу рекомбинантной ДНК, с другим растением и отбора растений-потомков, которые унаследовали рекомбинантную молекулу ДНК или трансген. Еще одно альтернативное значение термина «получение растения» относится к введению рекомбинантной молекулы ДНК с использованием таких методов, как слияние протопластов, при необходимости с последующей регенерацией растения из слитых протопластов.

[83] Очевидно, что используемые методы трансформации имеют второстепенное значение для настоящего изобретения. В настоящее время трансформация растений является широко распространенной технологией. Преимущественно, любой из способов трансформации может использоваться для внедрения интересующей нуклеиновой кислоты/гена в подходящую клетку-предка. Способы трансформации включают использование липосом, электропорацию, химикаты, увеличивающие поглощение свободной ДНК, ввод ДНК непосредственно в растение (клетку растения), например, путем микроинъекции, обстрел из генной пушки, трансформацию с использованием вирусов, пыльцы или микропроекции. Способы могут быть выбраны из способов использования кальция/полиэтиленгликоля для протопластов (Krens et al. (1982) *Nature* 296: 72-74; Negrutiu et al. (1987) *Plant. Mol. Biol.* 8: 363-373); электропорация протопластов (Shillito et al. (1985) *Bio/Technol.* 3: 1099-1102); микроинъекция в растительный материал (Crossway et al. (1986) *Mol. Gen. Genet.* 202: 179-185); обстрел частицами, покрытыми ДНК или РНК (Klein et al. (1987) *Nature* 327: 70) заражение (неинтегративными) вирусами и другие подобные способы.

[84] Способы трансформации растений пшеницы также известны специалистам. Для различных зерновых культур могут быть созданы различные системы трансформации: электропорация тканей, трансформация протопластов и перенос ДНК путем бомбардировки частицами в регенерируемых тканях и клетках (for an overview see Jane, *Euphytica* 85 (1995), 35-44). Трансформация пшеницы неоднократно описана в литературе (for an overview see Maheshwari, *Critical Reviews in Plant Science* 14 (2) (1995), 149-178, Nehra et al., *Plant J.* 5 (1994), 285-297). Эффективный метод агробактериальной трансформации, описанный Ishida et al. 2015 *Agrobacterium protocols*: том 1, *Methods in Molecular Biology*, том 1223: 189-198.

[85] При использовании по тексту настоящего документа термин «мутагенез» относится к процессу, при котором клетки растений (например, множество семян или других частей пшеницы) подвергаются воздействию с использованием техники, вызывающей мутации в ДНК клеток, такой, например, как контакт с мутагенным агентом, таким как, например, химическое вещество (например, этилметилсульфонат (EMS), этилнитрозомочевина (ENU), и т.д.) или ионизирующее излучение (нейтронами (например, мутагенез с использованием быстрых нейтронов и т.д.),

альфа-излучение, гамма-излучение (например, излучение, которое генерируется источником с кобальтом 60), рентгеновское излучение, ультрафиолетовое излучение и т.д.), Т-ДНК-инсерционный мутагенез (Azpiroz-Leehan et al. (1997) Trends Genet 13:152-156), транспозонный мутагенез (McKenzie et al. (2002) Theor Appl Genet 105:23-33), или мутагенез в культуре ткани (индукция соматональных вариаций), или комбинация двух или нескольких таких техник. Таким образом, необходимый мутагенез одного или нескольких аллелей *АРО1* может обеспечиваться с использованием одного из вышеуказанных методов. В то время как мутации, полученные с использованием облучения, зачастую вызывают большие делеции или другие обширные поражения, такие как транслокации или сложные перестройки, мутации, полученные с использованием химических мутагенов, зачастую обеспечивают местные поражения, такие как точечные мутации. Например, EMS алкилирует гуаниновые основания, что приводит к ошибочному спариванию оснований: алкилированный гуанин спаривается с тиминным основанием, что обычно приводит к транзициям G/C в A/T. После мутагенеза осуществляется получение растений пшеницы из обработанных клеток с использованием известных техник. Например, может осуществляться посев полученных семян растений пшеницы в соответствии с традиционными способами выращивания растений, после самоопыления на таких растениях образуются семена. Может осуществляться сбор дополнительных семян, которые образуются в результате такого самоопыления в настоящем поколении или в будущих поколениях, а также их скриннинг на присутствие мутантных *apo1* аллелей. Известны различные техники скриннинга для получения специфических мутантных аллелей, например, в методе Deleteagene™ (Delete-a-gene; Li et al., 2001, Plant J 27: 235-242) используют анализ на основе полимеразной цепной реакции (ПЦР) для скриннинга делеционных мутантов, которые получены путем мутагенеза с использованием быстрых нейтронов, в методе TILLING (поиск индуцированных локальных нарушений в геномах; McCallum et al., 2000, Nat Biotechnol 18:455-457) осуществляют идентификацию EMS-индуцированных точечных мутаций и т.д.

[86] Термин «генный таргетинг» в настоящем документе означает к направленной модификации генов, в которой используется такие механизмы, как гомологическая

рекомбинация, исправление ошибок спаривания или сайт-специфический мутагенез. Этот способ может использоваться для замены, вставки и делеции эндогенных последовательностей или последовательностей, которые присутствуют или были ранее введены в клетки растения. Способы генного таргетинга описаны, например, в WO 2006/105946 или WO2009/002150. Для создания мутантных или искусственных *apo1* аллелей может использоваться генный таргетинг.

[87] Генный таргетинг также может использоваться для создания новых гаплотипов или блоков гаплотипов. Например, блоки гаплотипов, содержащие ген APO1 на хромосоме 7A, которые могут использоваться для повышения урожайности несколькими способами, но которые при этом содержат вышележащую делецию и/или вставку, которые связаны с низкими числами SPS, с использованием генного таргетинга могут быть модифицированы для замещения такой вышележащей делеции и/или вставки.

[88] При использовании по тексту настоящего документа выражение «дикого типа» относится к типичной форме растения или гена, которая наиболее распространена в природе. Термин «растение дикого типа» относится к растению с наиболее распространенным фенотипом такого растения в природной популяции. Термин «аллель дикого типа» относится к аллелю гена, который необходим для получения фенотипа дикого типа. Напротив, термин «мутантное растение» относится к редкому, отличающемуся фенотипу такого растения, которое было получено путем вмешательства человека, например, путем мутагенеза, а термин «мутантный аллель» относится к аллелю гена, который требуется для получения мутантного фенотипа.

[89] При использовании по тексту настоящего документа термин «мутант» относится к форме растения или гена, которая отличается от такого растения или гена в природной популяции, и которая получена в результате вмешательства человека, например путем мутагенеза, а термин «мутантный аллель» относится к аллелю, который не обнаруживается у растений в природной популяции или в селекционной популяции, но который образуется в результате вмешательства человека, например, в результате мутагенеза или генного таргетинга.

[90] При использовании по тексту настоящего документа термин «аллель дикого типа» (например, «аллель *APO1* дикого типа») означает природный аллель, который находится в растениях, в частности, в растениях пшеницы, который кодирует функциональный белок (например, функциональный белок *APO1*). Напротив, при использовании по тексту настоящего документа термин «мутантный аллель» (например, «мутантный аллель *apo1*») относится к аллелю, который не кодирует функциональный белок, т.е. к аллелю *apo1*, кодирующему нефункциональный белок *APO1*, т.е. термин «нефункциональный белок *APO1*» при использовании по тексту настоящего документа относится к белку *APO1* без биологической активности или со существенно сниженной биологической активностью по сравнению с соответствующим функциональным белком *APO1* дикого типа, или к аллелю *apo1*, который совсем не кодирует белок *APO1*.

[91] При использовании по тексту настоящего документа термин «нокаут-аллель» или «инактивированный аллель» относится к мутантному аллелю, который кодирует белок без биологической активности по сравнению с соответствующим функциональным белком дикого типа, или к аллелю, который совсем не кодирует белок. Термин «полностью инактивированный мутантный аллель» означает, например, аллель дикого типа, который содержит одну или несколько мутаций в нуклеотидной последовательности, например, одну или несколько нонсенс-мутаций или миссенс-мутаций. В частности, такой полностью инактивированный мутантный аллель *apo1* представляет собой аллель *APO1* дикого типа, который содержит мутацию, которая предпочтительно приводит к продуцированию белка *APO1*, лишенного, по меньшей мере, одного функционального домена, такого как F-box домен, или лишенного, по меньшей мере, одной аминокислоты, критически важной для его функции таким образом, чтобы биологическая активность белка *APO1* полностью ликвидируется, или мутацию, которая предпочтительно приводит к тому, что белок *APO1* не продуцируется.

[92] При использовании по тексту настоящего документа термин «частично инактивированный мутантный аллель» означает, например, мутантный аллель, который кодирует белок со существенно сниженной биологической активностью по сравнению с соответствующим функциональным белком дикого типа. Термин

«частично инактивированный мутантный аллель» означает, например, аллель дикого типа, который содержит одну или несколько мутаций в нуклеотидной последовательности, например, одну или несколько миссенс-мутаций. В частности, такой частично инактивированный мутантный аллель представляет собой аллель дикого типа, который содержит мутацию, которая предпочтительно приводит к продуцированию белка, в котором, по меньшей мере, одна консервативная и/или функциональная аминокислота заменена другой аминокислотой таким образом, чтобы его биологическая активность была значительно снижена, но полностью не ликвидирована.

[93] Специалист сможет определить уровень экспрессии гена, например, с использованием анализа накопления РНК, продуцируемой из нуклеиновой кислоты. Степень накопления РНК или уровни РНК, например, мРНК, могут измеряться в одной временной точке или в нескольких временных точках, в одной ткани или в нескольких тканях, поэтому кратное увеличение может представлять собой среднее значение кратного увеличения или экстраполированное значение, полученное из экспериментально полученных значений. Уровень экспрессии может быть определен такими методами, как ОТ-кПЦР, или с использованием микрочипов на основе гибридизации. Уровень экспрессии также может оцениваться путем полного транскриптомного секвенирования методом дробовика с использованием секвенирования следующего поколения для выявления присутствия и количества РНК, которая может быть выбрана для полиаденилирования РНК или для деплеции рибосомальной РНК.

[94] В некоторых вариантах осуществления этап модификации эндогенного гена *Arp1* может включать нуклеотидные модификации в эндогенном гене *Arp1* для увеличения или уменьшения SPS в растении.

[95] В некоторых вариантах осуществления растений или способов, описанных в настоящем документе, эндогенный ген *Arp1* может быть модифицирован путем редактирования генома. В некоторых вариантах осуществления редактирование генома может осуществляться с помощью одной или нескольких нуклеаз, созданных методами генной инженерии, выбранных из группы, состоящей из РНК-направляемых

нуклеаз, мегануклеаз, нуклеаз «цинковые пальцы» (ZFN) и нуклеаз, которые основаны на эффекторах, подобных активаторам транскрипции (TALEN).

[96] В некоторых вариантах осуществления этап получения растения может включать следующие стадии: получение растения дикого типа; модификация эндогенного гена *Aro1* в растении путем редактирования генома для получения растения, содержащего нуклеиновую кислоту в соответствии с определением в настоящем документе.

[97] Термин «редактирование генома» или «редактирование генома с помощью нуклеаз, созданных методами генной инженерии» обычно относится к типу генной инженерии, при котором осуществляется вставка, делеция или замена ДНК в геноме живого организма с использованием нуклеаз (созданных методами генной инженерии). Такие нуклеазы создают сайт-специфические, например, двухцепочечные разрывы (DSB), в определенных положениях в геноме.

[98] В некоторых вариантах осуществления эндогенный ген *Aro1* может быть модифицирован путем создания сайт-специфических, например, двухцепочечных разрывов (DSB), в одном или нескольких определенных положениях в геноме. Индуцированные двухцепочечные разрывы могут быть восстановлены посредством негомологичного соединения концов (NHEJ) или гомологически направленной репарации (HDR).

[99] В некоторых вариантах осуществления эндогенный ген *Aro1* может быть модифицирован способом редактирования генома, то есть способом модификации генома, предпочтительно ядерного генома, клетки растения в заранее выбранном сайте, причем указанный способ включает следующие этапы:

- введение двухцепочечного разрыва ДНК (DSB) в геном указанной клетки в сайте расщепления на участке узнавания (или рядом с ним) для фермента, индуцирующего двухцепочечный разрыв ДНК (DSBI), путем экспрессии в указанной клетке фермента DSBI, распознающего указанный сайт распознавания и индуцирующего указанный DSB в указанном сайте расщепления;
- введение в указанную клетку молекулы репаративной нуклеиновой кислоты, содержащей вышележащий фланкирующий участок гомологичный участку ДНК выше указанного заранее выбранного сайта и/или нижележащий фланкирующий

участок ДНК гомологичный участку ДНК ниже указанного заранее выбранного сайта для обеспечения гомологичной рекомбинации между указанным фланкирующим участком или участками и указанным участком или участками ДНК, фланкирующими указанный заранее выбранный сайт; и

- выбор клетки, в которой указанная молекула репаративной нуклеиновой кислоты использована в качестве матрицы для модификации указанного генома в указанном заранее выбранном сайте.

[100] при этом указанная модификация выбрана из замены, по меньшей мере, одного нуклеотида, делеции, по меньшей мере, одного нуклеотида, вставки, по меньшей мере, одного нуклеотида или любой их комбинации.

[101] При использовании по тексту настоящего документа термин «фермент, индуцирующий разрыв двухцепочечной ДНК» означает фермент, способный индуцировать разрыв двухцепочечной ДНК в определенной нуклеотидной последовательности, называемой «сайтом узнавания».

[102] Эндонуклеазы с редким расщеплением – это ферменты DSBI с сайтом узнавания длиной приблизительно 14 - 70 последовательных нуклеотидов, следовательно, они имеют очень низкую частоту расщепления даже в крупных геномах, таких как большинство геномов растений. «Хоуминг-эндонуклеазы, также называемые мегануклеазами» составляют семейство таких эндонуклеаз с редким расщеплением. Они могут кодироваться интронами, независимыми генами или вставочными последовательностями и демонстрируют поразительные структурные и функциональные свойства, которые отличают их от более классических рестрикционных ферментов, обычно от бактериальных систем рестрикции-модификации типа II. Их сайты распознавания имеют общую асимметрию, которая противоположна характерной диадной симметрии большинства сайтов распознавания рестриктаз. Было показано, что несколько хоуминг-эндонуклеаз, кодируемых интронами или интеинами, способствуют хоумингу их соответствующие генетические элементы в аллельные сайты без интронов или без интеинов. Путем сайт-специфического двухцепочечного разрыва в аллелях без интронов или без интеинов, эти нуклеазы создают рекомбиногенные концы, которые участвуют в процессе

конверсии генов, который дублирует кодирующую последовательность и приводит к вставке интрона или вставочной последовательности на уровне ДНК.

[103] Перечень других мегануклеаз с редким расщеплением и их соответствующих сайтов узнавания представлен в Таблице I WO03/004659 (стр. 17 – 20) (включен в настоящий документ посредством ссылки). Они включают I-Sce I, I-Chu I, I-Dmo I, I-Cre I, I-Csm I, PI-Fli I, Pt-Mtu I, I-Ceu I, I-Sce II, I-Sce III, HO, PI-Civ I, PI-Ctr I, PI-Aae I, PI-BSU I, PI-DhaI, PI-Dra I, PI-Mav I, PI-Mch I, PI-Mfu I, PI-Mfl I, PI-Mga I, PI-Mgo I, PI-Min I, PI-Mka I, PI-Mle I, PI-Mma I, PI-Msh I, PI-Msm I, PI-Mth I, PI-Mtu I, PI-Mxe I, PI-Npu I, PI-Pfu I, PI-Rma I, PI-Spb I, PI-Ssp I, PI-Fac I, PI-Mja I, PI-Pho I, PI-Tag I, PI-Thy I, PI-Tko I или PI-Tsp I.

[104] Кроме того, имеются способы конструирования «индивидуально подобранных эндонуклеаз с редким расщеплением», которые распознают практически любую определенную целевую нуклеотидную последовательность. В немногих словах, химерные рестрикционные ферменты могут быть получены с использованием гибридов между цинк-пальцевым доменом, сконструированным, чтобы распознавать специфическую нуклеотидную последовательность, и доменом неспецифического расщепления ДНК из природного рестрикционного фермента, такого как FokI. Такие способы описаны, например, в документах WO 03/080809, WO94/18313 или WO95/09233 и в Isalan et al., 2001, Nature Biotechnology 19, 656- 660; Liu et al. 1997, Proc. Natl. Acad. Sci. USA, 94, 5525-5530).

[105] Индивидуально подобранные мегануклеазы могут быть получены путем отбора из библиотеки вариантов в соответствии с описанием в документе WO2004/067736. Индивидуально подобранные мегануклеазы с измененной специфичностью последовательности и сродством к связыванию с ДНК могут быть получены также путем рационального дизайна в соответствии с описанием в документе WO2007/047859.

[106] Другие примеры индивидуально подобранных эндонуклеаз включают так называемые нуклеазы TALE (TALEN), которые основаны на эффекторах, подобных активаторам транскрипции (TALE), из рода бактерий *Xanthomonas*, которые гибридизированы с каталитическим доменом нуклеазы (например, FOKI).

Специфичность ДНК-связывания этих TALE определяется парами высоковариабельных аминокислотных остатков (RVD) из тандемно расположенных единиц повтора аминокислот 34/35, так что один RVD специфически распознает один нуклеотид в ДНК-мишени. Единицы повтора могут быть собраны для распознавания практически любых целевых последовательностей и гибридизированы с каталитическим доменом нуклеазы для получения эндонуклеаз, специфических к последовательности (смотрите, например, Boch et al., 2009, Science, 326:p1509-1512; Moscou and Bogdanove, 2009, Science, 326:p1501; Christian et al., 2010, Genetics, 186:p757-761; и WO10/079430, WO11/072246, WO2011/154393, WO11/146121, WO2012/001527, WO2012/093833, WO2012/104729, WO2012/138927, WO2012/138939). В документе WO2012/138927 также описаны мономерные (компактные) TALEN и TALEN с различными каталитическими доменами, а также их комбинации.

[107] Описан новый тип настраиваемой эндонуклеазной системы – так называемая «система CRISPR/Cas» (короткие палиндромные повторы, регулярно расположенные группами), в которой используется особая молекула РНК (сгРНК), придающая специфичность к последовательности для расщепления ассоциированной РНК-направляемой эндонуклеазы. Такие специально разработанные эндонуклеазы с редким расщеплением также называются неприродными эндонуклеазами с редким расщеплением.

[108] При использовании по тексту настоящего документа термин «РНК-направляемая нуклеаза» или «РНК-направляемая эндонуклеаза» (RGEN) означает РНК-управляемый ДНК-модифицирующий полипептид, обладающий (эндо)нуклеазной активностью.

[109] RGEN обычно получают из систем коротких палиндромных повторов, регулярно расположенных группами (CRISPR), которые представляют собой широко распространенный класс бактериальных систем для защиты от чужеродных нуклеиновых кислот. Системы CRISPR обнаружены в широком диапазоне организмов эубактерий и архей. CRISPR-системы включают системы подтипа I, II, III и V (смотрите, например, WO2007025097; WO2013098244; WO2014022702;

WO2014093479; WO2015155686; EP3009511; US2016208243). В системах CRISPR/Cas II дикого типа используется РНК-направляемая нуклеаза, например, Cas9, в комплексе с направляющей и активирующей РНК для распознавания и расщепления чужеродной нуклеиновой кислоты (Jinek et al., 2012, *Science*, 337(6096):816-21).

[110] Гомологи Cas9 обнаружены у самых разных зубактерий, включая, помимо прочего, бактерии следующих таксономических групп: Actinobacteria, Aquificae, Bacteroidetes-Chlorobi, Chlamydiae-Verrucomicrobia, Chloflexi, Cyanobacteria, Firmicutes, Proteobacteria, Spirochaetes и Thermotogae. Примером белка Cas9 является белок Cas9 *Streptococcus pyogenes*. Другие белки Cas9, их гомологи и варианты, а также методы их использования при редактировании генома описаны, например, в Chylinski, et al., 2013, *RNA Biol.*, 10(5): 726-737; Makarova et al., 2011, *Nat. Rev. Microbiol.*, 9(6): 467-477; Hou, et al., 2013, *Proc Natl Acad Sci U S A*, 110(39):15644-9; Sampson et al., 2013, *Nature*, 497(7448):254-7; Jinek, et al., 2012, выше; WO2013142578; WO2013176772; WO2014065596; WO2014089290; WO2014093709; WO2014093622; WO2014093655; WO2014093701; WO2014093712; WO2014093635; WO2014093595; WO2014093694; WO2014093661; WO2014093718; WO2014093709; WO2014099750; WO2014113493; WO2014190181; WO2015006294; WO2015071474; WO2015077318; WO2015089406; WO2015103153; WO201621973; WO201633298; WO201649258, все включенные в настоящий документ посредством ссылок.

[111] Другие РНК-направляемые нуклеазы включают, например, Cpf1 и ее гомологи и варианты (как описано, например, в Zetsche et al., 2015, *Cell*, Том 163, выпуск 3, 759-771; EP3009511; US2016208243; Kleinstiver et al., 2016, *Nat Biotechnol.*, 34(8):869-74; Gao et al., 2016, *Cell Res.*, 6(8):901-13; Hur et al., 2016, *Nat Biotechnol.*, 34(8):807; Kim et al., 2016, *Nat Biotechnol.*, 34(8):863-8.; Yamano et al., 2016, *Cell*, 165(4):949-62), а также C2c1 и C2c3 (Shmakov et al., 2015, *Mol Cell.*, 60(3):385-97), все включенные в настоящий документ посредством ссылок.

[112] Другие РНК-направляемые нуклеазы могут включать Argonaut-подобные белки, в соответствии с описанием, например, как в WO2015157534.

[113] Другие РНК-направляемые нуклеазы и другие полипептиды описаны в WO2013088446.

[114] В соответствии с одним из вариантов осуществления изобретения RGEN также может быть РНК-управляемым ник-ферментом (никазой) или парой РНК-управляемых ник-ферментов, каждый из которых вводит разрыв только в одну цепь двухцепочечной ДНК в заранее выбранном сайте или рядом с ним. Из пары никаз один фермент вносит разрыв в одну цепь ДНК в заранее выбранном сайте или рядом с ним, а другой фермент вносит разрыв в другую цепь ДНК в заранее выбранном сайте или рядом с ним. Два одноцепочечных разрыва могут быть введены в одно и то же положение нуклеотида на обеих цепях, что приводит к двухцепочечному разрыву ДНК с тупым концом, но два одноцепочечных разрыва также могут быть введены в разные положения нуклеотидов в каждой цепи, в результате чего получается 5' или 3' выступ в месте разрыва («липкие концы» или «ступенчатый разрыв»). Ник-мутанты (мутанты с односторонним разрывом) и способы их использования, например, описаны в вышеуказанных документах и, в частности, в WO2014191518, WO2014204725 и WO201628682. Кроме того, ник-мутант с разрывом только в одной из двух цепей ДНК (то есть односторонний разрыв ДНК), может усилить гомологически направленную репарацию (HDR) с помощью донорного полинуклеотида (Richardson et al. 2016, Nature Biotechnology 34, 339-344; US62/262,189).

[115] В качестве альтернативы нуклеазы или никазы для увеличения направленного введения донорного полинуклеотида также могут быть использованы варианты описанных выше нуклеаз, дефицитные по нуклеазе (также называемые «мертвыми» или каталитически неактивными), такие как dCas9, в соответствии с описанием, например, в Richardson et al. 2016, Nature Biotechnology 34, 339-344; US62/262,189). Такие варианты не обладают способностью расщеплять ДНК или производить односторонний разрыв ДНК, но могут быть нацелены на ДНК и связывать ее (смотрите, например, WO2013176772, EP3009511). Считается, что такие «мертвые» нуклеазы вызывают замещение цепи путем связывания с одной из двух цепей («плавление ДНК»), тем самым усиливая рекомбинацию с донорным полинуклеотидом путем соединения донорского полинуклеотида с другой «свободной» цепью ДНК.

[116] Ник-мутанты описаны для различных RGEN и включают одну или несколько мутаций в каталитическом домене, таком как домены HNH и RuvC (например, Cas9) RuvC-подобного домена (например, Cpf1). Например, SpCas9 может быть

преобразована в никазу путем мутации D10A в RuvC, а 863A в нуклеазном домене HNH преобразует SpCas9 в ДНК-никазу, в то время как инактивация обоих нуклеазных доменов приводит к образованию каталитически неактивного белка (Jinek et al., 2012, выше, Gasiunas et al., 2012, Proc. Natl. Acad. Sci. USA 109, E2579-E2586). В случае Cpf1 было обнаружено, что мутация D917A, а также мутация E1006A полностью инактивировали активность расщепления ДНК FnCpf1, а D1255A значительно снижал нуклеолитическую активность (Zetsche et al., 2015, выше). Соответствующие остатки других вариантов RGEN (например, Cas9 или Cpf1) могут быть определены оптимальным выравниванием.

[117] Сайт расщепления фермента DSBI относится к точному месту на ДНК, где индуцируется двухцепочечный разрыв ДНК. Сайт расщепления может входить в сайт узнавания фермента DSBI (перекрываться сайтом узнавания фермента DSBI), а может и не входить в сайт узнавания фермента DSBI, и, следовательно, считается, что сайт расщепления фермента DSBI расположен на его сайте узнавания или рядом с ним. Сайт узнавания фермента DSBI, также иногда называемый сайтом связывания, представляет собой нуклеотидную последовательность, которая (специфически) распознается ферментом DSBI и определяет его специфичность связывания. Например, мономер TALEN или ZNF имеет сайт узнавания, который определяется их повторами RVD или ZF соответственно, тогда как сайт расщепления мономера определяется его нуклеазным доменом (например, FOKI) и обычно находится за пределами сайта узнавания. В случае димерных TALEN или ZFN сайт расщепления расположен между двумя сайтами узнавания/связывания соответствующих мономеров, причем эта промежуточная область ДНК, где происходит расщепление, называется спейсерной областью. Для мегануклеаз, с другой стороны, расщепление ДНК происходит в пределах ее специфической связывающей области, и, следовательно, сайт связывания и сайт расщепления перекрываются.

[118] Специалист сможет выбрать фермент DSBI, распознающий определенный сайт узнавания и индуцирующий DSB в сайте расщепления в заранее выбранном сайте или рядом с ним, или может сконструировать такой фермент DSBI. В качестве альтернативы сайт узнавания фермента DSBI может быть введен в целевой геном с использованием любого обычного метода трансформации или путем скрещивания с

организмом, имеющим сайт узнавания фермента DSB1 в своем геноме, и затем в сайт расщепления этого фермента DSB1 или рядом с ним может быть введена любая целевая ДНК.

[119] При использовании по тексту настоящего документа термин «молекула репаративной нуклеиновой кислоты» представляет собой одноцепочечную или двухцепочечную молекулу ДНК или молекулу РНК, которая используется в качестве матрицы для модификации геномной ДНК в предварительно выбранном сайте в непосредственной близости от сайта расщепления или на нем. При использовании по тексту настоящего документа термин «использование в качестве матрицы для модификации геномной ДНК» означает, что молекула репаративной нуклеиновой кислоты копируется или интегрируется в заранее выбранный сайт посредством гомологичной рекомбинации между фланкирующей областью и соответствующей областью гомологии в целевом геноме, фланкирующей заранее выбранный сайт, при необходимости в комбинации с нехомологичным соединением концов (NHEJ) на одном из двух концов молекулы репаративной нуклеиновой кислоты (например, в случае, если имеется только одна фланкирующая область). Интеграция путем гомологичной рекомбинации позволит точно присоединить молекулу репаративной нуклеиновой кислоты к целевому геному до уровня нуклеотидов, в то время как NHEJ может привести к небольшим вставкам/делециям на стыке между молекулой репаративной нуклеиновой кислоты и геномной ДНК.

[120] При использовании по тексту настоящего документа термин «модификация генома» означает, что геном был изменен, по меньшей мере, в одном нуклеотиде (в соответствии с одним из вариантов осуществления изобретения указанное изменение не происходит в немодифицированном растении/растении дикого типа). Модификация может происходить путем замены, по меньшей мере, одного нуклеотида и/или делеции, по меньшей мере, одного нуклеотида и/или вставки, по меньшей мере, одного нуклеотида, если это приводит к полному изменению, по меньшей мере, одного нуклеотида по сравнению с нуклеотидной последовательностью заранее выбранного целевого геномного сайта перед модификацией, что позволяет идентифицировать модификацию методами известными специалистам, например такими как секвенирование или ПЦР-анализ и т.д.

[121] В соответствии с другими вариантами осуществления изобретения описываются способы идентификации и/или отбора растения пшеницы, содержащего аллель гена, положительно или отрицательно влияющего на количество колосков на колос, которые включают, соответственно, этап идентификации наличия или отсутствия, соответственно, в геноме растения пшеницы нуклеиновой кислоты, имеющей нуклеотиды от положения 4399 до положения 4513 SEQ ID NO: 5, или нуклеиновой кислоты, имеющей нуклеотиды от положения 7816 до положения 7930 в SEQ ID NO: 19, или нуклеотидной последовательности, имеющей, по меньшей мере, 90% идентичности последовательности с указанной нуклеиновой кислотой.

[122] Растения пшеницы по настоящему изобретению можно выращивать или собирать для получения зерна, в первую очередь для использования в качестве пищевого продукта или в качестве корма для животных, или для ферментации или промышленного производства сырья, например, помимо прочего, для производства этанола. В качестве альтернативы, растения пшеницы могут использоваться непосредственно в качестве корма. Растение по настоящему изобретению предпочтительно может использоваться для производства пищевых продуктов, в частности, для промышленного производства пищевых продуктов. Такое производство пищевых продуктов может включать производство муки, теста, манной крупы или других продуктов из зерна, которые могут использоваться в качестве ингредиента при коммерческом производстве пищевых продуктов. Настоящим изобретением также предоставляется мука, крупа или другие продукты, которые производятся из зерна. Такие продукты могут не подвергаться обработке или подвергаться обработке, например, путем фракционирования или отбеливания.

[123] Настоящим изобретением также предоставляются продукты, полученные из растений или зерна/семян по настоящему изобретению, например, пищевые продукты, которые могут использоваться в качестве ингредиентов пищевых продуктов. Примеры пищевых продуктов включают муку, крахмал, квасной или пресный хлеб, макаронные изделия, лапшу, корм для животных, блюда из зернового продукта для завтрака, закусочные пищевые продукты, торты, солод, выпечку и продукты, содержащие соусы на основе муки. Указанный пищевой продукт может представлять собой бублик, несладкое печенье, хлеб, булочку, круассан, клецку, маффин, английский маффин,

питу, хлеб из теста, приготовленного ускоренным способом, продукт из охлажденного/замороженного теста, тесто, консервированную фасоль, буррито, острый соус чили, тако, тамале, тортилью, закрытый пирог, зерновой продукт, готовый к употреблению, сухой паек, начинка, блюдо, предназначенное для приготовления в СВЧ печи, пирожное, торт, чизкейк, кофейный торт, сладкое печенье, десерт, выпечку, сладкий рулет, шоколадный батончик, основу для пирога, начинка для пирога, детское питание, смесь для выпечки, жидкое тесто, панировку, смесь для подливки, добавку для увеличения сроков хранения мясных продуктов, заменитель мяса, смесь приправ, суповую смесь, подливку, заправку для салата, суп, сметану, лапшу, пасту, рамен, чоу-мейн, ло-мейн, добавку для мороженого, брикет мороженого, конусный вафельный стаканчик для мороженого, вафлю для мороженого, крекер, гренок, пончик, яичный рулет, экструдированную закуску, батончик мюсли, закусочный продукт, предназначенный для приготовления в СВЧ печи, питательный батончик, блин, частично обжаренную выпечку, крендель, пудинг, продукт на основе гранолы, закуску из чипсов, закусочный пищевой продукт, закусочную смесь, вафли, основу для пиццы, корм для животных, в том числе корм для домашних животных. Пищевой продукт может быть приготовлен путем смешивания зерна или муки, цельнозерновой муки или отрубей из указанного зерна с другим ингредиентом. Другим видом продуктов является корм для животных, например, собранное зерно, сено, солома или силос. Растения по настоящему изобретению могут использоваться непосредственно в качестве корма для животных, например, при выращивании в поле.

[124] В соответствии с одним из вариантов осуществления изобретения настоящим изобретением предоставляется способ получения пшеничной муки, цельнозерновой муки, крахмала, крахмального зерна или отрубей, включающий получение зерна растения по изобретению и переработку зерна для получения муки, цельнозерновой муки, крахмала, крахмального зерна или отрубей, а также предоставляется пшеничная мука, цельнозерновая мука, крахмал, крахмальное зерно или отруби, которые получены с применением такого способа или которые содержат молекулу нуклеиновой кислоты Apo1 по изобретению и/или полипептид APO1 по изобретению.

[125] Кроме того, настоящим изобретением предоставляется способ получения пищевого продукта, включающий смешивание зерна растений по изобретению или

пшеничной муки, цельнозерновой муки, крахмала, крахмального зерна или отрубей, по меньшей мере, с одним другим пищевым ингредиентом с получением пищевого продукта. Также настоящим изобретением предоставляется способ производства крахмала, включающий получение зерна растений по изобретению и обработку зерна для производства крахмала, а также способ производства этанола, который включает ферментацию указанного крахмала с получением этанола.

[126] Кроме того, настоящим изобретением предоставляется способ кормления животных, включающий кормление животного растениями пшеницы по изобретению, зерном пшеницы по изобретению, клетками пшеницы по изобретению или кормовым продуктом, содержащим указанную пшеничную муку, цельнозерновую муку, крахмал, крахмальное зерно или отруби.

[127] Настоящим изобретением также предоставляется пищевой продукт, содержащий растение пшеницы по изобретению или его часть, зерно пшеницы по изобретению, клетки пшеницы по изобретению, молекулу нуклеиновой кислоты по изобретению, полипептид по изобретению, ингредиент, который представляет собой указанную пшеничную муку, цельнозерновую муку, крахмал, крахмальное зерно или отруби, такой как указанный пищевой продукт, при этом указанный пищевой продукт представляет собой квасной или пресный хлеб, макаронные изделия, лапшу, блюда из зернового продукта для завтрака, закусочные пищевые продукты, торты, выпечку и соусы на основе муки.

[128] Кроме того, настоящим изобретением предоставляются семена растений по изобретению, содержащие аллель *Aro1* по изобретению, а также продукты пшеницы, полученные из таких семян, которые содержат аллель *Aro1*. Такой пшеничный продукт может представлять собой или может содержать муку, молотые семена, порошок, хлопья и т.д. В частности, такой продукт из пшеницы содержит нуклеиновую кислоту, которая продуцирует ампликон, который является диагностическим или специфическим в отношении аллеля *Aro1* по изобретению.

[129] Настоящим изобретением также предоставляется способ изменения количества колосков на колос растения пшеницы, включающий этап изменения количества белка *ARO1* по изобретению в указанном растении пшеницы, в частности, способ,

отличающийся тем, что увеличивается количество указанного белка и увеличивается количество колосков на колос по сравнению с количеством колосков на колос указанного растения пшеницы с неизменным количеством белка.

[130] Способ по п. выше, отличающийся тем, что количество указанного белка уменьшается и уменьшается количество колосков на колос по сравнению с количеством колосков на колос указанного растения пшеницы с неизменным количеством белка, например, способ, отличающаяся тем, что количество указанного белка увеличивается путем получения указанного растения пшеницы с:

- a. рекомбинантным геном по изобретению; или
- b. гетерологичным геном, кодирующим белок APO1 по изобретению, при этом степень экспрессии указанного гетерологичного гена выше, чем степень экспрессии соответствующего эндогенного гена, например, когда указанный гетерологичный ген содержит нуклеотидную последовательность SEQ ID NO: 4 или SEQ ID NO: 5 или нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с указанной последовательностью.

[131] Настоящим изобретением также предоставляется способ по двум пп. выше, отличающийся тем, что количество указанного белка уменьшается путем получения указанного растения пшеницы с:

- a. гетерологичным геном, кодирующим белок APO1 по изобретению, при этом степень экспрессии указанного гетерологичного гена ниже, чем степень экспрессии эндогенного гена; или
- b. мутантным аллелем эндогенного гена, кодирующего белок APO1 по изобретению.

[132] Способ по предыдущему п., отличающийся тем, что промотор указанного гетерологичного гена содержит нуклеотидную последовательность SEQ ID NO: 9 или нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с указанной последовательностью, и не содержит нуклеотидную последовательность от нулкотидного положения 4399 до нулкотидного положения 4513 SEQ ID NO: 5, и не содержит нуклеотидную последовательность, имеющую, по

меньшей мере, 90% идентичности последовательности с указанной последовательностью, например, где указанный мутантный аллель является нокаут-аллелем.

[133] Способ по любому из пп. выше, где этап получения растения включает получение растения путем трансформации, скрещивания, обратного скрещивания, интрогрессии, редактирования генома или мутагенеза.

[134] Трансформированные растения и клетки растений, полученные способами, описанными в настоящем документе, могут дополнительно использоваться в процедурах селекции, известных специалистам, таких как скрещивание, самоопыление и обратное скрещивание. Селекционные программы могут включать скрещивание для получения поколения F1 (первого потомка), за которым следует несколько поколений самоопыления (поколения F2, F3 и т.д.). Селекционная программа может также включать этапы обратного скрещивания (BC), в ходе которых потомство скрещивается с одной из родительских линий, которая называется рекуррентным родителем.

[135] В определенных юрисдикциях патентоспособности могут быть лишены растения по изобретению, которые были получены исключительно с помощью преимущественно биологических процессов, при этом процесс получения растений считается преимущественно биологическим, если он включает только природные процессы, такие как скрещивание или селекция. Таким образом, растения по изобретению также включают растения, полученные не исключительно преимущественно биологическими способами.

[136] Перечень последовательностей, который находится в файле с названием "BCS18-2001-WO1\_ST25.txt", размером 87 килобайт, который содержит 31 последовательность, с SEQ ID NO: 1 по SEQ ID NO: 31, подан для регистрации вместе с настоящим документом путем подачи в электронном виде, указанный перечень включен в настоящую заявку посредством ссылки.

[137] В описании изобретения и в разделе «Примеры» имеются ссылки на следующие последовательности:

SEQ ID NO: 1: нуклеотидная последовательность кодирующей ДНК Apo1-7A из Chinese Spring, Westonia или Baxter.

SEQ ID NO: 2: нуклеотидная последовательность геномной ДНК Apo1-7A из Chinese Spring, Westonia или Baxter.

SEQ ID NO: 3: аминокислотная последовательность белка APO1-7A из Chinese Spring, Westonia или Baxter.

SEQ ID NO: 4: нуклеотидная последовательность 5' апстрим последовательности Apo1-7A из Westonia.

SEQ ID NO: 5: нуклеотидная последовательность 5' апстрим последовательности Apo1-7A из Baxter.

SEQ ID NO: 6: нуклеотидная последовательность кодирующей ДНК Apo1-7A из Chara или Yitpi.

SEQ ID NO: 7: нуклеотидная последовательность геномной ДНК Apo1-7A из Chara или Yitpi.

SEQ ID NO: 8: аминокислотная последовательность белка APO1-7A из Chara или Yitpi.

SEQ ID NO: 9: нуклеотидная последовательность 5' апстрим последовательности Apo1-7A из Chara или Yitpi.

SEQ ID NO: 10: нуклеотидная последовательность молекулярного маркера wsnr\_Ku\_c19943\_29512612.

SEQ ID NO: 11: нуклеотидная последовательность молекулярного маркера Excalibur\_c95707\_285.

SEQ ID NO: 12: нуклеотидная последовательность молекулярного маркера mTRI00073530.

SEQ ID NO: 13: нуклеотидная последовательность молекулярного маркера mTRI00055675.

SEQ ID NO: 14: нуклеотидная последовательность молекулярного маркера mTRI00055678.

SEQ ID NO: 15: нуклеотидная последовательность 7B последовательности, кодирующей гомологичный APO1 ген (Chinese Spring).

SEQ ID NO: 16: нуклеотидная последовательность 7D последовательности, кодирующей гомологичный APO1 ген (Chinese Spring).

SEQ ID NO: 17: аминокислотная последовательность белка APO1-7B (Chinese Spring).

SEQ ID NO: 18: аминокислотная последовательность белка APO1-7D (Chinese Spring).

SEQ ID NO: 19: нуклеотидная последовательность 5' апстрим последовательности Apo1-7A из Chinese Spring.

SEQ ID NO: 20: 1242 нуклеотидная последовательность кодирующей ДНК Apo1-7B из Chinese Spring.

SEQ ID NO: 21: нуклеотидная последовательность геномной ДНК Apo1-7B из Chinese Spring.

SEQ ID NO: 22: нуклеотидная последовательность 5' апстрим последовательности Apo1-7B из Chinese Spring.

SEQ ID NO: 23: нуклеотидная последовательность маркера CAP7\_c2350\_105.

SEQ ID NO: 24: нуклеотидная последовательность маркера wsnr\_Ku\_rep\_c104159\_90704469.

SEQ ID NO: 25: нуклеотидная последовательность маркера BS00021657\_51.

SEQ ID NO: 26: нуклеотидная последовательность маркера BS00066288\_51.

SEQ ID NO: 27: нуклеотидная последовательность маркера BS00039502\_51.

SEQ ID NO: 28: нуклеотидная последовательность кодирующей ДНК Apo1-7A из Chinese Spring (более короткая версия).

SEQ ID NO: 29: аминокислотная последовательность белка APO1-7A из Chinese Spring (более короткая версия).

SEQ ID NO: 30: нуклеотидная последовательность кодирующей ДНК Apo1-7B из Chinese Spring (более короткая версия).

SEQ ID NO: 31: аминокислотная последовательность белка APO1-7B из Chinese Spring (более короткая версия).

## ПРИМЕРЫ

[138] Если иное не указано в разделе «Примеры», все технологии рекомбинантных ДНК осуществляются согласно стандартным протоколам в соответствии с описанием в работе Sambrook and Russell (2001) *Molecular Cloning: A Laboratory Manual*, Third Edition, Cold Spring Harbor Laboratory Press, NY, Том 1 и 2 Ausubel et al. (1994) *Current Protocols in Molecular Biology*, Current Protocols, США и Том I и II Brown (1998) *Molecular Biology LabFax*, Второе Издание, Academic Press (UK). Стандартные методы и материалы молекулярной биологии для работы с растениями описаны в *Plant Molecular Biology Labfax* (1993), R.D.D. Croy, совместно опубликованные BIOS Scientific Publications Ltd (UK) и Blackwell Scientific Publications, UK. Стандартные материалы и способы осуществления полимеразной цепной реакции описаны в публикации Dieffenbach and Dveksler (1995) *PCR Primer: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, и в McPherson et al. (2000) *PCR - Basics: From Background to Bench*, Первое Издание, Springer Verlag, Германия. Стандартные процедуры для анализа AFLP (полиморфизм длины амплифицированных фрагментов, ПДАФ) описаны в Vos et al. (1995, NAR 23:4407-4414) и опубликованы в заявке на Европейский патент EP 534858.

[139] В разделе «Примеры» представлены результаты, полученные с использованием 2 различных популяций пшеницы: одна основана на анализе группы растений яровой пшеницы (раздел А ниже), а другая – на анализе группы растений озимой пшеницы (раздел В ниже); эти результаты демонстрируют, что идентифицированный фенотип SPS (SPS- или SPS +), связанный с присутствующим аллелем типа Apo1, применим ко всем популяциям/генотипам пшеницы.

### **А. Анализ АРО1 в линиях яровой пшеницы**

#### **Пример 1: Картирование QTL на хромосоме 7А, который управляет количеством колосков на колосе**

[140] Осуществлялось фенотипирование двойной межлинейной гибридной популяции яровой пшеницы MAGIC (Huang et al. 2012 *Plant Biotechnology Journal* 10:826-839) путем подсчета количества колосков на колос на разных линиях растений.

[141] С использованием генетической карты нескольких SNP, был проведен QTL-анализ для проверки влияния вариации количества колосков на колос по всем маркерам. Значимые связи маркер-признак различаются р-значениями, преобразованными с помощью  $-\log$  выше 3. Таким образом был выделен интервал значимо связанных маркеров, включая маркеры фланкирования (SEQ ID NO: 10 и SEQ ID NO: 11). Интервал значимо связанных маркеров определяли с использованием следующих критериев: порог значимости 2,5, падение значимости 1,5 и падение значимости между пиками 2. Это ограничило интервал для 7A до 2,1 сМ по левому и правому фланкирующим маркерам.

[142] Были получены гетерогенные инбредные семейства (НIF) с контрастным наличием 7A SPS QTL (Fam1\_A\_1, Fam1\_B\_1, Fam2\_B\_1, Fam2\_C\_1, Fam2\_H\_1, Fam3\_E\_1, Fam3\_I\_1, Fam4\_A, Fam4\_G, Fam5\_C\_1 и Fam5\_F\_1), которые затем использовались для точного картирования и анализа экспрессии ниже 7A QTL.

[143] Фенотипирование НIF с контрастирующим присутствием аллелей, влияющих на высокое и низкое количество колосков для 7A SPS QTL, осуществлялось в соответствии с описанием выше. Были разработаны дополнительные анализы SNP для увеличения плотности маркеров в интервале QTL. Лocus SPS может быть далее ограничен областью около 2,1 сМ на 7A (от 58.7 до 60.8 сМ вдоль хромосомы 7A), ограниченной фланкирующими маркерами (SEQ ID NO: 12 и SEQ ID NO: 13 или SEQ ID NO: 14).

[144] Последовательность точно картированных маркеров использовали для поиска контигов и скаффолдов геномной последовательности Chinese Spring с помощью BLAST. Для позиционирования SNP в частичной геномной последовательности применялись строгие критерии BLAST и синтаксического анализа, такие как > 98% идентичности, длина выравнивания > 158 п.о., совпадение в последовательности 7A и дополнительные критерии для невыровненного выступа. Скаффолды использовались для точного картирования (и создания дополнительных генетических карт). 16 аннотированных генов в пределах интервала, определенного точным картированием, подвергали анализу экспрессии в соответствии с описанием в Примере 2.

## **Пример 2: Анализ экспрессии и идентификация Aro1**

[145] Анализ экспрессии выполняли путем полного транскриптомного секвенирования методом дробовика образцов РНК, полученных из контрастирующих семейств НІF, как описано по существу Wang et al. (2009) Nature Review Genetics 10, 57-63. Количественную оценку экспрессии осуществляли путем подсчета нормализованного количества считываемых фрагментов, которые сопоставлены с интервалом QTL, определенным в Примере 1.

[146] Осуществлялась количественная оценка уровня экспрессии 16 генов, аннотированных в интервале, определенном точным картированием, у различных родителей картируемой популяции, а также в 11 НІF. Из них существенно более высокая степень экспрессии (в среднем в 1,8 раза) наблюдалась только у одного кандидата, ортолога APO1 риса, в линиях, демонстрирующих фенотип большого количества колосков на колос (сокращенно в настоящем документе иногда обозначается как SPS+ (фенотип) по сравнению с линиями, имеющими низкое количество колосков на колосья (которые иногда сокращенно в настоящем документе обозначаются как SPS- (фенотип). Этот ген впоследствии был идентифицирован как базовый ген для QTL в отношении количества колосков на колос на хромосоме 7A.

[147] На Фиг. 1 показаны подробные результаты анализа уровня экспрессии гена Aro1 с помощью анализа транскрипции последовательности РНК в анализируемых генотипах яровой пшеницы. Разница в содержании транскриптов APO1 в контрастных линиях составляет, по меньшей мере, 1,5 - 2,75 раза. Родители сортов Chara и Yitpi имеют низкое количество колосков на колос и низкий уровень экспрессии Aro1, тогда как родители Westonia и Baxter имеют высокое количество колосков на колос и имеют более высокий уровень экспрессии Aro1 (разница в 1,6 - 2,6 раза и выше). Аналогичным образом, линии НІF с низким количеством колосков на колос имеют низкий уровень экспрессии Aro1, тогда как линии НІF с высоким количеством колосков на колос имеют более высокий уровень экспрессии Aro1.

[148] Последовательность гена APO1 была получена из эталонной линии пшеницы Chinese Spring, а также из четырех родительских сортов MAGIC. APO1 является высококонсервативным и имеет более 99% идентичности последовательности между

последовательностью аллеля от сортов с низким количеством колосков на колос и последовательностью аллеля от сортов с высоким количеством колосков на колос. В Таблице 1 показаны 3 однонуклеотидных полиморфизма, обнаруженных между проанализированными кодирующими последовательностями APO1. Соответствующие аминокислотные последовательности также имеют 99% идентичности последовательностей. SNP в положении 140 на SEQ ID NO: 2 или 7 приводит к тому, что белковая последовательность Yitpi и Chara (SEQ ID NO: 8) имеет цистеин в положении 47, в то время как белковые последовательности Baxter, Westonia и Chinese Spring (SEQ ID NO: 3) имеют фенилаланин в положении 47. SNP в положении 842 на SEQ ID NO: 2 или 7 не приводит к разнице в аминокислотных последовательностях, как в интроне. SNP в положении 1284 на SEQ ID NO: 2 или 7 приводит к тому, что белковая последовательность Yitpi и Chara (SEQ ID NO: 8) имеет аспарагин в положении 384, в то время как белковые последовательности Baxter, Westonia и Chinese Spring (SEQ ID NO: 3) имеют аспарагиновую кислоту в положении 384. Прогнозируется, что эти различия в белковых последовательностях генотипов с высоким и низким количеством колосков на колос существенно не изменят функцию белка APO1.

**Таблица 1:** Однонуклеотидные полиморфизмы (SNP), идентифицированные между генными последовательностями APO1 сортов, имеющих низкое количество колосков на колос (Yitpi и Chara), и сортов, имеющих высокое количество колосков на колос (Baxter и Westonia). \*означает SNP в последовательности интрона.

Положение SEQ ID NO: 2 или 7	Yitpi/ Chara	Baxter/ Westonia	Тип
140	G	T	SNP
842*	T	C	SNP
1284	A	G	SNP

[149] Также была получена нуклеотидная последовательность длиной приблизительно 5 т.п.о. перед геном APO1, а затем она сравнивалась с четырьмя родительскими сортами. В Таблице 2 приведен перечень однонуклеотидных

полиморфизмов и вставок/делеций, обнаруженных между последовательностями из генотипов низкого количества колосков на колос и последовательностями из генотипов высокого количества колосков на колос. Неожиданно было обнаружено, что в последовательностях генотипов сортов, имеющих низкое количество колосков на колосок, отсутствуют приблизительно 115 п.о. по сравнению с последовательностями генотипов сортов, имеющих высокое количество колосков на колос, приблизительно на 500 п.о. перед сайтом начала трансляции (что соответствует сайту начала трансляции в эталонной последовательности SEQ ID NO: 1). Предполагается, что эта делеция объясняет более низкий измеренный уровень экспрессии в этих линиях.

**Таблица 2:** Однонуклеотидные полиморфизмы (SNP) и вставки/делеции (инсерционно-делеционные полиморфизмы), идентифицированные между вышележащими последовательностями Apo1 длиной приблизительно 5 т.п.о. сортов, имеющих низкое количество колосков на колос (Yitpi и Chara), и сортов, имеющих высокое количество колосков на колос (Baxter и Westonia).

Положение SEQ ID NO: 9	Yitpi/ Chara	Baxter/ Westonia	Положение SEQ ID NO: 4	Положение SEQ ID NO: 5	Тип
32	G	A	32	32	SNP
33	G	A	33	33	SNP
520	G	T	520	520	SNP
-	-	T	551	551	вст/дел
651	A	G	652	652	SNP
1063	A	-	-	-	вст/дел
1482	T	C	1482	1482	SNP
1639	C	T	1639	1639	SNP
2093	T	-	-	-	вст/дел
2094	C	-	-	-	вст/дел
2095	T	-	-	-	вст/дел
2096	C	-	-	-	вст/дел
2097	T	-/T	2093	-	вст/дел
2098	C	-/C	2094	-	вст/дел
2660	A	G	2656	2654	SNP
2730	A	C	2726	2724	SNP
2747	A	G	2743	2741	SNP
2759	T	G	2755	2753	SNP
2785	C	T	2781	2779	SNP
2792	T	C	2788	2786	SNP
3000	T	C	2996	2994	SNP
3241	G	A	3237	3235	SNP
3456	C	T	3452	3450	SNP
3493	C	T	3489	3487	SNP
3603	G	A	3599	3597	SNP
-	-	G	4108	4106	вст/дел
-	-	C	4109	4107	вст/дел

-	-	CAATTTACTCTAG TTGCATCCCAACA TCGTGCCCTACC TCGCCTCCGGCTA GGTCATTCCAAGC CCTAGTCGCCGAC GTCGCAACCCTGT CTCATGCTCGGCG GCTATCTAATT	4401-4515	4399-4513	вст/дел
4403	A	C	4516	4514	SNP
4427	G	A	4540	4538	SNP
4643	A	G	4756	4754	SNP
4753	G	T	4866	4864	SNP

[150] Идентифицированные SNP и инсерционно-делеционные полиморфизмы между генотипами высокого и низкого количества колосков на колос, также можно использовать в качестве маркеров для определения того, какой аллель гена APO1 входит в какой-либо определенный генотип пшеницы.

[151] Выращивание двух линий NIL яровой пшеницы (NIL), контрастирующих в локусе APO1-7A в различных средах, показало, что аллель APO1-7A, обеспечивающий снижение количества колосков на колос (SPS-), был связан со значительным увеличением урожайности в полевых испытаниях (3 - 6 повторов для каждой тестируемой линии) при выращивании в Австралии, по сравнению с контрастирующими NIL, несущими аллель APO1-7A, обеспечивающий увеличение количества колосков на колос (SPS+), на том же генетическом фоне (выращенными в тех же испытаниях). Эта связь была обратной, когда те же самые NIL выращивали в полевых испытаниях во Франции (3 - 6 повторов для каждой тестируемой линии), где линии с аллелем APO1, обеспечивающим увеличение количества колосков на колос (SPS+), демонстрировали значительное увеличение урожайности в сравнении с родственными линиями с аллелем APO1 SPS- в том же генетическом фоне (выращенными в тех же испытаниях). При том, что эффекты на урожайность были

противоположными, эффекты каждого из 2 аллелей APO1-7A для фенотипа SPS были одинаковыми во всех средах.

**Пример 3: Проверка Apo1 в качестве гена, определяющего количество колосков на колос у растений пшеницы с исходным низким количеством колосков на колос (ГМ-подход)**

[152] С использованием стандартных технологий рекомбинантных ДНК были функционально связаны следующие области ДНК:

- a. промоторная область CaMV35S (P35S)
- b. ДНК-область, кодирующая TaAPO1
- c. ДНК-область, представляющая терминатор октопин синтазы нетранслируемой 3'-последовательности

Рекомбинантный ген вводили в вектор Т-ДНК, который содержит каскету селективируемых маркеров, с получением вектора Т-ДНК P35S::APO1.

[153] С использованием стандартных технологий рекомбинантных ДНК были функционально связаны следующие области ДНК:

- a. область убиквитинового промотора (PUbi)
- b. ДНК-область, кодирующая TaAPO1
- c. ДНК-область, представляющая терминатор октопин синтазы нетранслируемой 3'-последовательности

Рекомбинантный ген вводили в вектор Т-ДНК, который содержит каскету селективируемых маркеров, с получением Т-ДНК PUbi::APO1.

[154] С использованием стандартных технологий рекомбинантных ДНК были функционально связаны следующие области ДНК:

- a. промоторная область Apo1 размером около 5 т.п.о. из сорта пшеницы Westonia (SEQ ID NO: 4)
- b. ДНК-область, кодирующая TaAPO1
- c. ДНК-область, представляющая терминатор октопин синтазы нетранслируемой 3'-последовательности

Рекомбинантный ген вводили в вектор Т-ДНК, который содержит каскету селективируемых маркеров, с получением Т-ДНК Apo1::APO1.

[155] В агробактерии с использованием стандартных методик вводили три вектора T-ДНК, содержащие вспомогательные Ti-плазмиды, и применяли их при трансформации пшеницы в соответствии с описанием в работе с использованием стандартных методик Том 1, *Methods in Molecular Biology*, том 1223 : 189-198. Трансформацию осуществляют либо непосредственно Chara или Yitpi, либо осуществляют трансформацию любого другого сорта, и затем такой сорт используют в качестве донора для введения рекомбинантного гена в сорт Chara или Yitpi путем скрещивания и отбора. Сорт пшеницы Fielder используют в качестве контроля эффективности трансформации. Также осуществляли фенотипирование трансформантов сорта Fielder для оценки влияния сверхэкспрессии гена APO1 на количество колосков на колос. Трансформанты Fielder могут использоваться для интрогрессии рекомбинантного гена в Chara или Yitpi.

[156] После каждой трансформации получают независимые события и фенотипируют их в соответствии с методом, описанным в Примере 1.

#### **Пример 4: Идентификация гомеологов Apo1 в пшенице**

[157] С использованием нуклеотидной последовательности гена, кодирующего APO1, расположенного на хромосоме 7A, можно было обнаружить гомеологичные нуклеотидные последовательности, которые расположены на хромосомах 7B и 7D соответственно в эталонных геномах пшеницы сорта Chinese Spring. Нуклеотидные последовательности для кодирующих областей этих генов включены в перечень последовательностей: SEQ ID NO: 15 (7B Apo1) и 16 (7D Apo1), соответственно. Аминокислотные последовательности включены в перечень последовательностей: SEQ ID NO: 17 (7B Apo1) и SEQ ID NO: 18 (7D Apo1). Согласно более короткой модели гена для 7B Apo1, нуклеотидная последовательность соответствует SEQ ID NO: 15 от нуклеотида 130 до нуклеотида 1452, а аминокислотная последовательность соответствует SEQ ID NO: 17 от аминокислоты 45 до аминокислоты 483.

[158] Соответствующие значения идентичности кодирующих нуклеотидных последовательностей представлены в Таблице 3, а значения идентичности аминокислотных последовательностей кодируемых белков представлены в Таблице 4.

**Таблица 3:** % идентичности последовательности между гомологичными генами Apo1.

	<b>Apo1 7A</b> (SEQ ID NO: 1)	<b>Apo1 7B</b> (SEQ ID NO: 15)	<b>Apo1 7B укороч.</b> (SEQ ID NO: 15 от nt 130)	<b>Apo1 7D</b> (SEQ ID NO: 16)
<b>Apo1 7A</b> (SEQ ID NO: 1)	100			
<b>Apo1 7B</b> (SEQ ID NO: 15)	88	100		
<b>Apo1 7B укороч.</b> (SEQ ID NO: 15 от nt 130)	97	91	100	
<b>Apo1 7D</b> (SEQ ID NO: 16)	96	87	96	100

**Таблица 4:** % идентичности последовательности между белками Apo1, кодированными гомологичными генами.

	<b>Apo1 7A</b> (SEQ ID NO: 3)	<b>Apo1 7B</b> (SEQ ID NO: 17)	<b>Apo1 7B укороч.</b> (SEQ ID NO: 17 от aa 45)	<b>Apo1 7D</b> (SEQ ID NO: 18)
<b>Apo1 7A</b> (SEQ ID NO: 3)	100			
<b>Apo1 7B</b> (SEQ ID NO: 17)	89	100		
<b>Apo1 7B укороч.</b> (SEQ ID NO: 17 от aa 45)	97	90	100	
<b>Apo1 7D</b> (SEQ ID NO: 18)	97	88	97	100

## **В. Анализ APO1 в линиях озимой пшеницы**

### **Пример 1: приблизительное картирование QTL на хромосоме 7A, который управляет количеством колосков на колосе**

#### Фенотипирование

[159] В течение полевого сезона 2013/2014 осуществляли полностью повторное испытание 784 линий F<sub>7</sub> MAGIC из популяции озимой пшеницы MAGIC согласно Mackay et al. (2014, G3-Genes Genomes Genetics, 4(9): 1603-1610) и осуществляли выращивание восьми основателей этих линий. С 1000 из 1600 участков поля были собраны десять репрезентативных колосьев пшеницы, которые затем высушивали при комнатной температуре. Сбор осуществляли в соответствии с дизайном исследования с частичным воспроизведением, при этом сбор 200 RIL и родителей MAGIC осуществляли в двух экземплярах. Осуществляли скриннинг колосьев пшеницы по морфологическому признаку – общему количеству колосков на колос (сокращенно «SPS»).

[160] В 2014/2015 в питомнике осуществлялся скриннинг 1091 линии F<sub>8</sub> MAGIC и основателей на те же характеристики колоса с использованием выборки из шести репрезентативных колосьев пшеницы на участок.

[161] Для минимизации или удаления пространственных эффектов в данных фенотипа из-за вариаций поля использовали Asreml-R 3.0 (Gilmour et al. 1997, Journal of Agricultural Biological and Environmental Statistics, том 2(3), 269-293). В то время как пакет `mrwgaim` для анализа QTL обеспечивает возможность одноэтапного сопоставления QTL, другие пакеты для анализа QTL, используемые в этом исследовании, требовали предварительного расчета характеристики BLUPS (наилучший линейный объективный прогноз).

[162] Общее количество колосков варьировало в диапазоне 18 - 30 колосков на колос в RIL. Родителей MAGIC можно в общих чертах разделить на группы с фенотипом высокого и низкого количества колосков, при этом сорта Soissons, Robigus и Brompton имеют меньшее количество колосков по сравнению с другими пятью родителями MAGIC (Фиг. 2). Средний фенотип сорта Soissons даже ниже, чем у сорта Robigus и Brompton, и лишь на 2,6 колоска больше, чем зарегистрированный минимальный

фенотип в RIL (рекомбинантные инбредные линии). Уменьшение общего количества колосков у сорта Soissons связано с тем, что в отличие от других сортов он обладает нечувствительным к фотопериоду аллелем Ppd-D1, который обеспечивает как более раннее цветение, так и уменьшенное количество колосков (González et al, 2005, *Euphytica* 146(3):253-269). Остальные 7 родителей MAGIC не несут этот аллель, и, следовательно, причина уменьшения количества колосков у Robigus и Brompton не связана с этим аллелем Ppd-D1.

Генетическое картирование:

[163] Анализ QTL проводили с использованием трех различных методологий: (i) с использованием простой регрессии средних линий с маркерными баллами при учете пересекающейся воронкообразной структуры MAGIC (Mackay et al. 2014) с использованием пакета R Asreml-R (Gilmour, 1997), (ii) анализа Байесовской сети с использованием пакета R bnlearn (Scutari et al., 2014, *Genetics*, 198(1):129-137)) и (iii) картирования среднего интервала всего генома с использованием пакета R mrwgaim (Verbyla et al., 2014, *G3*, 4(9):1569-1584).

[164] Использовали генотипы маркеров и их соответствующие хромосомные группы согласно Gardner et al., 2016 (2016, *Plant Biotechnol J*, 14(6):1406-1417).

[165] Все три метода идентифицировали основной QTL на хромосоме 7A между 257.05cM и 257.21cM на MAGICmapv14.4, далее – *Q<sub>Tsn,jbl-7A</sub>* (Table 5).

**Таблица 5:** Сводные данные по значимым QTL, идентифицированным для характеристики общего числа колосков (SPS) с помощью регрессии [17], анализ с использованием байесовской сети [23] или полногеномного картирования интервалов [22]. Пиковый маркер в регрессионном анализе – это маркер с самым низким или с совместным самым низким значением *p*. Значимые маркеры могут отходить дальше от показанного пикового маркера. Mrwgaim дает *p*-значения <0,0005 до 0. *q*-значения 0 регрессии <2,2E-16. Сокращения: хромосома (chr) и сантиморган (cM).

Метод	Пиковый маркер/левый маркер	Хромосома	cM	q- значение, регрессия	p- значение mpwgaim	Правый маркер	cM	%вар.
Регрессия	wsnp_Ku_rep_c104159_90704469	7A	257.21	0				
mpwgaim	CAP7_c2350_105	7A	257.05		0	BS00021657_51	257.21	35
Bnlearn	wsnp_Ku_rep_c104159_90704469	7A	257.21					
mpwgaim	BS00066288_51	7B	144.34		1.00E-03	BS00039502_51	144.50	1.9

[166] Информация о маркерах

CAP7\_c2350\_105

([https://triticeaetoolbox.org/wheat/view.php?table=markers&name=CAP7\\_c2350\\_105](https://triticeaetoolbox.org/wheat/view.php?table=markers&name=CAP7_c2350_105))

TAGTAAGCTCTTCAACGAGGATGGATGTTGTGTAATTTGGACAAGTGCGA[C/T]  
GTATGTCACATCTTTTTTTAATGATCCTAATCTATGATCGAAGTTCGTT (SEQ ID  
NO: 23).

wsnp\_Ku\_rep\_c104159\_90704469

<https://triticeaetoolbox.org/wheat//view.php?table=markers&name=IWA7409>

TGCCGGCCTGCAAGCCGATCCTTACTCCAAARTGGGTTGTCTCGGTGTTTTTCCT  
TGTCGGCGTCGTCTTTGTCCCAGTTGGTGTTCGTTTCGCTACTAGC[C/T]gcacaagatg  
ttgtgagatcattgatcggtatgatcatgcatgtgtccacctaacatgactgataacaagcttgcgtacatccagaatgagactatac  
(SEQ ID NO: 24).

Маркер BS00021657\_51

[https://triticeaetoolbox.org/wheat//view.php?table=markers&name=BS00021657\\_51](https://triticeaetoolbox.org/wheat//view.php?table=markers&name=BS00021657_51)

TCCACAAGAAAAGAGCAAGACACTCCGGCCGTTGTAGAGCTGATGGTGCG[C/T]  
GGTGATTTACCATAGACATGGTAGACGGCGCCCGTCCTCGTGGCATCAT (SEQ  
ID NO: 25).

Маркер BS00066288\_51

[https://triticeaetoolbox.org/wheat///view.php?table=markers&name=BS00066288\\_51](https://triticeaetoolbox.org/wheat///view.php?table=markers&name=BS00066288_51)

GGCACGTA CTCCCTTTCAGGACCCGACGAACAACGGCAATTCAGGTAAT[A/G]  
CATA CATCACGTA CTCTTACATACTTCAATCTTGTAATCCATAATATAT (SEQ  
ID NO: 26).

Маркер BS00039502\_51

[https://triticeaetoolbox.org/wheat//view.php?table=markers&name=BS00039502\\_51](https://triticeaetoolbox.org/wheat//view.php?table=markers&name=BS00039502_51)

ATCCCAGGGGGCGAGATTCAGAGCTTCTCGGCCATCCTGCGCAGCAGCGC[A/G]  
 GCCCCTAGTGGCTCCTCGGTCGGGTTCTTGGTGAGCCATGCCTGCGCGGC (SEQ  
 ID NO: 27).

[167] QTsn.jbl-7A объясняет значительную генетическую вариативность SPS в 35% в популяции MAGIC при  $-\log_{10}(p)$  62,64 с использованием mrwgaim. Фенотипы колосков, собранные в питомнике MAGIC в 2015 г., подтвердили наличие QTL с  $-\log_{10}(p)$  37,82 для SPS с использованием mrwgaim (Таблица 6).

**Таблица 6:** Результаты Mrwgaim по QTL в отношении общего количества колосков в питомнике MAGIC в 2015 г. Сокращения: LOGP:  $-\log_{10}(p)$ , % вар. – объясненный процент генетической вариативности.

ХРОМОСОМА	ЛЕВЫЙ МАРКЕР	РАССТ. (СМ)	ПРАВЫЙ МАРКЕР	РАССТ. (СМ)	ЗОНД	% ВАР.	LOGP
7A	CAP7_c2350_105	257.05	BS00021657_51	257.21	0	19.2	37.82

[168] Гаплотипы Brompton и Robigus вызывают относительное снижение SPS потомства более чем на 1,5 колоска как в 2014, так и в 2015 году (Таблица 7).

**Таблица 7: Сводные данные по SPS QTL для *QTsn.jbl-7A*.** Использовали данные исследования фенотипа урожайности MAGIC за 2014 г., проведенного Национальным институтом сельскохозяйственной ботаники. Прогнозные эффекты родительских гаплотипов на RIL BLUP по результатам анализа trwgaim. Сокращения: LOGP: –  $\log_{10}(p)$ . 2 и 0 – коды аллелей для соответствующих показанных маркеров.

Основатель	Эффекты основателя	Вероятность основателя	LOGP основателя	объяснение % вар.	LOGP	CAP7_c2350_105	wsnp_Ku_rep_c10 4159_90704469
<i>Alchemy</i>	0.634	0.083	1.08	35	62.64	2 (GG)	0 (AA)
<i>Brompton</i>	-1.607	0	5.31			0 (AA)	2 (GG)
<i>Claire</i>	0.566	0.102	0.99			2 (GG)	0 (AA)
<i>Hereward</i>	0.559	0.065	1.19			2 (GG)	0 (AA)
<i>Rialto</i>	0.242	0.273	0.56			2 (GG)	0 (AA)
<i>Robigus</i>	-1.766	0	6.14			0 (AA)	2 (GG)
<i>Soissons</i>	0.01	0.489	0.31			2 (GG)	0 (AA)
<i>Xi-19</i>	1.007	0.005	2.31			2 (GG)	0 (AA)

[169] Интервал генетического картирования 0,16 сМ соответствует прогнозной физической длине приблизительно 2,3 Мб и фланкирующим маркерам CAP7\_c2350\_105 (SEQ ID NO: 23) и wsnp\_Ku\_rep\_c104159\_90704469 (SEQ ID NO: 24). Увеличенное общее количество колосков наиболее близко совпадает с wsnp\_Ku\_rep\_c104159\_90704469 маркером.

[170] В дополнение к *QTsn.jbl-7A*, анализ QTL с помощью trwgaim подтвердил еще один QTL на 7B (*QTsn.jbl-7B*) для общего количества колосков в 2014 (LOGP 3.07) между фланкирующими маркерами BS00066288\_51 (144.34сМ; SEQ ID NO: 26) и BS00039502\_51 (144.50сМ; SEQ ID NO: 27), которые определяют интервал 5 Мб, прямо гомеологичный 7A QTL (смотрите Таблицу 8).

**Таблица 8:** Сводные данные по SPS QTL для QTsn.jbl-7B. Использовали данные исследования фенотипа урожайности MAGIC за 2014 г. Прогнозные эффекты родительских гаплотипов на RIL BLUP по результатам анализа mrwgaim. Сокращения: LOGP:  $-\log_{10}(p)$ . 2 и 0 – коды аллелей для соответствующих показанных маркеров.

Основатель	Эффекты основателя	Вероятность основателя	LOGP основателя	объяснение % вар.	LOGP	BS00039502_51	BS00066288_51
<i>Alchemy</i>	-0.37	0.012	1.92	1.9	3.07	0 (ТТ)	0 (ТТ)
<i>Brompton</i>	-0.198	0.12	0.92			0 (ТТ)	0 (ТТ)
<i>Claire</i>	0.283	0.048	1.32			2 (СС)	0 (ТТ)
<i>Hereward</i>	-0.043	0.395	0.4			0 (ТТ)	0 (ТТ)
<i>Rialto</i>	-0.052	0.365	0.44			0 (ТТ)	2 (СС)
<i>Robigus</i>	-0.12	0.214	0.67			0 (ТТ)	2 (СС)
<i>Soissons</i>	0.187	0.11	0.96			0 (ТТ)	0 (ТТ)
<i>Xi19</i>	0.289	0.044	1.36			2 (СС)	0 (ТТ)

## Пример 2: Идентификация гена-кандидата APO1

### 25 кандидатных генов в QTsn.jbl-7A

[171] В этом интервале размером 2,3 Мб было аннотировано 25 генов. Идентификация ортолога выявила семь генов с хорошо аннотированными ортологами и функциями: g109255 (*AtFTT/AtDTX35*), g109235 (*AtRAN1*), g109240 (*AtCHLI*), g109250 (*AtAAN*), g109253 (*AtSYP132*), g109256 (*AtALIS4*) и g109251 (*AtUFO*). *AtUFO* является ортологом *APO1* риса (*ABERRANT PANICLE ORGANIZATION 1*). Еще десять генов имели повторяющиеся аннотации как белки At5g07610, связанные с F-box. Каждый ген содержал домен F-box и демонстрировал значительную консервацию последовательностей ДНК между собой до 72,5%.

### Анализ синтении QTsn.jbl-7A демонстрирует, что APO1 является геном-кандидатом

[172] В дополнение к *QTsn.jbl-7A* (Table 8), QTL, анализ QTL с помощью mrwgaim подтвердил еще один QTL на 7B (*QTsn.jbl-7B*) относящийся к общему количеству колосков в 2014 (LOGP 3.07) между фланкирующими маркерами BS00066288\_51 (144.34cM) и BS00039502\_51 (144.50cM), которые определяют интервал 5 Мб, непосредственно гомеологичный 7A QTL.

[173] В этом интервале размером 5 Мб QTL на 7В (*QTsn.jbl-7B*) было идентифицировано 39 генов, из которых 15 были гомеологичными для 7А (Фиг. 4). Из 15 гомеологических генов ни один не имеет идентифицируемой делеционной мутации кодирующей последовательности, предсказанной с помощью PROVEAN, которая могла бы объяснить оба QTL.

[174] *QTsn.jbl-7A* и *QTsn.jbl-7B* являются синтеничными по отношению к хромосоме 6 риса, которая содержит четыре позиционно консервативных ортолога chr7A.g109235 (*AtRANI*), g109250 (*AtAAH*), g109251 (*AP01 / AtUFO*) и g109256 (*AtALIS4*).

Полиморфизмы последовательностей of *TaAP01-7A* выделяются вместе с *QTsn.jbl-7A*

[175] *TaAP01-7A* имеет два больших инсерционно-делеционного полиморфизма перед предсказанным сайтом начала транскрипции в Robigus по сравнению с Claire и Chinese Spring: делеция 115 п.о., вверх на 565 п.о. и вставка приблизительно 5-7,5 т.п.о. (7343 п.о., но размер варьируется в зависимости от качества используемой эталонной последовательности 4970 п.о. с учетом исключения N/X-прогонов) приблизительно 7 т.п.о. (7565 п.о. перед сайтом начала транскрипции, 7513 п.о. перед стартовым кодоном относительно последовательности SEQ ID NO: 1 (эталонная последовательность CS) перед сайтом начала транскрипции (TSS). Делеция 115 п.о. также присутствует в сортах пшеницы Cadenza и Paragon, выделяясь вместе с ВА00589872 в массиве 35k Breeders. Длинную вставку в Robigus, Cadenza и Paragon примерно на 7 т.п.о. выше TSS труднее охарактеризовать из-за отсутствия некоторых распознанных нуклеотидных оснований в сборках Robigus, Cadenza и Paragon TGAC, но аналогичная большая (> 5 т.п.о.) вставка также присутствует в сортах Yipti и Chara. Промотор Claire несет один блок (CC(A/T)<sub>6</sub>GG) на 2346 п.о. выше, который отсутствует в Robigus. Вставка размером приблизительно 5 - 7,5 т.п.о. также несет бокс CArG (Фиг. 3). Кроме того, сорт Robigus имеет те же SNP и инсерционно-делеционные полиморфизмы, что и сортов Yipti/Chara в Таблице 2, а сорт Claire имеет те же SNP и инсерционно-делеционные полиморфизмы, что и Westonia в Таблице 2.

[176] При сравнении нуклеотидной последовательности размером приблизительно 5 т.п.о. перед геном AP01 в других линиях пшеницы, у Robigus, Claire, Cadenza, Paragon и Fielder также были обнаружены вариации, показанные в Таблице 2. Сорт Claire имел

такие же SNP и аллели с инсерционно-делеционными полиморфизмами, что и Westonia, Fielder имел такие же аллели SNP и инсерционно-делеционного полиморфизма, что и Вахтер. Robigus, Cadenza и Paragon имели такие же аллели SNP и инсерционно-делеционного полиморфизма, что и Yitpi/Chara. Это подтверждает, что в последовательностях генотипов сортов озимой пшеницы, имеющих низкое количество колосков на колосок, отсутствуют приблизительно 115 п.о. по сравнению с последовательностями генотипов сортов озимой пшеницы, имеющих высокое количество колосков на колос, приблизительно на 500 п.о. перед сайтом начала трансляции (по отношению к сайту начала трансляции в SEQ ID NO: 1). Предполагается, что эта промоторная делеция объясняет более низкий измеренный уровень экспрессии в этих линиях. Вставка на 5-7,5 п.о. выше, идентифицированная у Robigus, Cadenza и Paragon (7565 п.о. выше TSS), также является общей для Yitpi и Chara.

[177] Согласно прогнозу PROVEAN, аминокислотные изменения (F20C, D357N), связанные с SNPs в TaAPO1-7A в Robigus, не являются делеционными.

### **Пример 3: Экспрессия *TaAPO1-7A*, коррелирующая с общим числом колосков**

[178] Три реплики целых образцов колосьев были собраны из рассеечения побегов питомника NIAB MAGIC 2017 на стадии роста gS32 (Zadoks et al., 1974, Weed Research, 14(6): 415-421) для родителей MAGIC Alchemy, Brompton, Claire, Hereward, Rialto Robigus и Xi-19. На дату сбора развитие растений сорта Soissons достигли стадии gS34. После вскрытия колосья немедленно замораживали в жидком азоте. Праймеры были разработаны с использованием плагина Primer3 (Koressaar et al., 2007, Bioinformatics, 23(10): 1289-1291) на платформе Geneious. Образцы дважды гомогенизировали при замораживании с использованием шариков из нержавеющей стали диаметром 5 мм на TissueLyser II (QIAGEN, UK) при 20Hz в течение двух минут. РНК экстрагировали с помощью набора для экстракции RNeasy Micro (QIAGEN, UK) а расщепление ДНК проводили на колонке с использованием набора RNase-free DNase (QIAGEN, UK). РНК элюировали водой без РНКазы/ДНКазы, концентрацию определяли с помощью спектрофотометра NanoDrop 1000 (Thermo Scientific, UK). Второе расщепление ДНК проводили с использованием ezDNase (Invitrogen, UK) с последующим синтезом кДНК

из 500 нг РНК с использованием набора для синтеза кДНК SuperScript IV Vilo Master Mix (Invitrogen, UK). ОТ-кПЦР осуществляли с использованием набора для ПЦР Rotor-Gene SYBR Green на аппарате для ПЦР в реальном времени Rotor-Gene Q, оснащённом Rotor-Disc 100 (QIAGEN, UK). Все реакции проводились в виде технических дубликатов при конечном реакционном объеме 10 мкл, добавляли раствор бетаина APO1 (Sigma-Aldrich) в конечной концентрации 1 М, чтобы преодолеть высокое содержание ампликонов при ГХ. Эффективность амплификации пар праймеров определяли с помощью нескольких серийных двукратных разведений образцов кДНК в восьми точках. Для подтверждения специфичности реакций ОТ-кПЦР кривые плавления для каждой реакции проверяли на наличие только одного пика. Специфичность анализов подтверждалась относительно геномной нуль-тетрасомной ДНК, полученной от Seedstor.ac.uk (WPGS1289-PG-1, WPGS1296-PG-1, WPGS1301-PG-1). Уровни экспрессии Apo1 рассчитывались относительно экспрессии генов «домашнего хозяйства» *TaRP15* (Shaw et al., 2012, Plant J, 71(1): 71-84) и *Ta2291* (Paolacci et al., 2009, BMC Molecular Biology, 10(1): 11) с использованием эффективности амплификации, рассчитанной для каждого анализа.

[179] Было обнаружено, что экспрессия TaAPO1 в Xi-19 была самой высокой из всех основателей MAGIC. Гаплотип Xi-19 также статистически значимо связан с положительным эффектом основателя на *Qtsn.jbl-7A* в 2014 (Table 8).

[180] На Фиг. 5 показаны результаты анализа уровня экспрессии гена Apo1 в исследуемых генотипах. Разница в содержании транскриптов APO1 в контрастных линиях Brompton и Xi-19 составляет до 3,8 раза.

## Формула изобретения

1. Белок, участвующий в определении количества колосков на колос пшеницы, который является ортологом белка Aberrant panicle organization 1 (APO1) из риса.

2. Белок по п. 1, содержащий аминокислотную последовательность, выбранную из:

- a. аминокислотной последовательности SEQ ID NO: 3, 8 или 29 или ее функционального варианта, или
- b. аминокислотной последовательности, имеющей, по меньшей мере, 85% идентичности последовательности с аминокислотной последовательностью SEQ ID NO: 3, 8 или 29, или ее функционального варианта.

3. Изолированная нуклеиновая кислота, кодирующая белок по п. 1 или 2.

4. Нуклеиновая кислота по п. 3, содержащая нуклеотидную последовательность, выбранную из:

- a. нуклеотидной последовательности любой из SEQ ID NO: 1 или SEQ ID NO: 2,
- b. нуклеотидной последовательности, имеющей, по меньшей мере, 80% идентичности с нуклеотидной последовательностью любой из SEQ ID NO: 1 или SEQ ID NO: 2;
- c. нуклеотидной последовательности SEQ ID NO: 6, 7 или 28;
- d. нуклеиновой кислоты, имеющей комплементарную последовательность с любой из нуклеиновых кислот по пункту а) или b).

5. Нуклеиновая кислота по п. 3 или 4, которая расположена на хромосоме 7A пшеницы в интервале, включающем нуклеотидную последовательность между нуклеотидом в положении 674,081,462 и нуклеотидом в положении 674,082,918 в эталонной геномной последовательности Chinese Spring.

6. Рекомбинантный ген, содержащий экспрессируемый в растении промотор, такой как гетерологичный экспрессируемый в растении промотор, функционально связанный с нуклеотидной последовательностью, кодирующей белок по п. 1 или 2, и, при необходимости, последовательность терминации транскрипции и полиаденилирования, предпочтительно участок терминации транскрипции и полиаденилирования, которые являются функциональными в растениях.

7. Рекомбинантный ген по п. 6, **отличающийся тем**, что нуклеиновая кислота выбрана из следующих последовательностей:

- a. последовательности нуклеиновой кислоты, имеющей нуклеотидную последовательность любой из SEQ ID NO: 1, 2 или SEQ ID NO: 28,
- b. последовательности нуклеиновой кислоты, имеющей, по меньшей мере, 80% идентичности с нуклеотидной последовательностью любой из SEQ ID NO: 1, 2 или SEQ ID NO: 28; или
- c. нуклеиновой кислоты, имеющей комплементарную последовательность с любой из нуклеиновых кислот по пункту а) или b).

8. Рекомбинантный ген по п. 6 или 7, **отличающийся тем**, что указанный экспрессируемый в растении промотор выбран из группы, состоящей из конститутивного промотора, индуцируемого промотора или тканеспецифического промотора.

9. Рекомбинантный ген по любому из пп. 6 - 8, **отличающийся тем**, что указанный экспрессируемый в растении промотор представляет собой промотор CaMV35S или убиквитиновый промотор.

10. Вектор, содержащий рекомбинантный ген по любому из пп. 6 - 9.

11. Клетка-хозяин, содержащая рекомбинантный ген по любому из пп. 6 - 9 или вектор по п. 10.

12. Клетка-хозяин по п. 11, которая является бактериальной клеткой или клеткой растения пшеницы.

13. Растение пшеницы, часть растения или семя, состоящее из клеток растения по п. 12.

14. Способ получения растения пшеницы с измененным количеством колосков на колос, включающий этап изменения содержания белка по п. 1 или 2 в указанном растении.

15. Способ по п. 14, **отличающийся тем**, что увеличивается количество указанного белка, и увеличивается количество колосков на колос по сравнению с количеством колосков на колос растения пшеницы с неизменным количеством белка.

16. Способ по п. 14, **отличающийся тем**, что количество указанного белка уменьшается, и уменьшается количество колосков на колос по сравнению с количеством колосков на колос растения пшеницы с неизменным количеством белка.

17. Способ по п. 14 или 15, **отличающийся тем**, что количество указанного белка увеличивается путем получения указанного растения пшеницы с:

- a. рекомбинантным геном по любому из пп. 6 - 9, или
- b. гетерологичным геном, кодирующим белок по п. 1 или 2, при этом степень экспрессии указанного гетерологичного гена выше, чем степень экспрессии соответствующего эндогенного гена.

18. Способ по п. 17, **отличающийся тем**, что указанный гетерологичный ген примерно на 500 п.о. выше точки начала трансляции содержит нуклеотидную последовательность, имеющую нуклеотиды от положения 4399 до положения 4513 SEQ ID NO: 5, или нуклеотидную последовательность, имеющую, по меньшей мере, 90% идентичности последовательности с указанной последовательностью.

19. Способ по п. 14 или 16, **отличающийся тем**, что количество указанного белка уменьшается путем получения указанного растения пшеницы с:

- a. гетерологичным геном, кодирующим белок по п. 1 или 2, при этом степень экспрессии указанного гетерологичного гена ниже, чем степень экспрессии эндогенного гена, или
- b. мутантным аллелем эндогенного гена, кодирующего белок по п. 1 или 2.

20. Способ по п. 19, **отличающийся тем**, что степень экспрессии указанного гетерологичного гена ниже из-за отсутствия нуклеотидной последовательности от нуклеотидного положения 4399 до нуклеотидного положения 4513 SEQ ID NO: 5.

21. Способ по п. 19, **отличающийся тем**, что указанный мутантный аллель представляет собой нокаут-аллель.

22. Способ по пп. 17 - 21, **отличающийся тем**, что этап получения растения включает получение растения путем трансформации, скрещивания, обратного скрещивания, интрогрессии, нацеленного редактирования генома или мутагенеза.

23. Продукт из пшеницы, полученный из семян по п. 13, **отличающийся тем**, что указанный продукт из пшеницы содержит или представляет собой муку, молотые семена, порошок или хлопья.

24. Продукт из пшеницы по п. 23, **отличающийся тем**, что указанный продукт из пшеницы содержит искусственную нуклеиновую кислоту, которая продуцирует ампликон, который является диагностическим или специфическим в отношении любой нуклеотидной последовательности из SEQ ID NO: 1, 2, 6, 7 или 28, или последовательности, имеющей, по меньшей мере, 80% идентичности последовательности с любой из указанных последовательностей.

25. Способ получения продукта из пшеницы по п. 23, включающий получение семян, содержащих искусственную нуклеиновую кислоту, полученную из любой нуклеотидной последовательности из SEQ ID NO: 1, 2, 6, 7 или 28, или последовательности, имеющей, по меньшей мере, 80% идентичности последовательности с любой из указанных последовательностей, и получение из семян указанного продукта из пшеницы.

26. Способ получения пшеничной муки, цельнозерновой муки, крахмала, крахмального зерна или отрубей, включающий получение семян по п. 13, содержащих искусственную нуклеиновую кислоту Apo1, и переработку семян для получения муки, цельнозерновой муки, крахмала, крахмального зерна или отрубей.

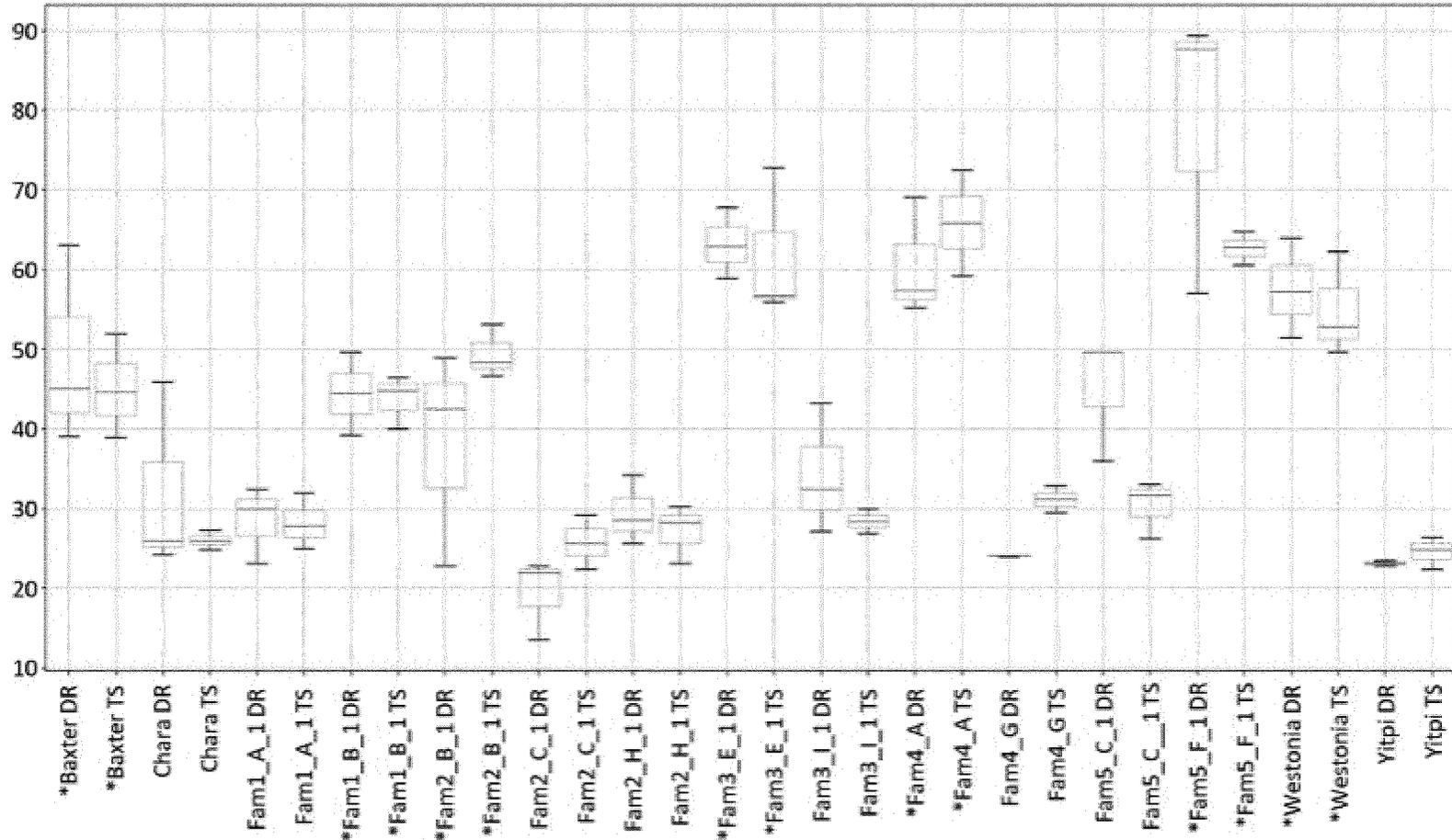
27. Пшеничная мука, цельнозерновая мука, крахмал, крахмальное зерно или отруби, полученные способом по п. 26 или содержащие искусственную нуклеиновую

кислоту, полученную из любой нуклеотидной последовательности из SEQ ID NO: 1, 2, 6, 7 или 28, или последовательности, имеющей, по меньшей мере, 80% идентичности последовательности с любой из указанных последовательностей.

28. Способ получения пищевого продукта, включающий смешивание семян по п. 13 или пшеничной муки, цельнозерновой муки, крахмала, крахмального зерна или отрубей по п. 27, по меньшей мере, с одним другим пищевым ингредиентом с получением пищевого продукта.

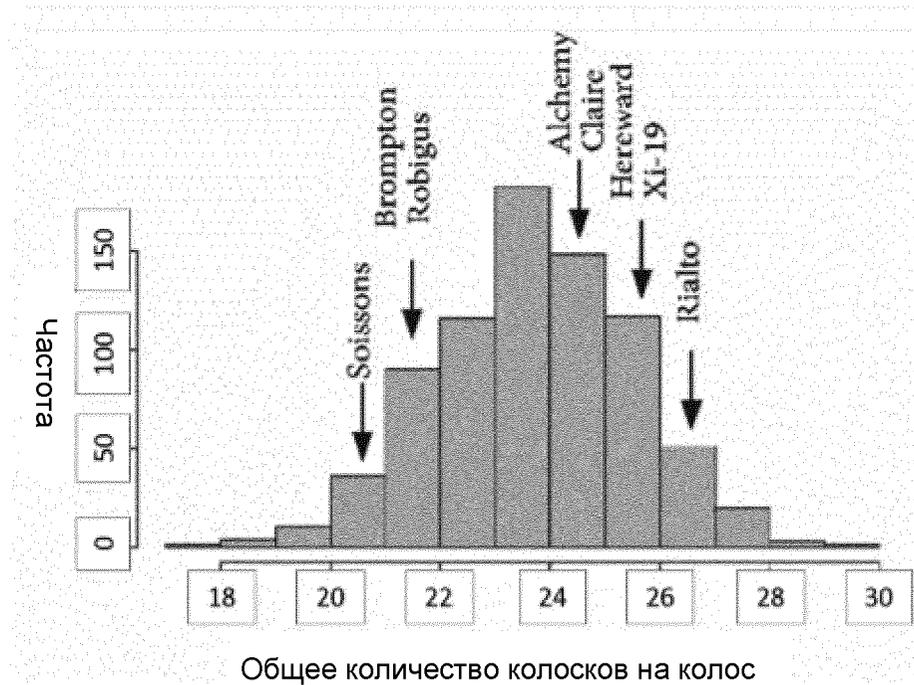
29. Способ идентификации и/или отбора растения пшеницы, содержащего аллель гена, положительно влияющего на количество колосков на колос, включающий этап идентификации наличия в геноме растения пшеницы нуклеиновой кислоты, имеющей нуклеотидную последовательность SEQ ID NO: 5 от нуклеотидного положения 4399 до нуклеотидного положения 4513, или нуклеотидную последовательность, имеющую с ней, по меньшей мере, 90% идентичности последовательности.

30. Способ идентификации и/или отбора растения пшеницы, содержащего аллель гена, отрицательно влияющего на количество колосков на колос, включающий этап идентификации отсутствия в геноме указанного растения пшеницы нуклеиновой кислоты, имеющей нуклеотидную последовательность SEQ ID NO: 5 от нуклеотидного положения 4399 до нуклеотидного положения 4513.



Фигура 1

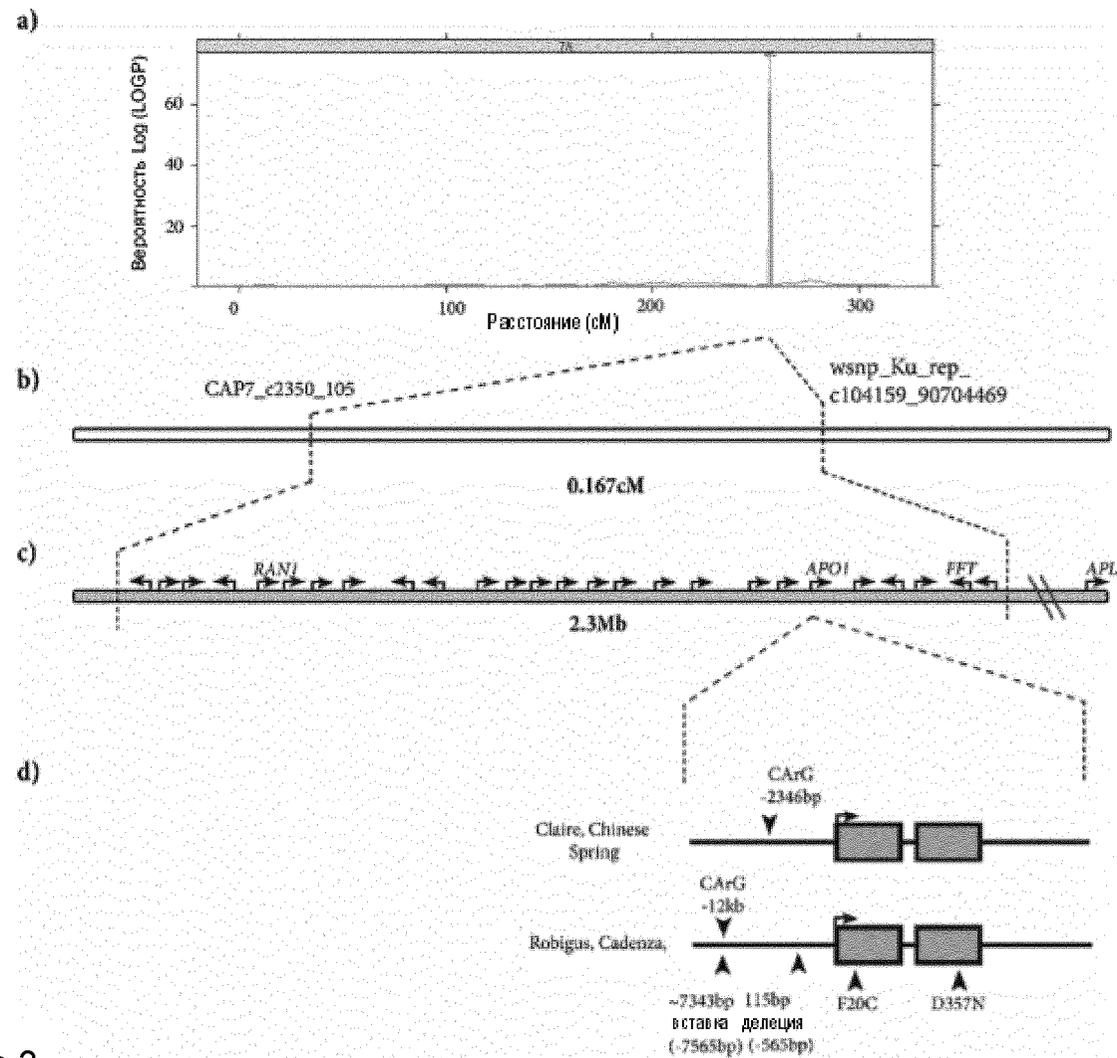
A.



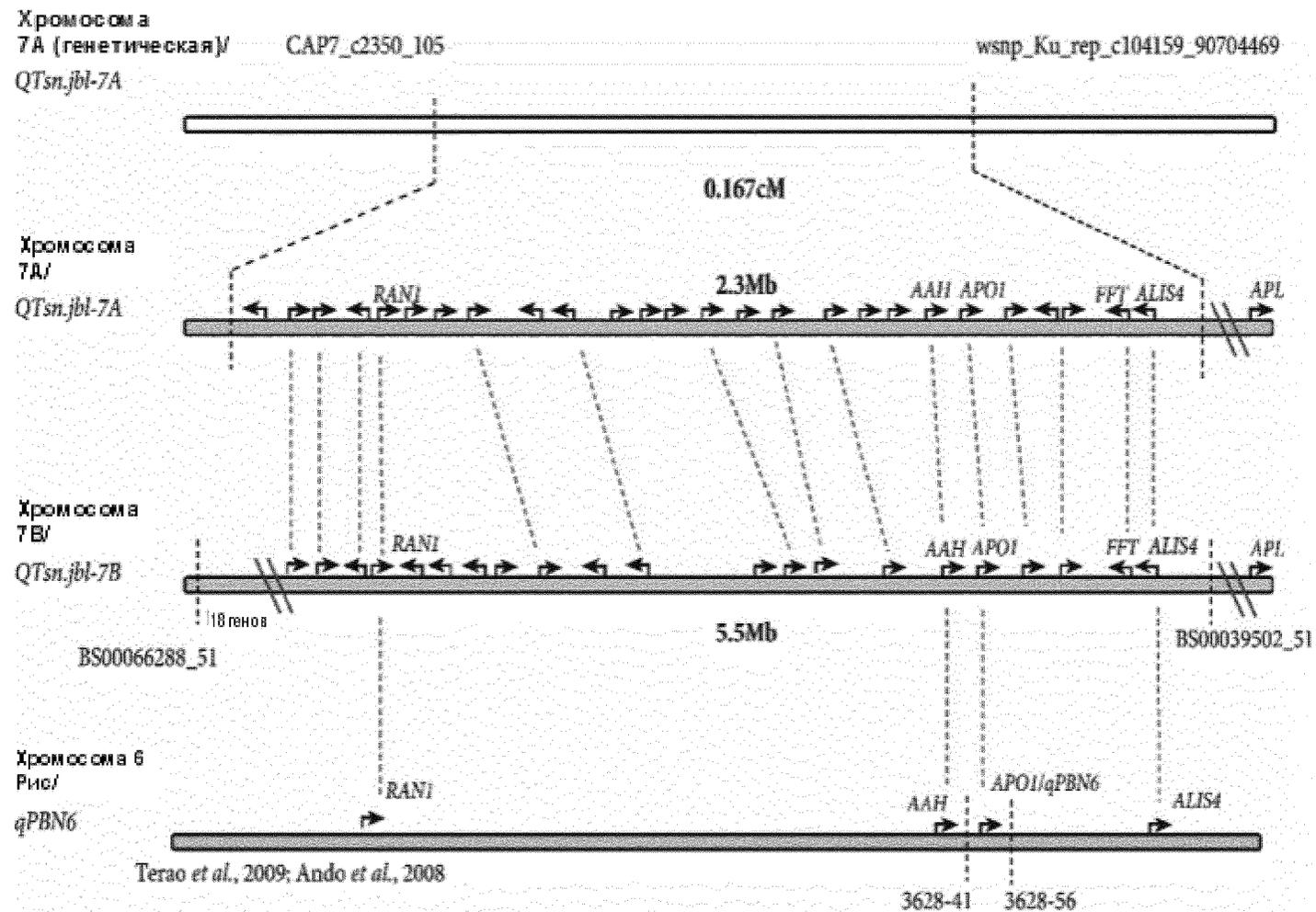
B.

Признак	Родители			MAGIC RILs			Наследуемость H <sup>2</sup> (%)
	Макс	Мин	Диапазон	Макс	Мин	Диапазон	
Общее количество колосков на колос (SPS)	26.40	20.32	1.30	29.80	17.70	1.68	84%

Фигура 2



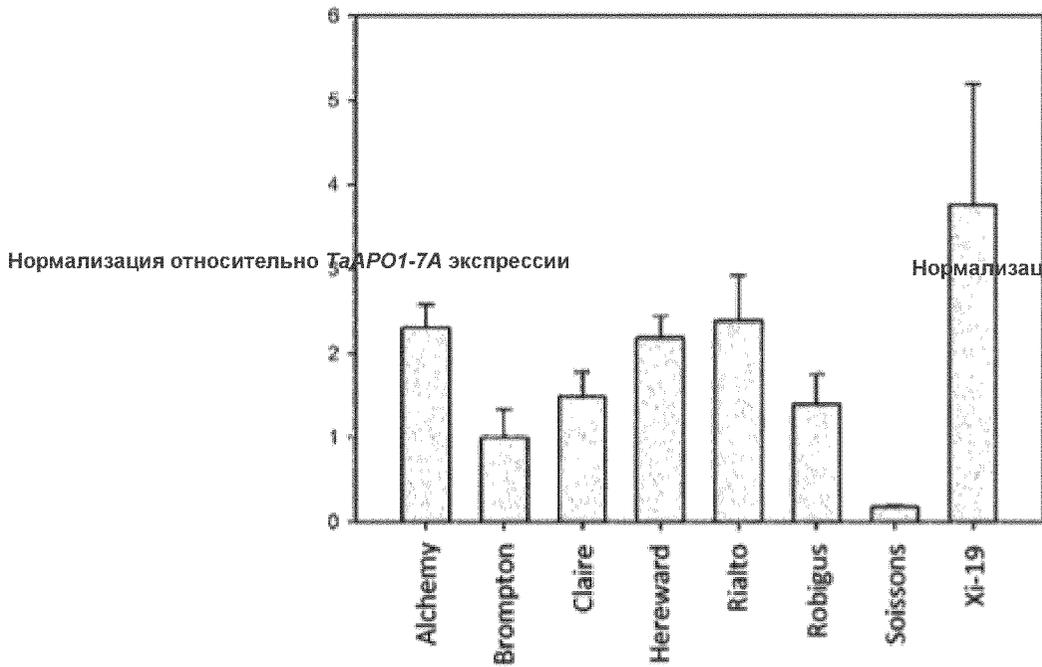
Фигура 3



Фигура 4

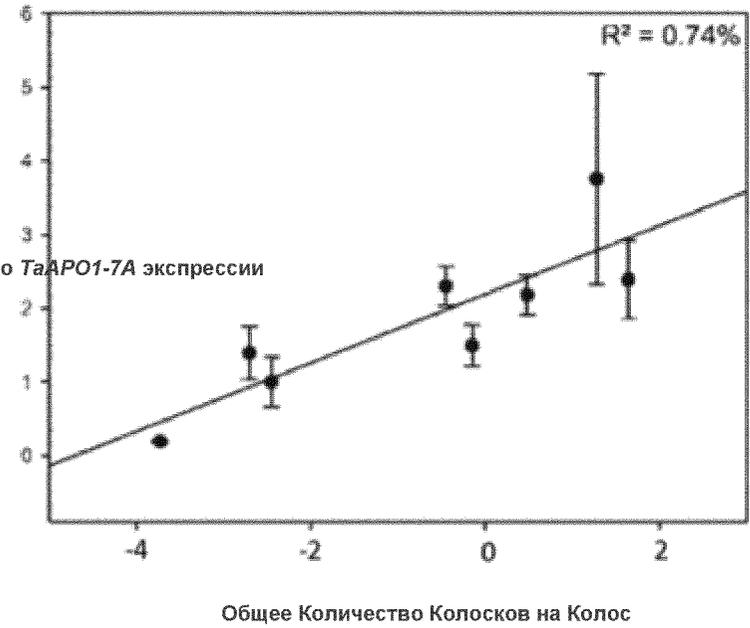
a)

*TaAPO1-7A* экспрессия в Основателях MAGIC



b)

*TaAPO1-7A* экспрессия против Общего Количества Колосков



Фигура 5