

(19)



**Евразийское
патентное
ведомство**

(21) **202291119** (13) **A1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2023.01.26

(51) Int. Cl. *A01H 1/04* (2006.01)
A01H 6/46 (2018.01)
A01H 5/10 (2018.01)

(22) Дата подачи заявки
2020.10.08

(54) СПОСОБЫ ПОЛУЧЕНИЯ МУТАНТНЫХ РАСТЕНИЙ

(31) 19202380.2

(32) 2019.10.10

(33) EP

(86) PCT/EP2020/078321

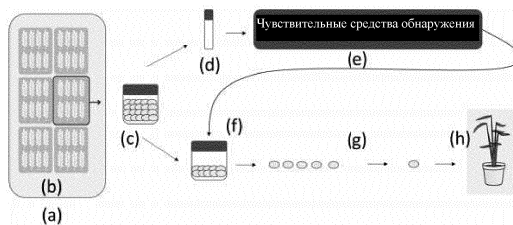
(87) WO 2021/069614 2021.04.15

(71) Заявитель:
КАРЛСБЕРГ А/С (DK)

(72) Изобретатель:
Скадхауге Биргитте, Кнудсен Сорен,
Хамбраеус Густав, Вендт Тони,
Расмуссен Магнус, Тулин Остерберг
Джеппе (DK)

(74) Представитель:
Гизатуллин Ш.Ф., Угрюмов В.М.,
Гизатуллина Е.М., Строкова О.В. (RU)

(57) Изобретение обеспечивает способы получения растений со специфической заранее заданной(ыми) мутацией(ями) в одном или нескольких NOI(s). Специфическая заданная(ые) мутация(и) предпочтительно может привести к идентификации растений, обладающих желаемыми признаками.



A1

202291119

202291119

A1

Способы получения мутантных растений

Область, к которой относится изобретения

Настоящее изобретение обеспечивает способы получения растений со специфической заданной мутацией(ями) в одном или нескольких представляющих интерес нуклеотидах (NOI(s)). Конкретная заданная мутация(и) предпочтительно может привести к идентификации растений, обладающих желаемыми признаками.

Предпосылки создания изобретения

Генетическая изменчивость ответственна почти за все фенотипические и генотипические различия в размножающейся популяции. Генетическая изменчивость возникает естественным путем, обычно вследствие экологических стрессов, воздействующих на организм, например, вследствие повышенного воздействия солнечной радиации, экстремальных условий почвы или других абиотических или биотических стрессов. Эта вызванная стрессом изменчивость является одной из движущих сил природного разнообразия, адаптации и эволюции организма.

Традиционный метод селекции использует эту генетическую изменчивость для выведения сельскохозяйственных культур, адаптированных к местным условиям. Одним из способов увеличения генетической изменчивости сверх того, что уже имеется в размножающейся популяции, является повышение уровня воздействия стресса, чтобы активно стимулировать дополнительную генетическую изменчивость. Это было успешно сделано за последние 90 лет путем мутационной селекции, в результате которой были созданы тысячи коммерческих сортов сельскохозяйственных культур, обладающих полезными признаками, основанными на индуцированной генетической изменчивости. Этот тип индукции мутаций обычно используют посредством создания уникальной размножающейся популяции, состоящей из мутагенизированной зародышевой плазмы, которая затем функционирует в качестве источника для идентификации признаков и дальнейшего обеспечения родительских организмов для генетического скрещивания с существующими элитными селекционными линиями. Примечательной характеристикой мутагенизированной размножающейся популяции является то, что популяция охватывает значительно увеличенную генетическую изменчивость по сравнению с естественной популяцией. Для достижения высокой частоты мутаций часто выбирают обработку методом мутагенеза, который вызывает гибель значительной части обработанного материала вследствие накопления слишком большого количества мутаций с вредными

фенотипическими эффектами. Чтобы выбрать оптимальную обработку методом мутагенеза, обычно готовят серии мутационных обработок и сравнивают способность обработанного организма прорасти в нормальных условиях. Это обычно называют кривой уничтожения «kill-curve». Кривую уничтожения используют в качестве инструмента для определения того, при какой концентрации и продолжительности обработки погибает 50% мутагенизированного материала. Эту обработку затем используют для создания мутагенизированной размножающейся популяции, например, злаковых зерен.

Мутационные обработки могут быть осуществлены с помощью различных химических веществ или доз облучения. Конечный эффект на организм определяется временем подвергания воздействию мутагена, а также относительной концентрацией мутагена. Эта обработка во многом зависит от организма.

Выбор мутагена может зависеть от природы организма, подлежащего мутации, поскольку некоторые мутагены лучше действуют на одни организмы, чем на другие. Время воздействия также может варьироваться от нескольких часов до многих часов, и химические концентрационные нагрузки могут различаться в зависимости от времени воздействия, природы организма, ткани и химического мутагена.

Индукцированные мутации обычно являются гетерозиготными и вызывают химерный генотип растения в изначально мутагенизированной культуре. Это означает, что зерна, собранные из мутагенизированной регенеративной части растения, такой как мутагенизированное семя, регенеративные части M1, обладают высокой степенью генетического разнообразия. Следовательно, мутагенизированные размножающиеся популяции обычно представлены регенеративными частями M2 или регенеративными частями (такими как семена) из более поздних поколений. Таким образом, большинство потомства растения сегрегируют в соответствии с менделевской генетикой только по двум аллелям (мутантный и дикий тип), чтобы позволить гомозиготным рецессивным мутациям быть выраженными (видимыми) на фенотипическом уровне. Это также позволяет селекционерам использовать методы скрининга, характеризующиеся низкой чувствительностью. По мере дальнейшего развития методов скрининга селекционеры начали использовать повышенную чувствительность инструментов скрининга, комбинируя ДНК нескольких регенеративных частей M2, таких как семена или растения, для скрининга. В то время как известные средства скрининга генетических вариаций в мутагенизированных популяциях были ограничены скринингом пулов из 5-10 растений (гетеродуплексный

анализ, секвенирование по Сэнгеру, анализ кривой плавления с высоким разрешением), сегодня некоторые из этих пулов могут достигать 100 образцов ДНК при анализе с помощью передовых технологий секвенирования. Однако, несмотря на заметные успехи, остаются проблемы не только с подготовкой многочисленных отдельных образцов ДНК, но и с идентификацией потенциальных мутантных растений, которые могут войти в программу селекции с рядом фоновых мутаций, т.е. тех, которые не имеют отношения к интересующему признаку, но потенциально влияют на дополнительные характеристики сельскохозяйственных культур, такие как урожайность и характеристики размножения.

Не существует четкого определения насыщенной библиотеки мутаций культур, но основная идея мутагенизированной размножающейся популяции состоит в том, что во всей популяции представлены несколько мутантных аллелей для каждого гена. Чтобы размер библиотеки был небольшим, каждый член популяции должен нести мутации в нескольких генах. В типичной библиотеке ячменя каждый член популяции несет 100-200 мутаций в кодирующих областях. Эти дополнительные кодирующие мутации являются основной проблемой для селекционера после того, как желаемый признак был идентифицирован. По-настоящему насыщенная библиотека мутаций будет включать мутантный аллель для каждой аминокислоты в каждом гене. Чтобы достичь этого, частота мутаций или плотность мутаций должны быть выше или популяция больше. Однако заметно более низкая приспособленность растений остается ограничивающим фактором, связанным с использованием библиотек, приготовленных с повышенными дозами, вызывающими мутации. Кроме того, труд и время нескольких тысяч человек, затрачиваемые на сбор библиотеки, а также затраты на скрининг такой библиотеки ограничивают возможности расширения библиотек. Один из подходов к преодолению упомянутых выше барьеров относится к применению более высоких уровней мутагена при создании библиотеки мутантов. В этом отношении дальнейшее уменьшение приспособленности растений ограничивает увеличение дозы, вызывающей мутации, для каждой особи. Следовательно, действительно насыщенных мутагенизированных библиотек не существует.

После мутагенеза с обработкой, которая обеспечивает высокую плотность мутаций или высокую частоту мутаций, обычно проводят обратное скрещивание мутантных культур с культурой дикого типа или родительской культурой для устранения мутаций, которые могли возникнуть в других частях генома. Это может

помочь устранить «побочные эффекты», возникающие в результате нежелательных нецелевых мутаций.

Краткое описание изобретения

До сих пор целью растениеводов-селекционеров было использование размножающихся популяций с максимально возможной скоростью или частотой мутаций, чтобы увеличить вероятность идентификации растений, несущих конкретный признак. Напротив, настоящее изобретение обеспечивает способы уменьшения числа кодирующих мутаций на особь в мутагенизированной размножающейся популяции. Другими словами, изобретение обеспечивает способы, позволяющие использовать популяцию растений или их способных к регенерации частей, имеющих низкую скорость мутагенеза.

После мутагенеза с обработкой, которая обеспечивает высокую плотность мутаций или высокую частоту мутаций, обычно проводят обратное скрещивание мутантных культур с культурой дикого типа или родительской культурой для устранения мутаций, которые могли произойти в других частях генома. Это может помочь устранить «побочные эффекты», возникающие в результате нежелательных нецелевых мутаций. В практических условиях улучшение проблем, связанных с селекцией, как описано в настоящем документе, приводит к радикально меньшему количеству фенотипических недостатков у организмов размножающейся популяции. Будет меньше тех индуцированных мутаций, которые со временем, т.е. после фактического процесса идентификации мутантов, придают сложные или даже невыгодные фенотипы - особенно мутации, которые придают фенотипические недостатки в ответ на экологические стрессы, связанные с температурой и осадками. Кроме того, можно ожидать, что мутанты, идентифицированные с помощью настоящих способов, будут иметь повышенную жизнеспособность, поскольку они могут избежать генерирования мутаций, которые могут стать очевидными или актуальными только спустя много времени после того, как мутанты были получены. Растения, несущие множество мутаций, могут сначала казаться такими же жизнеспособными, как и родительское растение, но в определенных условиях, например, засуха, холод, обильные дожди и т.д. может быть выявлено, что они содержат дополнительные, еще не обнаруженные мутации, которые могут создавать проблемы для жизнеспособности.

Низкая скорость мутагенеза может выражаться в малом количестве несинонимичных мутаций в кодирующих областях. Небольшое количество кодирующих мутаций на особь снижает пагубные эффекты дополнительных мутаций в

мутагенизированных размножающихся популяциях. Для химических мутагенов плотность мутаций на растение зависит от времени воздействия мутагена, культуры, ткани, природы и концентрации химического мутагена. Для конкретного растения можно создать библиотеки с особями с низкой плотностью мутаций путем определения оптимальной концентрации и/или времени воздействия конкретного химического мутагена с помощью исследования «кривой уничтожения». Для способов по изобретению предпочтительно использовать концентрацию и/или время воздействия данного мутагена, которые не приводят к потере приспособленности или утрате жизнеспособности мутагенизированной популяции растений. В идеальном случае каждое растение в мутагенизированной популяции растений, подлежащей использованию со способами по настоящему изобретению, имеет от 1 до 30 несинонимичных мутаций, т.е. мутаций в кодирующих областях, вызывающих аминокислотную замену в транслируемом белке.

Вместо количества несинонимичных мутаций, низкая скорость мутагенеза также может быть выражена в виде высокого процента генов, не содержащих несинонимичных мутаций. Таким образом, в некоторых вариантах осуществления низкая скорость мутагенеза может означать, что по меньшей мере 99%, такое как по меньшей мере 99,8%, например, по меньшей мере 99,9% генов растения, например, в способных к регенерации частях поколения M1 свободны от несинонимичных мутаций.

То, что конкретно можно считать низкой скоростью мутагенеза, может отличаться от растения к растению, в частности, исходя из плоидности растения. Таким образом, растения с высокой плоидностью обычно могут выдерживать более высокую скорость мутагенеза, и, таким образом, «низкая скорость мутагенеза» у таких растений может быть намного выше, чем «низкая скорость мутагенеза» у диплоидного растения. В целом, низкая скорость мутагенеза в диплоидном растении может означать, что по меньшей мере 99,8%, предпочтительно по меньшей мере 99,9% всех генов в M1 свободны от несинонимичных мутаций. В целом, низкая скорость мутагенеза у растений со значением плоидности выше двух, например у тетраплоидных или гексаплоидных растений, может означать, что по меньшей мере 98,0%, предпочтительно по меньшей мере 99,0% всех генов в M1 свободны от несинонимичных мутаций.

Низкая скорость мутагенеза также может выражаться в виде низкой частоты мутаций. Например, в диплоидном растении низкой скоростью мутагенеза может

являться частота мутаций 1 на по меньшей мере 500000, например 1 на по меньшей мере 750000, например 1 на по меньшей мере 1000000. В тетраплоидном растении низкой скоростью мутагенеза может быть частота мутаций 1 на по меньшей мере 70000, например 1 на по меньшей мере 100000, например 1 на по меньшей мере 300000. В гексаплоидном растении низкой скоростью мутагенеза может быть частота мутаций 1 на по меньшей мере 40000, например 1 на по меньшей мере 80000, например 1 на по меньшей мере 200000.

Низкая скорость мутагенеза может быть достигнута путем использования мягких условий мутагенеза. Например, библиотека, содержащая растения заданных видов, имеющих заданное количество генов, подвергается случайному мутагенезу в определенных условиях, например, специфической концентрации мутагена в течение определенного периода времени, как показано в примерах 3-5 и 12. Мутанты идентифицируют, например, в заданной области заданной длины, как описано в настоящем документе. Количество мутантов, идентифицированных в этой области, может быть использовано для расчета плотности мутаций для заданных условий мутагенеза (концентрация и время).

Плотность мутаций также можно определить путем секвенирования по меньшей мере всех открытых рамок считывания во всех мутагенизированных растениях или их способных к регенерации частях и подсчета количества полученных мутантов. Таким образом, зная размер генома, можно также рассчитать плотность мутаций.

Размер используемой библиотеки обычно зависит от количества генов для данного растения. Например, ячмень имеет приблизительно 30000 генов, что означает, что библиотека мутантов, состоящая из особей, каждая из которых содержит только 1 несинонимичную мутацию в кодирующих областях, должна содержать по меньшей мере 30000 особей, чтобы обеспечить мутантный аллель для каждого гена. Однако в библиотеке, полученной с помощью случайного мутагенеза, мутации не будут статистически полностью равномерно распределены, и, таким образом, требуется значительно больше, чем 30000 подвергнутых случайному мутагенезу особей, чтобы обеспечить библиотеку, содержащую один мутантный аллель для каждого гена. Если учесть, что гены ячменя в среднем содержат 1000 кодонов, то для идентификации одной мутации в каждом кодоне каждого гена требуется библиотека по меньшей мере из 30 миллионов особей. Настоящие способы можно применять для скрининга больших библиотек, при этом значительно сокращая время обратного скрещивания или полностью устраняя необходимость в обратном скрещивании.

В качестве альтернативы оптимальный размер библиотеки можно рассчитать на основе частоты мутаций и количества мутаций, на предмет чего был проведен скрининг, как описано в настоящем документе. Для получения оптимального размера библиотеки определенное количество способных к регенерации частей (N_p) должно быть подвергнуто мутагенезу. Оптимальное количество N_p можно рассчитать на основе оптимального размера библиотеки, числа GECN и уровня сбора видов растений после мутагенеза.

Для скрининга таких больших библиотек изобретение обеспечивает способ объединения в пулы растений подвергнутой мутагенезу размножающейся популяции с последующим использованием высокочувствительных способов обнаружения, например способов, используемых в настоящее время в диагностике, основанных на обнаружении редкой ДНК, т.е. циркулирующей опухолевой ДНК (ctDNA). Такие способы включают, например, цифровую ПЦР (dPCR) и/или использование сложной химии олигонуклеотидов, которые являются ультраспецифическими, и способны идентифицировать и амплифицировать очень редкие фрагменты ПЦР в сложной смеси молекул ДНК. Таким образом, пул растений может быть разделен на подпулы для скрининга. Оптимальные размеры подпулы могут быть рассчитаны на основе числа GECN и чувствительности обнаружения используемых средств обнаружения.

Описанные в настоящем документе способы представляют основу для высокопроизводительных методов скрининга, которые применимы для скрининга очень больших библиотек мутантов, которые могут быть генерированы из элитных сортов или передовых селекционных линий, а также для скрининга всех размножающихся популяций или коллекций зародышевой плазмы. Эти способы также можно использовать для идентификации генетических вариантов у нетрансформируемых видов сельскохозяйственных культур, которые не поддаются генетической модификации и технологиям редактирования генов.

Таким образом, изобретение обеспечивает способ идентификации мутантного растения предварительно определенного вида, несущего одну или более мутаций в представляющем интерес нуклеотиде(ах) [NOI(s)], в заданной целевой последовательности, при этом указанный способ включает следующие стадии:

а) обеспечение способных к регенерации частей родительского растения указанного вида, где обеспечено по меньшей мере 5000, такое как по меньшей мере 10000, такое как по меньшей мере 15000, такое как по меньшей мере 20000, такое как

по меньшей мере 25000, такое как по меньшей мере 30000 способных к регенерации частей;

b) подвергание указанных способных к регенерации частей стадии мутагенеза, предпочтительно случайного мутагенеза, тем самым создавая пул способных к регенерации частей поколения M0, представляющих множество генотипов;

c) выращивание указанных способных к регенерации частей поколения M0 в зрелые растения поколения M1 и получение способных к регенерации частей поколения M1 из указанных зрелых растений;

d) необязательно повторение предыдущей стадии X раз для получения растений поколения M(1+X), включающих способные к регенерации части поколения M(1+X);

e) разделение способных к регенерации частей множества растений поколения M1 или M(1+X) на подпулы, при этом каждый подпул содержит способные к регенерации части, представляющие множество генотипов, где все способные к регенерации части от одного данного зрелого растения помещают в один и тот же подпул;

f) подготовку образцов ДНК, таких как образцы геномной ДНК (gDNA) или образцы кДНК (cDNA), полученные из образцов мРНК, из каждого подпула;

g) идентификацию подпула(ов), включающего ДНК, содержащую мутацию(и), с использованием чувствительных средств обнаружения;

h) идентификацию способных к регенерации частей в указанном подпуле, содержащем указанную мутацию, тем самым идентифицируя указанное мутантное растение,

где указанная стадия мутагенеза приводит к скорости мутагенеза от 1 до 30 несинонимичных мутаций на клетку способной к регенерации части поколения M1, предпочтительно, где способная к регенерации часть представляет собой вегетативную ткань, семя, зерно, зародышевую клетку в семени, зерне или зародыше указанного растения, пыльцу или зародыш, предпочтительно способная к регенерации часть представляет собой семя.

ОПИСАНИЕ ЧЕРТЕЖЕЙ

Фигура 1. Пример процедуры скрининга. (a) Подвергнутые мутагенезу способные к регенерации части, в данном случае семена (M0), выращивали до зрелости

в поле, (b) собирали в подпулы по 300 растений, и (c) семена, собранные в виде образцов насыпью, представляющих каждый пул. Образцы насыпью разделяли на фракции, одну фракцию (d) использовали для выделения ДНК, другую сохраняли в качестве банка семян для последующего анализа. ДНК из образцов насыпью (d) анализировали с использованием чувствительных средств обнаружения (e). Исходный пул из 300 растений, несущих представляющую интерес мутацию, идентифицировали (f) и использовали для анализа отдельных семян (g). (h) Мутантное растение можно прорастить из идентифицированного семени.

Фигура 2. Размножение мутантных зерен ячменя. Эта фигура иллюстрирует размножение мутантных зерен, в данном случае ячменя, подвергнутых мутагенезу азидом натрия. Мутантные семена поколения M0 дали рост растению поколения M1. Семена растения поколения M1 дали рост растению поколения M2. Семена растения M2 дали рост растению поколения M3, которое можно использовать для скрининга и скрещивания.

Фигура 3. Разделение пула на подпулы. Приводится пример способа разделения пула на подпулы.

ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

Определения

Термин «молчащая мутация» относится к мутации, возникающей в последовательности, например, геномной последовательности организма, которая не приводит к изменению аминокислотной последовательности. Таким образом, молчащая мутация может представлять собой любую мутацию в некодирующей области генома, или она может представлять собой мутацию в кодирующей области генома, которая не изменяет аминокислотную последовательность продукта трансляции. Синонимичные мутации относятся к замене нуклеотидов, которая не изменяет аминокислотную последовательность продукта трансляции и, следовательно, относится к типу молчащих мутаций.

Термин «несинонимичная мутация», используемый в настоящем документе, относится к мутации, возникающей в последовательности, например, в геномной последовательности организма, которая действительно приводит к изменению аминокислотной последовательности. Таким образом, несинонимичная мутация возникает в кодирующих областях и поэтому может представлять собой любую мутацию в кодирующей области генома, которая изменяет аминокислотную

последовательность продукта трансляции, например, путем кодирования другой аминокислоты (например, миссенс-мутации, сквозные мутации), путем введения преждевременного стоп-кодона (нонсенс-мутации) и/или путем введения сдвига рамки считывания. В настоящем документе этот термин будет использован взаимозаменяемо с термином «кодирующая мутация».

Термин «умеренный уровень мутагенеза» или «низкий уровень мутагенеза», используемый в настоящем документе, относится к обработке методом мутагенеза, приводящей к плотности мутаций от 1 до 30 несинонимичных мутаций (в кодирующих областях, за исключением интронов), предпочтительно при измерении в клетке способной к регенерации части, например, поколения M1. Плотность мутаций может быть измерена путем секвенирования всех открытых рамок считывания генома(ов). То, что считается «низким уровнем мутагенеза», может зависеть от типа растения. Таким образом, термин «низкий уровень мутагенеза» может означать в отношении диплоидных растений, что в среднем по меньшей мере 99,8%, предпочтительно по меньшей мере 99,9% всех генов свободны от несинонимичных мутаций в способных к регенерации частях поколения M1, и в отношении растений более высокой плоидности «низкий уровень мутагенеза» может означать, что в среднем по меньшей мере 98,0%, предпочтительно по меньшей мере 99,0% всех генов свободны от несинонимичных мутаций в способных к регенерации частях поколения M1.

Термин «частота мутаций» — сокращенно «Mf» — относится к количеству мутаций в пределах данной длины нуклеиновой кислоты, например, количеству мутаций, полученных во всем геноме. В частности, частота мутаций может быть рассчитана как:

$$Mf = \frac{N_{nc}}{N_{na}}$$

где

N_{nc} представляет собой общее количество нуклеотидных замен;

N_{na} представляет собой количество подвергнутых анализу нуклеотидов.

Другими словами, $Mf = 1/\text{Количество нуклеотидов на мутацию}$.

В вариантах осуществления изобретения, где точная частота мутаций неизвестна, может быть использована предполагаемая частота мутаций. Таким образом, частота мутаций (Mf) может представлять собой фактическую частоту мутаций или предполагаемую частоту мутаций.

Используемый в настоящем документе термин «скорость мутагенеза» может относиться либо к плотности мутаций, либо к частоте мутаций.

Термин «аллель» относится к конкретной версии или состоянию гена. Термин «мутантный аллель», используемый в настоящем документе, относится к гену, несущему одну или несколько заданных мутаций в NOI. Когда мутация представляет собой делецию всего гена, мутантный аллель может также представлять собой аллель, в котором отсутствует указанный ген.

Термин «приблизительно» используется в настоящем документе по отношению к числам, относящимся к $\pm 10\%$, предпочтительно $\pm 5\%$, например, к $\pm 1\%$.

Используемый в настоящем документе термин «блокирующий зонд» относится к олигонуклеотиду, который не может быть удлинен на 3'-конце с помощью ДНК-полимеразы. Блокирующий зонд, как правило, представляет собой олигонуклеотид, идентичный или комплементарный целевой последовательности, включая эталонный NOI, связанный с блокирующим агентом, который ингибирует удлинение блокирующего зонда ДНК-полимеразой.

Термин «кодирующая область» или «кодирующая последовательность (CDS)», используемый в настоящем документе, относится к части ДНК или РНК, за исключением интронов, которая при трансляции приводит к образованию белка.

Термин «генотип», используемый в настоящем документе, относится к организму, содержащему определенный набор генов. Таким образом, два организма, содержащие идентичные геномы, относятся к одному и тому же генотипу. Генотип организма по отношению к конкретному гену определяется аллелями, которые несет этот организм. У диплоидных организмов генотип данного гена может быть AA (гомозиготный, доминантный), Aa (гетерозиготный) или aa (гомозиготный, рецессивный).

Используемый в настоящем документе термин «фенотип» представляет собой совокупность наблюдаемых характеристик или признаков организма.

Термин «генетически эффективное количество клеток» — сокращенно «GECN» — относится к количеству клеток растения, которые в конечном итоге образуют гаметы. Такие клетки в настоящем документе также называются как «клетки зародышевой линии». Здесь он рассматривается как число событий, дающих начало независимым генотипам в потомстве. Показатель GECN является важным для расчета ожидаемого разнообразия в мутантной популяции, поскольку каждая генетически эффективная клетка будет вносить в библиотеку различные события мутации.

Например, показатель GECN может быть рассчитан, как описано ниже в настоящем документе в примере 12. Для способов по изобретению не требуется использовать фактический показатель GECN. Таким образом, GECN может представлять собой фактический показатель GECN или ожидаемый показатель GECN. Чаще всего показатель GECN равен по меньшей мере 2 и, таким образом, ожидаемый показатель GECN может быть равен, например, 2. Показатели GECN для ряда выбранных растений представлены в ниже в таблице 1.

Таблица 1

<i>Brachypodium distachyon</i> (коротконожка)	-4
<i>Avena sativa</i> (овес посевной)	6
<i>Arabidopsis thaliana</i> (резуховидка таля)	2
<i>Glycine max</i> (соя культурная)	2
<i>Nicotiana glauca</i> (табак)	2
<i>Medicago sativa</i> (люцерна посевная)	3
<i>Linum usitatissimum</i> (лен обыкновенный)	4
<i>Zea mays</i> (кукуруза)	4+
<i>Hordeum vulgare</i> (ячмень обыкновенный)	5-6
<i>Triticum aestivum</i> (пшеница мягкая)	5+
<i>Brassica napus</i> (рапс)	4

Термин «средство детекции для обнаружения последовательностей нуклеиновых кислот» относится к любым средствам детекции, позволяющим обнаружить определенные последовательности нуклеиновых кислот. Таким образом, указанные средства детекции могут обнаруживать мутации в представляющих интерес нуклеотидах и дифференцировать их от последовательностей дикого типа. Многие средства детекции основаны на ПЦР.

Термин «скорость прорастания» — сокращенно «Gr» — относится к средней скорости прорастания данных видов растений после мутагенеза. Для определения скорости прорастания ряд способных к регенерации частей инкубировали в условиях, обеспечивающих рост, и определяли процент способных к регенерации частей, инициирующих рост. Например, если способные к регенерации части представляют собой семена, то заданное количество семян инкубируют в условиях, обеспечивающих прорастание, и после соответствующей инкубации определяют процент проросших

семян. Если не указано иное, скорость прорастания представляют в %. Для способов по изобретению не требуется использовать фактическую скорость прорастания (Gr). Таким образом, скорость прорастания (Gr) может представлять собой фактическую скорость прорастания или ожидаемую скорость прорастания.

Термин «уровень сбора» — сокращенно «Hr» — относится к среднему количеству растений, собранных на одну способную к регенерации часть, подвергнутую мутагенезу. Для некоторых растений уровень сбора приблизительно равен скорости прорастания. Однако для других видов растений высевается гораздо большее количество способных к регенерации частей по сравнению с количеством фактически собранных растений. Если не указано иное, уровень сбора представляют в %. Для способов по изобретению не требуется использовать фактический уровень сбора (Hr). Таким образом, уровень сбора (Hr) может представлять собой фактический уровень сбора или ожидаемый уровень сбора.

Термин «клетки зародышевой линии» относится к клеткам, которые передают свои хромосомы или наследуемый генетический материал своему потомству, своим потомкам или следующему поколению.

Термин «чувствительные средства детекции» относится к средствам детекции, которые позволяют обнаруживать один мутантный организм в библиотеке, содержащей по меньшей мере 300 организмов, даже в тех случаях, когда генотип мутантного организма отличается от генотипов немутантных организмов максимум на один нуклеотид. Предпочтительно они позволяют обнаруживать один мутантный геном в библиотеке из 300 геномов.

Используемый в настоящем документе термин «зонд для обнаружения мутантов» относится к олигонуклеотиду, необязательно связанному с обнаруживаемым средством, причем указанный олигонуклеотид идентичен или комплементарен целевой последовательности, включая заданную мутацию NOI.

Используемый в настоящем документе термин «ПЦР» относится к полимеразной цепной реакции. ПЦР представляет собой реакцию амплификации нуклеиновых кислот. Метод основан на циклическом температурном воздействии и состоит из циклов повторного нагревания и охлаждения реакционной смеси для получения последовательного плавления и ферментативной репликации указанной ДНК. На первой стадии две нити, образующие двойную спираль ДНК, физически разделяются при высокой температуре в процессе, также известном как плавление ДНК. На второй стадии температуру снижают, обеспечивая ферментативную

репликацию ДНК. ПЦР может также включать инкубацию при дополнительной температуре, чтобы усилить отжиг праймеров и/или оптимизировать температуру (температуры) для репликации. В ПЦР температура обычно колеблется между различными температурами в течение нескольких циклов.

Используемый в настоящем документе термин «реагенты для ПЦР» относится к реагентам, которые добавляют к ПЦР в дополнение к образцу и набору праймеров. Реагенты для ПЦР содержат по меньшей мере нуклеотиды и полимеразу нуклеиновых кислот. Кроме того, реагенты для ПЦР могут содержать другие соединения, такие как соль(и) и буфер(ы).

Термин «dPCR» относится к цифровой полимеразной цепной реакции. Ее можно использовать для прямого количественного определения и клональной амплификации нитей нуклеиновых кислот, включая ДНК, кДНК или РНК. dPCR измеряет количество нуклеиновых кислот более точно, чем ПЦР. В традиционной ПЦР одну реакцию проводят на один образец. dPCR также основывается на проведении одной реакции внутри образца, однако образец разделен на компартменты, т.е. образец разделен на большое количество сегментов, и реакцию проводят в каждом сегменте индивидуально.

Термин «ddPCR» относится к капельной цифровой полимеразной цепной реакции. В ddPCR проводят одну или несколько ПЦР-амплификаций, при этом каждая реакция разделена на множество капель водно-масляной эмульсии, так что ПЦР-амплификация целевой последовательности может происходить в каждой индивидуальной капле. ddPCR является примером разделенной на компартменты dPCR.

Термин «вероятность успеха» — сокращенно «PS» — относится к вероятности того, что желаемая мутация может быть идентифицирована из данной библиотеки способных к регенерации частей, которые были подвергнуты случайному мутагенезу. Квалифицированный специалист может выбрать подходящую вероятность успеха. В целом предпочтительно, чтобы вероятность успеха составляла по меньшей мере 85%, предпочтительно по меньшей мере 90%, более предпочтительно по меньшей мере 95%, такую как по меньшей мере 98%.

Используемый в настоящем документе термин «репродукция» относится как к половому, так и к бесполому размножению.

Таким образом, репродукция может представлять собой размножение организма клональным образом (также известное как «бесполая репродукция»). Репродукция

также может представлять собой получение потомства организма, где потомство содержит аллель(и) родительского организма. Таким образом, репродукция организма, содержащего мутантный аллель, может относиться к получению потомства указанного организма, где потомство содержит мутантный аллель. Предпочтительно мутантный аллель несет одну или несколько мутаций в NOI(s).

Используемый в настоящем документе термин «способные к регенерации части организма» относится к любой части организма, которая при соответствующих условиях может вырасти в целый организм. Например, если организм представляет собой растение, этот термин относится к любой части растения, которая может регенерироваться в целое растение. Таким образом, способная к регенерации часть растения может представлять собой любой «материал для размножения» указанного растения. Способная к регенерации часть растения может представлять собой, например, семя, зерно, клетку зародышевой линии в семени, зерне или зародыше указанного растения, вегетативную ткань, пыльцу или зародыш указанного растения. В настоящем документе этот термин будет использоваться взаимозаменяемо с термином «репродуктивная часть». Термин «клетки зародышевой линии» относится к клеткам, которые передают свои хромосомы или наследуемый генетический материал своему потомству, потомкам или следующему поколению.

Используемый в настоящем документе термин «набор праймеров, фланкирующих целевую последовательность» относится к набору из двух праймеров, фланкирующих целевую последовательность, так что один праймер содержит последовательность, идентичную 5'-концу последовательности-мишени (также называемую «прямой праймер»), и один праймер содержит последовательность, комплементарную 3'-концу последовательности-мишени (также называемую «обратным праймером»). «Набор праймеров» может амплифицировать целевую последовательность при добавлении к ПЦР вместе с нуклеиновой кислотой, содержащей целевую последовательность, и реагентами для ПЦР в условиях, позволяющих амплифицировать указанную целевую последовательность.

Используемый в настоящем документе термин «целевая последовательность» относится к любой последовательности нуклеиновой кислоты, в пределах которой желательно создать или идентифицировать мутацию. Кроме того, целевая последовательность предпочтительно представляет собой последовательность нуклеиновой кислоты, которую можно амплифицировать с помощью технологии ПЦР с использованием праймеров, фланкирующих целевую последовательность. Кроме

того, целевая последовательность обычно содержит одну или несколько NOI(s). Изобретение обеспечивает способы получения и/или идентификации организма(ов), несущего(их) мутацию в указанном NOI. Целевая последовательность может, например, представлять собой последовательность нуклеиновой кислоты, ассоциированную с определенным признаком.

Используемый в настоящем документе термин «эталонный зонд для обнаружения» относится к олигонуклеотиду, необязательно связанному с обнаруживаемым средством, при этом олигонуклеотид является идентичным или комплементарным целевой последовательности, включая эталонный NOI. Как правило, целевая последовательность, включая эталонный NOI, соответствует целевой последовательности до мутагенеза. «Эталонный зонд для обнаружения» может также называться как «зонд дикого типа».

Термин «ячмень» в отношении способа получения напитков на основе ячменя, таких как пиво, в частности, при использовании для описания способа соложения, означает зерна ячменя. Во всех других случаях, если не указано иное, «ячмень» означает растение ячменя (*Hordeum vulgare*, L.), включая любую селекционную линию, сорт или разновидность, в то время как часть растения ячменя может представлять собой любую часть растения ячменя, например, внешнюю структуру растения, такую как листья, стебли, корни, цветы и зерна, известные как органы растения. Он также может включать любую ткань или клетки.

Как определено в настоящем документе, «злаковое» растение» является представителем семейства растений Poaceae, выращиваемых преимущественно из-за их содержащих крахмал семян или зерен. Злаковые растения включают, помимо прочего, ячмень (*Hordeum*), пшеницу (*Triticum*), рис (*Oryza*), кукурузу (*Zea*), рожь (*Secale*), овес (*Avena*), сорго (*Sorghum*) и тритикале, гибрид ржи-пшеницы.

Термин «качество зерна» в настоящем документе относится к характеристикам зерна, полученного из растения, такого как мутантное растение. Таким образом, этот термин может относиться к содержанию влаги в зерне, массе зерна, проценту обесцвеченных, разрушенных или поврежденных зерен, ломкости зерна, качеству помола, содержанию белка, содержанию масла и/или жизнеспособности зерна. Другие внутренние факторы включают цвет, состав, объемную плотность, запах и аромат, размер и форму.

Термин «% стерильности» в настоящем документе относится к проценту цветков в популяции растений заданного генотипа, которые не дают семян.

Способ идентификации мутантного растения

В настоящем документе обеспечен способ идентификации мутантного растения предварительно определенного вида, несущего одну или более мутаций в представляющем интерес нуклеотиде(ах) [NOI(s)] в заданной целевой последовательности, при этом указанный способ включает следующие стадии:

- a) обеспечение способных к регенерации частей N_p родительского растения указанного вида, где N_p представляет собой целое число, равное по меньшей мере 5000;
- b) подвергание указанных способных к регенерации частей стадии мутагенеза, предпочтительно случайного мутагенеза, тем самым создавая пул способных к регенерации частей поколения M_0 , представляющих множество генотипов;
- c) выращивание указанных способных к регенерации частей поколения M_0 в зрелые растения, тем самым получая способные к регенерации части поколения M_1 из указанных зрелых растений;
- d) необязательное повторение предыдущей стадии X раз с получением растений поколения $M(1+X)$, включающих способные к регенерации части поколения $M(1+X)$;
- e) разделение способных к регенерации частей множества растений поколения M_1 или $M(1+X)$ на подпулы, при этом каждый подпул включает способные к регенерации части, представляющие множество генотипов, где все способные к регенерации части от одного данного зрелого растения помещают в один и тот же подпул;
- f) получение образцов ДНК, таких как образцы cDNA, полученные из образцов мРНК, или образцы gDNA, из каждого подпула;
- g) идентификацию субпула(ов), включающего ДНК, содержащую мутацию(и), с использованием чувствительных средств обнаружения;
- h) идентификацию способных к регенерации частей и/или потомства способных к регенерации частей в указанном подпуле, содержащем указанную мутацию, тем самым идентифицируя указанное мутантное растение,

где указанная стадия мутагенеза приводит к скорости мутагенеза от 1 до 30 несинонимичных мутаций на клетку способной к регенерации части поколения M_1 .

Следует понимать, что настоящие способы в предпочтительных вариантах осуществления могут быть выполнены с семенами. Однако в некоторых вариантах

осуществления настоящие способы могут быть выполнены с другими способными к регенерации частями родительского растения, отличными от семян, например, могут быть использованы зародыши родительского растения. В вариантах осуществления по изобретению, в которых растение представляет собой злаковое растение, семена представляют собой зерна указанного злакового растения.

Представляющие интерес мутантные растения могут быть идентифицированы, как описано, например, в примерах 3-5 ниже, с использованием подходящих реагентов для ПЦР. Реагенты для ПЦР, как правило, содержат по меньшей мере нуклеотиды и полимеразу нуклеиновых кислот. Кроме того, реагенты для ПЦР предпочтительно могут также содержать один или несколько зондов для обнаружения, например, зонд для обнаружения мутаций и/или эталонный зонд для обнаружения, как описано в настоящем документе выше в разделе «Обнаружение продукта(ов) с помощью ПЦР».

Альтернативно, представляющие интерес мутанты могут быть идентифицированы путем секвенирования всего генома подвергнутых мутагенезу растений или их способных к регенерации частей, как известно в данной области. Таким образом, плотность мутаций может быть также определена путем секвенирования по меньшей мере всех открытых рамок считывания некоторых или всех подвергнутых мутагенезу растений или их способных к регенерации частей и подсчета количества полученных мутантов. Таким образом, зная размер генома, можно также рассчитать плотность мутаций.

В дополнение к описанным выше стадиям способы по изобретению могут включать одну или несколько дополнительных стадий. Способы могут, например, включать стадию подготовки указанного пула способных к регенерации частей, таких как семена, т.е. путем мутагенеза. Способы подготовки пула организмов описаны в настоящем документе ниже в разделе «Пул».

Способы также могут включать стадию репродукции одного или нескольких из указанных растений или способных к регенерации частей в пуле или в подпуле растений или способных к регенерации частей. Эта стадия может включать культивирование указанных растений или способных к регенерации частей посредством одного или нескольких циклов.

В дальнейшем, часто делается ссылка только на «семена», «культура» или «растение». Однако эти же соображения применимы к способам, в которых используются другие способные к регенерации части культур или растений. Каждая стадия культивирования может привести к потомству, которое не идентично

исходному растению, семени или другой способной к регенерации части. Например, после случайного мутагенеза полиплоидных растений большинство растений будут нести любую случайную мутацию только на одном аллеле и, таким образом, будут генетически гетерозиготными по отношению к этой мутации. Для самоопыляющихся растений, как правило, организмы-потомки включают: (1) организмы без мутации, (2) организмы-потомки, генетически гетерозиготные по мутации, и (3) организмы-потомки, генетически гомозиготные по мутации. Таким образом, пул потомства исходного пула или растений, или способных к регенерации частей, таких как семена, не обязательно будет идентичен исходному пулу, но в целом будет по меньшей мере представлять любую мутацию в NOI(s), присутствующем в исходном пуле – в виде гетерозиготных и/или гомозиготных организмов. Таким образом, по меньшей мере часть указанного потомства будет содержать мутантный аллель. Точно так же потомство подпула может не быть идентичным исходному подпулу, но будет, как правило, по меньшей мере представлять любую мутацию в NOI(s), присутствующем в исходном подпуле, в форме гетерозигот или гомозигот.

Стадия е) разделения пула способных к регенерации частей на еще несколько подпулов может, таким образом, включать стадию репродукции растений или их способных к регенерации частей. Это может быть, например, выполнено одновременно с разделением пула на подпулы.

В качестве альтернативы это может быть сделано после разделения пула на подпулы. Таким образом, способы по изобретению могут после стадии е) включать стадию репродукции растений или их способных к регенерации частей в пределах указанных подпулов.

Точно так же способы по изобретению могут после стадии f) включать стадию репродукции растений или их регенеративных частей, таких как семена, в пределах указанных фракций подпулов. Однако в некоторых вариантах осуществления изобретения и, в частности, в вариантах осуществления по изобретению, где растение представляет собой злаковое растение, может быть предпочтительным, чтобы способы не включали стадию репродукции растений или их способных к регенерации частей между стадиями f) и g).

В некоторых вариантах осуществления изобретения способы включают стадию репродукции растений или их способных к регенерации частей, таких как семена, содержащихся в подпуле или его фракции, содержащих мутацию в NOI. Указанная

стадия может быть выполнена в любое полезное время, т.е. между стадиями h) и i), как указано выше.

Способы также могут включать стадию идентификации группы подпулов, содержащих мутацию в NOI. Указанная стадия может быть выполнена в любое полезное время, но часто выполняется после стадии c) описанных выше способов – и может, например, выполняться, как описано ниже в разделе «Суперпулы».

Представляющий интерес нуклеотид (NOI)

Изобретение относится к способам идентификации одного растения или нескольких растений, несущих мутацию(и) в одном или нескольких представляющих интерес нуклеотидах. В частности, способы позволяют идентифицировать растение(я), несущее одну или несколько заданных мутаций в NOI(s). NOI(s) могут находиться внутри кодирующих областей или вне кодирующих областей. Соответственно, способ позволяет идентифицировать растение, несущее конкретную мутацию, при этом все еще опираясь на обычные методы селекции. Мутация предпочтительно находится в заданной целевой последовательности.

Мутация может представлять собой любую мутацию, где указанная одна или несколько представляющих интерес NOI(s) отличаются от соответствующих NOI(s) в эталонной последовательности. Часто эталонная последовательность представляет собой последовательность дикого типа. Однако эталонная последовательность также может представлять собой любую другую последовательность. Следовательно, эталонная последовательность может уже содержать некоторые мутации по сравнению с последовательностью дикого типа, где указанные мутации могли возникнуть естественным образом или быть индуцированными, например, путем мутагенеза.

Мутация может представлять собой мутацию любого типа, например, делецию, вставку, замену или смесь вышеупомянутого. Мутация может представлять собой синонимичную мутацию или несинонимичную мутацию. Например, мутация может представлять собой (1) мутацию, приводящую к возникновению преждевременного стоп-кодона, (2) мутацию, приводящую к сдвигу рамки считывания, или (3) мутацию, приводящую к замене аминокислоты (аминокислот) в транслируемом белке.

NOI(s) могут состоять из одного нуклеотида или нескольких нуклеотидов и, таким образом, NOI(s) могут состоять по меньшей мере из одного, такого как 1, например 2, такого как 3, например 4, такого как 5, например 6, такого 7, например 8,

такого как 9, например 10, такого как от 10 до 20, например, от 20 до 50, такого как более 50 нуклеотидов.

В предпочтительных вариантах осуществления изобретения NOI состоит из одного нуклеотида, и в этом случае мутация, например, может представлять собой замену одного нуклеотида. Такие мутации также известны как точечные мутации.

В других вариантах осуществления изобретения мутация может представлять собой делецию указанного NOI. В других вариантах осуществления мутация может представлять собой вставку одного или нескольких нуклеотидов между двумя представляющими интерес нуклеотидами.

В одном варианте осуществления эталонная последовательность представляет собой последовательность дикого типа, то есть наиболее часто встречающуюся в природе последовательность, и мутация представляет собой, таким образом, мутацию по сравнению с указанной последовательностью дикого типа. В других вариантах осуществления эталонная последовательность уже содержит мутации, и мутация представляет собой дополнительную мутацию по сравнению с указанной эталонной последовательностью.

В одном варианте осуществления мутация может быть ассоциирована с желательным признаком в пределах указанного вида. В зависимости от типа вида, желаемый признак может быть выбран из множества различных признаков. В вариантах осуществления изобретения, где вид представляет собой домашнее растение, указанным признаком может быть, например, повышенная жизнеспособность, повышенная устойчивость к различным факторам окружающей среды, усиленный рост, повышенное качество зерна или более высокая урожайность. В вариантах осуществления изобретения, где вид представляет собой растение, используемое для производства пищевых продуктов, кормов или напитков, признак может также относиться к повышенной питательной ценности, улучшенным вкусовым качествам, улучшенным характеристикам хранения или повышенной полезности для производства указанных пищевых продуктов, кормов или напитков.

NOI может располагаться в любой целевой последовательности в любой нуклеиновой кислоте. Часто целевая последовательность является частью последовательности геномной ДНК (gDNA). Более предпочтительно, целевая последовательность является частью последовательности gDNA. Таким образом, мутация может представлять собой мутацию gDNA указанного организма. NOI может располагаться в любой части gDNA, как в кодирующей, так и не кодирующей областях.

Часто NOI может располагаться в любой части гена, т.е. в пределах кодирующей области (например, в пределах экзона), в интроне или в регуляторных областях гена, например, в промоторах, терминаторах и/или интронах. Предпочтительно мутация представляет собой кодирующую или несинонимичную мутацию.

Вместо геномной ДНК для идентификации желаемой мутации также можно использовать cDNA, полученную из образца мРНК.

Растение

В предпочтительных вариантах осуществления изобретения растение может представлять собой зеленое растение, например, растение, выбранное из группы, состоящей из цветковых растений, хвойных, голосеменных, папоротников, плауновых, роголистников, печеночников, мхов и зеленых водорослей. Растение может быть, например, однодольным или двудольным.

В частности, растение может быть одомашненным растением. Указанное одомашненное растение может представлять собой любое растение, культивируемое людьми, например, в качестве источника пищи, корма или сырья для производства товаров или в эстетических целях.

В одном предпочтительном варианте осуществления изобретения вид представляет собой злаковое растение. «Злаковое растение», как определено в настоящем документе, является представителем семейства растений Graminae, культивируемыми в первую очередь благодаря их крахмалосодержащим семенам или зернам. Злаковые растения включают, но без ограничения, ячмень (*Hordeum*), пшеницу (*Triticum*), рис (*Oryza*), кукурузу (*Zea*), рожь (*Secale*), овес (*Avena*), сорго (*Sorghum*) и пшенично-ржаной гибрид *Triticale*.

Растение также может представлять собой другие одомашненные растения, включая томат.

Родительское растение может представлять собой сельскохозяйственную культуру или цветковое растение, такое как растение, принадлежащее семейству *Brassicaceae*, семейству *Fabaceae*, семейству *Poaceae*, семейству *Amaranthaceae*, семейству *Asteraceae*, семейству *Solanaceae*, семейству *Rubiaceae* или семейству *Orchidaceae*. В конкретных вариантах осуществления растение представляет собой ячмень. Настоящие способы могут быть особенно выгодны для растений, которые не легко поддаются генетическому редактированию, т.е. путем трансформации.

Родительским растением также может быть дерево или кустарник. Термин «дерево» используется в настоящем документе в самом широком смысле и относится к многолетнему растению с удлиненным стеблем или стволом, поддерживающим ветви и листья у большинства видов. Таким образом, этот термин охватывает такие растения, как банановые деревья или кокосовые пальмы. Кустарники представляют собой многолетние древесные растения небольшого и среднего размера. Настоящие способы выгодны для идентификации мутантов деревьев или кустарников. Деревья и кустарники имеют медленный рост, поэтому особенно выгодно не проводить обратное скрещивание для выявления представляющих интерес мутантов, поскольку это приводит к значительному сокращению времени, необходимого для идентификации таких мутантов. В целом, настоящие способы особенно выгодны для любого растения, имеющего длительный цикл репродукции.

Пул

Способы по изобретению включают обеспечение способных к регенерации частей N_p данного растения и подвергание мутагенезу указанных способных к регенерации частей указанного растения с получением пула способных к регенерации частей, представляющих множество генотипов. Указанное растение может, например, представлять собой любое из растений, описанных выше в разделе «Растение». Таким образом, пул или способные к регенерации части относятся к подвергнутым мутагенезу способным к регенерации частям, которые получены из способных к регенерации частей, обеспеченных на стадии а) настоящих способов. Следовательно, пул способных к регенерации частей состоит из того же количества способных к регенерации частей, что и количество способных к регенерации частей, обеспеченных на стадии а). Например, если на стадии а) обеспечено 30000 способных к регенерации частей, пул (подвергнутых мутагенезу) способных к регенерации частей насчитывает 30000 подвергнутых мутагенезу способных к регенерации частей.

Таким образом, пул способных к регенерации частей содержит множество способных к регенерации частей, все из которых принадлежат одному и тому же виду, но представляют разные генотипы указанного растения. Пул может содержать более чем одну способную к регенерации часть каждого генотипа. Однако пул должен содержать множество способных к регенерации частей различных генотипов.

Как описано в примерах 6 и 7 ниже, для получения определенной вероятности успеха предпочтительно использовать определенный оптимальный размер библиотеки

(OLS). Термин «оптимальный» в данном контексте означает оптимальный размер библиотеки для получения желаемой вероятности успеха. Размер библиотеки представляет собой общее количество различных генотипов, представленных в способных к регенерации частях поколения M1 или M(1+x). OLS зависит от частоты мутаций и количества подвергнутых скринингу мутаций. В предпочтительных вариантах осуществления OLS определяют, как описано в примере 6 ниже.

На основе OLS можно определить оптимальное количество способных к регенерации частей, подлежащих мутагенезу (N_p). Оптимальное количество N_p зависит от количества клеток зародышевой линии конкретного растения, а также от уровня сбора после мутагенеза.

В частности, оптимальное количество N_p (ON_p) может быть определено, как описано в примере 7 ниже.

В предпочтительных вариантах осуществления N_p находится в диапазоне от $0,7 * ON_p$ до $1,3 * ON_p$, где

$$ON_p = \frac{\frac{OLS \times 100}{Hr}}{GECN} ; \text{ и}$$

где Hr представляет собой ожидаемый средний уровень сбора указанных видов растений после мутагенеза в %; и

OLS представляет собой оптимальный размер библиотеки, где

$$OLS = \frac{\log \left(1 - \frac{PS}{100} \right)}{\log \left(1 - (1 - (1 - (Mf \times n))) \right)}$$

где PS представляет собой вероятность успеха в %; и

Mf представляет собой частоту мутаций; и

n представляет собой количество подвергнутых скринингу мутаций; и

где стадия b) включает стадию случайного мутагенеза, приводящую к частоте мутагенеза Mf.

Для некоторых видов растений, например для злаковых растений, уровень сбора является приблизительно таким же, как и скорость прорастания, и для таких видов растений для расчета ON_p вместо уровня сбора можно использовать скорость прорастания Gr.

Предпочтительно указанная частота мутаций соответствует низкой скорости мутагенеза. Скорость мутагенеза в виде % генов, свободных от несинонимичных мутаций, можно рассчитать на основе частоты мутаций, количества генов и средней длины генов в конкретном растении, используя следующую формулу:

$$GFoNSM = Gn \times Gl \times Mf \times \frac{Nsnm}{100}$$

где *GFoNSM* представляет собой количество генов, свободных от несинонимичных мутаций, в %; и *Gn* представляет собой количество генов в конкретном растении, *Gl* представляет собой среднюю длину гена в bp (пара оснований) (Tiessen et al 2012), *Mf* представляет собой частоту мутаций, *Nsnm* представляет собой среднее количество несинонимичных мутаций, вызванных мутагенезом, в %.

Обычно *Nsnm* находится в диапазоне 60-70%, например, 66%. Если точное значение *Nsnm* неизвестно, можно использовать значение *Nsnm*, равное 66%.

В предшествующем уровне техники скорость мутагенеза часто описывают с точки зрения частоты мутаций. Например, иногда описывается частота мутаций 1 на 500000 у растений *Arabidopsis thaliana* поколения M2. Это соответствует приблизительно 57 генам с кодирующими мутациями у растений *Arabidopsis thaliana* поколения M1 или до 99,79% всех генов, свободных от несинонимичных мутаций у растений *Arabidopsis* поколения M1. Кроме того, иногда описывается частота мутаций 1 на 10000 у растений *Arabidopsis thaliana* поколения M2. Это соответствует приблизительно 2862 генам с кодирующими мутациями у растений *Arabidopsis thaliana* поколения M1 или до 89,65% всех генов, свободных от несинонимичных мутаций у растений *Arabidopsis* поколения M1.

В одном варианте осуществления *Np* находится в диапазоне от $0,9 * ONp$ до $1,1 * ONp$, предпочтительно *Np* находится в диапазоне от $0,9 * ONp$ до $1,1 * ONp$.

Иногда фактическое значение одного или нескольких параметров, необходимых для расчета *ONp*, неизвестно. В таких случаях можно использовать ожидаемое значение, а не фактическое значение. Таким образом, если уровень сбора после мутагенеза неизвестен, можно использовать ожидаемый уровень сбора. Ожидаемый уровень сбора может быть основан на знаниях специалиста в данной области, полученных в ходе аналогичных процедур мутагенеза. Аналогично, если точная частота мутаций неизвестна, можно использовать ожидаемую частоту мутаций. Ожидаемая частота мутаций может быть основана на знаниях специалиста в данной области, полученных в ходе аналогичных процессов мутагенеза.

Лицо, применяющее способы по изобретению, может выбрать подходящую вероятность успеха. Обычно вероятность успеха составляет по меньшей мере 85%, предпочтительно по меньшей мере 90%, более предпочтительно по меньшей мере 95%, такую как по меньшей мере 98%.

например, от 100000 до 500000 способных к регенерации частей, или от 500000 до 1000000 способных к регенерации частей, или от 600000 до 900000 способных к регенерации частей, или от 900000 до 1500000 способных к регенерации частей. В некоторых вариантах осуществления способные к регенерации части представляют собой семена.

Предпочтительно, чтобы пул организмов содержал достаточное количество способных к регенерации частей растения, с различными генотипами, так что пул теоретически может содержать все возможные мутации во всех генах указанного растения. Например, в вариантах осуществления изобретения, в которых видом является ячмень, полагают, что пул, содержащий ~1500000 подвергнутых случайному мутагенезу зерен ячменя, теоретически будет содержать все возможные мутации во всех генах указанного ячменя при плотности мутаций 1 мутация на 1500 kbp (т.п.н.) геномной ДНК. Для повышения эффективности способов по изобретению пул может содержать по меньшей мере 2×, например, по меньшей мере 3× способных к регенерации частей растения с различными генотипами, которые, как теоретически предполагается, содержат все возможные мутации. Таким образом, пул может содержать по меньшей мере 5000 способных к регенерации частей, например, по меньшей мере 10000 способных к регенерации частей, например, по меньшей мере 15000 способных к регенерации частей, например, по меньшей мере 20000 способных к регенерации частей, например, по меньшей мере 25000 способных к регенерации частей, например, по меньшей мере 30000 способных к регенерации частей, например, по меньшей мере 40000 способных к регенерации частей, например, по меньшей мере 50000 способных к регенерации частей, например, по меньшей мере 60000 способных к регенерации частей, например, по меньшей мере 70000 способных к регенерации частей, например, по меньшей мере 80000 способных к регенерации частей, например, по меньшей мере 90000 способных к регенерации частей, например, по меньшей мере 100000 способных к регенерации частей, например, по меньшей мере 200000 способных к регенерации частей, например, по меньшей мере 300000 способных к регенерации частей, например, по меньшей мере 400000 способных к регенерации частей, например, по меньшей мере 500000 способных к регенерации частей, например, по меньшей мере 600000 способных к регенерации частей, например, по меньшей мере 700000 способных к регенерации частей, например, по меньшей мере 800000 способных к регенерации частей, например, по меньшей мере 900000 способных к

регенерации частей, например, по меньшей мере 10^6 способных к регенерации частей, например, по меньшей мере $2 \cdot 10^6$ способных к регенерации частей, например, по меньшей мере $3 \cdot 10^6$ способных к регенерации частей, например, по меньшей мере $4 \cdot 10^6$ способных к регенерации частей, например, по меньшей мере $5 \cdot 10^6$ способных к регенерации частей, например, по меньшей мере $6 \cdot 10^6$ способных к регенерации частей, например, по меньшей мере $7 \cdot 10^6$ способных к регенерации частей, например, по меньшей мере $8 \cdot 10^6$ способных к регенерации частей, например, по меньшей мере $9 \cdot 10^6$ способных к регенерации частей, например, по меньшей мере 10^7 способных к регенерации частей или более с разными генотипами. В некоторых вариантах осуществления пул содержит от 5000 до 100000 способных к регенерации частей, например, от 10000 до 90000 способных к регенерации частей, например, от 15000 до 85000 способных к регенерации частей, например, от 20000 до 80000 способных к регенерации частей, например, от 25000 до 90000 способных к регенерации частей, например, от 30000 до 100000 способных к регенерации частей, например, от 30000 до 75000 способных к регенерации частей, например, от 50000 до 75000 способных к регенерации частей, например, от 50000 до 1000000 способных к регенерации частей, например, от 100000 до 500000 способных к регенерации частей, или от 500000 до 1000000 способных к регенерации частей, или от 600000 до 900000 способных к регенерации частей, или от 900000 до 1500000 способных к регенерации частей. В некоторых вариантах осуществления способные к регенерации части представляют собой семена. Это может, например, относиться к вариантам осуществления, в которых рассматриваемым видом является ячмень.

В некоторых вариантах осуществления пул способных к регенерации частей может содержать по меньшей мере 5000000, например, в диапазоне от 1000000 до 100000000 способных к регенерации частей растения с разными генотипами. В некоторых вариантах осуществления пул способных к регенерации частей представляет по меньшей мере 5000 различных генотипов, например, по меньшей мере 10000 различных генотипов, например, по меньшей мере 15000 различных генотипов, например, по меньшей мере 20000 различных генотипов, например, по меньшей мере 25000 различных генотипов, например, по меньшей мере 30000 различных генотипов, например, по меньшей мере 40000 различных генотипов, например, по меньшей мере 50000 различных генотипов, например, по меньшей мере 60000 различных генотипов, например, по меньшей мере 70000 различных генотипов, например, по меньшей мере 80000 различных генотипов, например, по меньшей мере 90000 различных генотипов,

например, по меньшей мере 100000 различных генотипов, например, по меньшей мере 200000 различных генотипов, например, по меньшей мере 300000 различных генотипов, например, по меньшей мере 400000 различных генотипов, например, по меньшей мере 500000 различных генотипов, например, по меньшей мере 600000 различных генотипов, например, по меньшей мере 700000 различных генотипов, например, по меньшей мере 800000 различных генотипов, например, по меньшей мере 900000 различных генотипов, например, по меньшей мере 10^6 различных генотипов, например, по меньшей мере $2 \cdot 10^6$ различных генотипов, например, по меньшей мере $3 \cdot 10^6$ различных генотипов, например, по меньшей мере $4 \cdot 10^6$ различных генотипов, например, по меньшей мере $5 \cdot 10^6$ различных генотипов, например, по меньшей мере $6 \cdot 10^6$ различных генотипов, например, по меньшей мере $7 \cdot 10^6$ различных генотипов, например, по меньшей мере $8 \cdot 10^6$ различных генотипов, например, по меньшей мере $9 \cdot 10^6$ различных генотипов, например, по меньшей мере 10^7 различных генотипов или более. В некоторых вариантах осуществления пул представляет от 5000 до 100000 различных генотипов, например, от 10000 до 90000 различных генотипов, например, от 15000 до 85000 различных генотипов, например, от 20000 до 80000 различных генотипов, например, от 25000 до 90000 различных генотипов, например, от 30000 до 100000, например, от 30000 до 75000 различных генотипов, например, от 50000 до 75000 различных генотипов, например, от 50000 до 1000000 различных генотипов, например, от 100000 до 500000 различных генотипов, или, например, от 500000 до 1000000 различных генотипов, или от 600000 до 900000 различных генотипов, или от 900000 до 1500000 различных генотипов.

Одним из преимуществ способов по настоящему изобретению является то, что способы позволяют проводить скрининг большого количества способных к регенерации частей растения (таких как семена) разных генотипов. Таким образом, пул может содержать очень большое количество способных к регенерации частей разных генотипов. Кроме того, способы по изобретению могут позволить идентифицировать растение с заданной мутацией в любом NOI. Чтобы иметь возможность идентифицировать растение с мутацией в любом NOI, может потребоваться, чтобы пул содержал большое количество различных генотипов, таких как вышеупомянутое количество различных генотипов.

В предпочтительном варианте осуществления изобретения пул способных к регенерации частей растения (таких как семена) получают путем мутагенеза, в частности, путем случайного мутагенеза. Таким образом, множество способных к

регенерации частей может быть подвергнуто мутагенезу, чтобы получить пул способных к регенерации частей. Указанный мутагенез может, в частности, представлять собой случайный мутагенез, который, например, может быть осуществлен, как описано ниже.

Может оказаться достаточным подвергнуть способную к регенерации часть указанного многоклеточного организма указанному случайному мутагенезу. В таких вариантах осуществления случайному мутагенезу подвергают либо множество способных к регенерации частей растения, такое как множество семян растения, либо их смесь.

Пул может содержать множество способных к регенерации частей, которые представляют множество генотипов. Таким образом, пул может включать множество способных к регенерации частей, которые были подвергнуты мутагенезу. Однако пул может также включать потомство способных к регенерации частей, подвергшихся мутагенезу.

В настоящем документе способные к регенерации части (например, семена, такие как зерна злаков), которые были подвергнуты мутагенезу, могут называться как поколение M0. Указанные способные к регенерации части, например, семена, такие как зерна злаков, можно высевать и дать им развиваться в зрелые растения, способные к регенерации части (например, семена, такие как зерна злаков) которых считаются поколением M1. Способные к регенерации части поколения M1, например, семена, можно посеять и дать им вырасти в зрелые растения, способные к регенерации части (например, семена) которых считаются поколением M2, и так далее. Растения сорта, использованного для создания библиотеки, которые еще не подвергались мутагенезу, также могут называться «материнскими растениями». Это показано на фигуре 2.

Пул способных к регенерации частей может включать способные к регенерации части, такие как семена, например, зерна злаков, любого из вышеперечисленных поколений. Пул способных к регенерации частей может также включать способные к регенерации части, например семена, такие как зерна злаков поколения M1, M2 или M3. Предпочтительно, чтобы пул способных к регенерации частей содержал способные к регенерации части поколения M1, предпочтительно, чтобы все способные к регенерации части в указанном пуле способных к регенерации частей относились к поколению M1.

Случайный мутагенез

Традиционно, эффективный протокол мутагенеза считается важнейшим, основным требованием для успешной программы селекции.

Однако в настоящем изобретении обеспечены способы, включающие использование «умеренного» мутагенеза, приводящего к генерации в диапазоне от 1 до 30 несинонимичных мутаций в кодирующих областях на одно растение.

Случайный мутагенез в настоящее время основан на использовании химического мутагена или облучения. Исторически, эффективность мутагена измеряли по биологическим эффектам. Однако желательно установить связь между наблюдаемым биологическим эффектом и четко определенной и легко измеримой физической величиной, характеризующей количество облучения или химического мутагена, придающего эффект.

Случайный мутагенез может быть проведен любым подходящим способом, например, путем облучения или химической обработки. Облучение может представлять собой УФ-облучение, рентгеновское облучение или радиоактивное облучение. Химический мутагенез может включать обработку любым мутагенизирующим химическим веществом, например, химическим веществом, выбранным из группы, состоящей из азидов натрия (NaN_3), алкилирующих агентов, таких как N-этил-N-нитрозомочевина (ENU), метилнитрозогуанидин (MNNG) и этилметансульфонат (EMS), или алкилирующих агентов, упомянутые ниже. NaN_3 , ENU и EMS часто используются для создания случайных мутантов.

Для индукции случайных мутаций в ДНК растения, в частности, в геномной ДНК, ядра или регенерируемые вегетативные ткани растений могут быть обработаны мутагеном или смесью мутагенов, включая, но без ограничения, алкилирующие агенты, такие как: сульфонаты, например, этилметансульфонат (EMS), диэтилсульфонат (DES); сернистые иприты, например, этил-2-хлорэтилсульфид; азотистые иприты, например, 2-хлорэтилдиметиламин и; эпоксиды, например, этиленоксид. Другие включают этиленимин, гидроксиламин (NH_2OH), N-метил-N'-нитро-N-нитрозогуанидин (MNNG), NaN_3 и диазометан. Затем обработанные ткани или ядра могут быть размножены для получения организмов потомства. Такой случайный мутагенез можно использовать с растениями, такими как сельскохозяйственные растения, или с любыми растениями, описанными в разделе «Растения».

Способы по изобретению не ограничиваются каким-либо конкретным типом мутагенеза при условии, что это «умеренный» или «низкий» мутагенез. Указанный низкий уровень мутагенеза может представлять собой мутагенез, приводящий к

скорости мутагенеза от 1 до 30 несинонимичных мутаций в кодирующих областях на растение. В некоторых вариантах осуществления это соответствует представляющим интерес мутациям, происходящим с частотой, соответствующей 1 мутации на 1,5 основания геномной ДНК. В некоторых вариантах осуществления низкий уровень мутагенеза у диплоидных растений может означать, что более 99,8%, предпочтительно более 99,9% всех генов свободны от несинонимичных мутаций в M1, и низкий уровень мутагенеза в растениях с уровнем плоидности выше двух, может означать, что более 98,0%, предпочтительно более 99,0% всех генов свободны от несинонимичных мутаций в M1.

Скорости мутаций известны для многих типов мутагенеза, и специалист в данной области легко сможет рассчитать оптимальную дозировку, концентрацию и/или время воздействия мутагена для каждого типа мутагенеза, чтобы получить желаемую плотность мутаций или частоту мутаций. Альтернативно, специалист в данной области может протестировать мутагенез для определения оптимальной дозы, концентрации и/или времени воздействия мутагена, чтобы получить желаемую низкую скорость мутагенеза для видов растений. Настоящие способы основаны на проведении мутагенеза любым из описанных в настоящем документе способов для получения низкой скорости мутагенеза. Упомянутая низкая скорость мутагенеза может, в частности, представлять собой скорость мутагенеза, приводящую к возникновению от 1 до 30 несинонимичных мутаций в кодирующих областях на одно растение. Альтернативно, низкая скорость мутагенеза может быть такой, что более 99,8%, например, 99,9% всех генов свободны от несинонимичных мутаций в диплоидных растениях в M1, и низкая скорость мутагенеза может быть такой, что по меньшей мере 98,0%, предпочтительно по меньшей мере 99,0% всех генов свободны от несинонимичных мутаций у растений с уровнем плоидности выше, чем два в M1.

Например, библиотеку, содержащую растения заданного вида, имеющие заданное количество генов, подвергают случайному мутагенезу в определенных условиях, например, при конкретной концентрации мутагена в течение определенного периода времени, как проиллюстрировано в примерах 3-5. Мутанты идентифицируют, например, в данной области данной длины, как описано в настоящем документе. Количество мутантов, идентифицированных в этой области, можно использовать для расчета плотности мутаций для заданных условий мутагенеза (концентрация и время). Например, если фрагмент размером 1000 bp анализируют для 10000 особей и среди проанализированных пар оснований идентифицируют всего 10 мутантов, частота

мутаций в этих условиях составляет $10/(1000 \times 10000) = 1.10^{-6}$ мутаций на бр. Другими словами, плотность мутаций в этих условиях можно считать равной количеству идентифицированных мутантов, деленному на произведение количества подвергнутых скринингу элементов (способных к регенерации частей) на длину подвергнутой скринингу области, как показано ниже:

Плотность мутаций = (количество идентифицированных мутантов) / ((количество мутагенизированных особей) x (длина анализируемой области)).

В некоторых вариантах осуществления мутагенез приводит к 1 несинонимичной мутации в кодирующих областях на растение. В других вариантах осуществления мутагенез приводит к 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 или 30 несинонимичным мутациям в кодирующих областях на одно растение. В дополнение к несинонимичным мутациям в кодирующих областях, растение может дополнительно содержать молчащие мутации, например, в некодирующих областях.

В предпочтительных вариантах осуществления стадия мутагенеза является такой, что скорость мутагенеза составляет не более 1 несинонимичной мутации на каждые 1000 генов, например, не более 1 несинонимичной мутации на каждые 5000 генов, например, не более 1 несинонимичной мутации на каждые 10000 генов, например, не более 1 несинонимичной мутации на каждые 20000 генов, например, не более 1 несинонимичной мутации на каждые 30000 генов, например не более 1 несинонимичной мутации на каждые 5000 генов. 1 несинонимичная мутация на каждые 1000 генов означает, что если растение имеет 6000 генов, ожидается, что 6 генов будут нести несинонимичные мутации, в то время как ожидается, что 5994 гена будут лишены несинонимичных мутаций в кодирующих областях, но все же могут содержать другие мутации, например, синонимичные мутации вне кодирующих областей. В других предпочтительных вариантах осуществления стадия мутагенеза является такой, что скорость мутагенеза такова, что по меньшей мере 99,8%, предпочтительно по меньшей мере 99,9% всех генов свободны от несинонимичных мутаций в диплоидных растениях в M1, и низкая скорость мутагенеза может быть такова, что по меньшей мере 98,0%, предпочтительно по меньшей мере 99,0% всех генов свободны от несинонимичных мутаций у растений с более высокой плоидностью, чем две в M1.

В некоторых вариантах осуществления стадия мутагенеза приводит не более чем к 30 несинонимичным мутациям в кодирующих областях, например, не более чем

к 25 несинонимичным мутациям в кодирующих областях, например, не более чем к 20 несинонимичным мутациям в кодирующих областях, например, не более чем к 15 несинонимичным мутациям в кодирующих областях, например, не более чем к 10 несинонимичным мутациям в кодирующих областях, например, не более чем к 5 несинонимичным мутациям в кодирующих областях, например, не более чем к 4 несинонимичным мутациям в кодирующих областях, например, не более чем к 3 несинонимичным мутациям в кодирующих областях, например, не более чем к 2 несинонимичным мутациям в кодирующих областях, например, 1 не молчащей мутации в кодирующей области у растений поколения M1 или M(1+X).

Способы, включающие мутагенез, часто выполняют в таких условиях, что значительная доля (до 50%) мутагенизированных организмов оказывается нежизнеспособной и погибает. В противоположность этому, поскольку предлагаемые способы можно выполнять на большом количестве растений или их способных к регенерации частях, предпочтительно проводить случайный мутагенез в таких условиях, при которых выживает большинство растений. Предпочтительно мутагенез осуществляют в таких условиях, что 90% способных к регенерации частей выдерживают стадию мутагенеза, например, по меньшей мере 91% способных к регенерации частей, например, по меньшей мере 92% способных к регенерации частей, например, по меньшей мере 93% способных к регенерации частей, например, по меньшей мере 94% способных к регенерации частей, например, по меньшей мере 95% способных к регенерации частей, например, по меньшей мере 96% способных к регенерации частей, например, по меньшей мере 97% способных к регенерации частей, например, по меньшей мере 98% способных к регенерации частей или по меньшей мере 99% способных к регенерации частей выдерживают стадию мутагенеза.

Для большинства культур предпочтительным мутагеном часто является этилметансульфонат (EMS). Однако мутагенез ячменя обычно проводят с использованием обработки путем мутагенеза на основе азиды натрия (NaN_3).

Обработка ячменя азидом натрия (1,5 mM в течение 2 часов) обычно приводит к плотности мутаций 1 мутация на каждые 500000 bp (Szurman-Zubrzycka et al., 2018). Это число обычно определяется путем анализа данных секвенирования различных генов в популяции. При плотности мутаций 1:500000 bp и генома ячменя 5,3 GB каждое растение ячменя будет нести приблизительно 10000 мутаций. Кодирующая последовательность ячменя составляет приблизительно 65 Mb, что соответствует 1-2%

всего генома. На основании этого можно предположить, что приблизительно 100-200 из 10000 индуцированных мутаций находятся в кодирующих областях.

Приблизительно одна треть вызванных азидом натрия мутаций в кодирующей области приводит к молчащей нуклеотидной замене. Таким образом, если мутагенез проводят с использованием азида натрия, Nsnm обычно составляет 66%. Поэтому можно предположить, что 50-100 мутаций из 10000 индуцированных мутаций будут вызывать несинонимичные мутации в кодирующей области. Поскольку мутации происходят случайным образом в геноме, можно предположить, что 50-100 аминокислотных замен в кодирующих областях находятся в разных кодирующих областях или генах генома. Плотность мутаций зависит от культуры и обработки. Для других культур плотность мутаций в типичной библиотеке TILLING может варьироваться от 1 на каждые bp до 1 на каждый 1 миллион bp общей геномной ДНК.

В некоторых вариантах осуществления мутаген представляет собой NaN_3 . В таких вариантах осуществления NaN_3 предпочтительно используют при концентрации менее 1 мМ или менее в течение 2 часов или менее, такой как 0,9 мМ или менее, такой как 0,8 мМ или менее, такой как 0,7 мМ или менее, такой как 0,6 мМ или менее, такой как 0,5 мМ или меньше, такой как 0,4 мМ или менее, такой как 0,3 мМ или менее. В одном варианте осуществления NaN_3 предпочтительно используют при концентрации приблизительно 0,3 мМ в течение периода времени от 1,5 до 2,5 часов.

Стадию случайного мутагенеза также можно проводить с EMS, предпочтительно при концентрации 4% или менее в течение 2 часов или менее, такой как 3% или менее, такой как 2% или менее, такой как 1,5% или менее, такой как 1% или менее, такой как 0,75% или менее, такой как 0,5% или менее, такой как 0,4% или менее, такой как 0,3% или менее, такой как 0,2% или менее, такой как 0,1% или менее, такой как 0,1%, в течение 24 часов или менее.

Случайный мутагенез также может быть осуществлен другими способами, известными в данной области, такими как случайный мутагенез CRISPR, в котором определенные массивы CRISPR можно использовать для случайного мутагенеза популяций растений (Lu et al., 2017). Например, массивы направляющих РНК CRISPR могут быть синтезированы и случайным образом трансформированы в растения, таким образом, создавая библиотеку растений, каждое из которых несет в геноме CRISPR-индуцированную мутацию.

Разделение пула способных к регенерации частей на подпулы

Способы по изобретению включают стадию разделения пула способных к регенерации частей растения, таких как семена, поколения $M1$ или $M(1+X)$ на один или несколько подпулов, где каждый подпул содержит множество способных к регенерации частей растения, таких как семена. Каждый подпул включает способные к регенерации части растения, представляющие множество генотипов, причем важно, чтобы каждый подпул содержал все способные к регенерации части от одного данного зрелого растения.

В некоторых вариантах осуществления, где способные к регенерации части относятся к поколению $M(1+X)$, может быть предпочтительным, чтобы все способные к регенерации части всего потомства данного растения поколения $M0$ помещались в один и тот же подпул.

На практике, например, когда способные к регенерации части представляют собой семена растений, выращенных в поле, после получения пула мутагенизированных семян (поколения $M0$) весь пул высевают в поле. Все поле представляет другой пул семян поколения $M1$, включающий все семена растений, полученных из пула семян поколения $M0$. Общий пул семян поколения $M1$ в настоящем документе называется «библиотекой». Размер библиотеки отличается от N_p в зависимости от количества GECN и уровня сбора. Библиотеку семян поколения $M1$ разделяют на подпулы, например, следующим образом: поле делят на разные участки, и семена с одного участка представляют один подпул. Эти семена в данном примере затем относят к поколению $M1$. Подпул семян затем делят на фракции, как подробно описано в настоящем документе в другом месте.

Если используют семена поколения $M2$ или более поздних поколений, растения для получения семян более поздних поколений можно в некоторых вариантах осуществления предпочтительно выращивать таким образом, чтобы можно было поместить все семена растений-потомков одного и того же растения $M0$ в один и тот же подпул.

Специалист сможет применить приведенный выше пример к растениям, выращиваемым в других условиях, например, в теплице, в отдельных горшках и т.д.

Как описано в настоящем документе ниже в примере 8, оптимальный размер подпула может быть определен исходя из чувствительности обнаружения используемых (чувствительных) средств обнаружения и количества GECN. Таким образом, в предпочтительных вариантах осуществления оптимальный размер подпула определяют, как описано в примере 8 ниже.

Таким образом, максимальный размер подпула (SPm) можно рассчитать следующим образом:

$$SPm = \frac{Y}{GECN}$$

где чувствительность обнаружения равна 1 на Y .

В одном варианте осуществления каждый подпул содержит способные к регенерации части от некоторого количества зрелых растений, при этом указанное количество находится в диапазоне от $0,5 * SPm$ до SPm , при этом все способные к регенерации части от одного данного зрелого растения помещают в один и тот же подпул.

В одном варианте осуществления каждый подпул включает способную к регенерации часть от зрелых растений в диапазоне от $0,7 * SPm$ до SPm .

Обычно пул делится на множество подпулов, предпочтительно по меньшей мере на 2 подпула, таких как по меньшей мере 50 подпулов, такое как по меньшей мере на 100 подпулов, предпочтительно по меньшей мере на 200 подпулов, более предпочтительно по меньшей мере на 300 подпулов, еще более предпочтительно по меньшей мере на 400 подпулов, еще более предпочтительно по меньшей мере на 500 подпулов, еще более предпочтительно по меньшей мере на 1000 подпулов, еще более предпочтительно по меньшей мере на 1500 подпулов.

В некоторых вариантах осуществления пул разделен по меньшей мере на 500, такое как по меньшей мере на 1000, например, по меньшей мере на 1500 подпулов. Это может, в частности, иметь место в вариантах осуществления изобретения, в которых пул включает большое количество организмов разных генотипов.

В принципе, не существует верхнего предела для количества подпулов. Тем не менее, обычно пул делится не более чем на 50000, например, не более чем на 25000, например, не более чем на 10000 подпулов.

В некоторых вариантах осуществления пул разделен на подпулы в диапазоне от 90 до 1000.

Каждый подпул предпочтительно включает множество способных к регенерации частей растения, представляющих множество генотипов. Предпочтительно каждый подпул включает по меньшей мере 10, более предпочтительно по меньшей мере 100, еще более предпочтительно по меньшей мере 200, еще более предпочтительно по меньшей мере 300, еще более предпочтительно по меньшей мере 400, еще более предпочтительно по меньшей мере 500, еще более

предпочтительно по меньшей мере 1000, еще более предпочтительно по меньшей мере 5000 способных к регенерации частей растения с различными генотипами. В некоторых вариантах осуществления каждый подпул содержит от 500 до 5000, такое как от 600 до 2000 способных к регенерации частей с разными генотипами.

В некоторых вариантах осуществления каждый подпул содержит от 1000 до 2000 способных к регенерации частей, представляющих различные генотипы.

Также предпочтительно, чтобы каждый подпул содержал более одной способной к регенерации части каждого генотипа. Это может быть достигнуто путем обеспечения того, чтобы все способные к регенерации части любого данного зрелого растения были помещены в один и тот же подпул. В частности, предпочтительно, чтобы каждый подпул содержал достаточное количество способных к регенерации частей каждого генотипа для того, чтобы произвольно разделить подпул на 2, 3 или 4 фракции таким образом, чтобы каждая фракция теоретически содержала способные к регенерации части, представляющие каждый генотип подпула. Таким образом, предпочтительно, чтобы каждый подпул содержал по меньшей мере 5, предпочтительно по меньшей мере 10, еще более предпочтительно по меньшей мере 15 способных к регенерации частей, представляющих каждый генотип. Это может, в частности, иметь место в вариантах осуществления изобретения, в которых вид представляет собой злаковое растение.

Иллюстрация способа деления пула на подпулы представлена на фигуре 3.

Как правило, подпул может быть приготовлен таким образом, чтобы все способные к регенерации части одного конкретного растения содержались в одном подпуле. Таким образом, в одном варианте осуществления изобретения способы включают следующие стадии:

- обеспечение множества способных к регенерации частей, таких как семена растения, например, зерна злакового растения;
- подвергание мутагенезу указанных семян, тем самым получая способные к регенерации семена, такие как семена поколения M0;
- выращивание указанных способных к регенерации частей, таких как семена поколения M0, в зрелые растения и получение способных к регенерации частей, таких как семена, от указанных зрелых растений, где указанные способные к регенерации части, такие как семена, представляют собой семена поколения M1;

- необязательное повторение предыдущей стадии X раз для получения растений, содержащих способные к регенерации части, такие как семена, поколения $M(1+X)$;

- получение способных к регенерации частей, таких как семена, поколения $M1$ или поколения $M(1+X)$ от указанных зрелых растений, тем самым получая пул способных к регенерации частей, таких как семена, (например, пул зерен злаковых растений);

- разделение указанного пула на подпулы, где все способные к регенерации части, такие как семена (например, зерна злаковых растений), от данного зрелого растения, помещают в один и тот же подпул.

Эти стадии проиллюстрированы на фигуре 2 с использованием в качестве примера ячменя. Способ может включать стадии, показанные на фигуре 2. Специалисту в данной области будет понятно, что аналогичные принципы можно применять и к другим растениям. Таким образом, стадии, показанные на фигуре 2, могут быть выполнены с использованием любого цветкового растения и, таким образом, не ограничиваются ячменем. Могут быть использованы другие протоколы мутагенеза, которые не ограничиваются мутацией с использованием азида натрия. Кроме того, стадии, показанные на фигуре 2, могут быть выполнены с использованием любого количества мутантных растений.

Соответственно, способ может включать следующие стадии:

- обеспечение множества способных к регенерации частей растения, таких как семена, например, зерна злаковых растений;

- подвергание мутагенезу указанных способных к регенерации частей, тем самым получая способные к регенерации части поколения $M0$;

- необязательно культивирование указанных способных к регенерации частей поколения $M0$ в зрелые растения и получение способных к регенерации частей, например, из указанных зрелых растений, где указанные способные к регенерации части представляют собой способные к регенерации части поколения $M1$;

- необязательное повторение предыдущей стадии X раз с получением растений, содержащих способные к регенерации части поколения $M(1+X)$;

- выращивание способных к регенерации частей, таких как семена поколения $M0$, $M1$ или $M(1+X)$, на дискретных полевых участках в зрелые растения, где способные к регенерации части указанных зрелых растений составляют пул способных к регенерации частей;

- разделение указанных участков на полевые подучастки;
- сбор всех способных к регенерации частей, таких как семена, от всех растений в пределах одного полевого подучастка, тем самым получая подпул способных к регенерации частей (например, подпул семян, таких как зерна злаковых растений), который можно назвать как «всего зерна».

Вместо выращивания способных к регенерации частей, таких как семена, на полевых участках, их можно выращивать любым полезным способом, позволяющим разделить растения на подгруппы, таким образом получая подпулы. Например, способные к регенерации части можно выращивать в отдельных контейнерах, каждый из которых содержит одно или несколько растений. Семена также можно выращивать в теплице.

Примеры способов разделения пула способных к регенерации частей на подпул обеспечены в РСТ/EP2017/065516, в частности, в разделе, озаглавленном «WS2: подготовка упорядоченной библиотеки зерен, полученных из мутантных злаковых растений».

Идентификация подпулов

Способы по изобретению включают стадии разделения пула способных к регенерации частей, таких как семена, на подпулы. Затем получают образцы ДНК из каждого подпула и используют для идентификации подпула(ов), содержащего ДНК, содержащую мутацию(и).

Предпочтительно образцы ДНК получают с помощью способа, включающего стадии (1) разделения каждого подпула на фракции; (2) получения образцов ДНК из всей фракции.

Предпочтительно фракции каждого подпула получают таким образом, чтобы по меньшей мере две фракции теоретически содержали способные к регенерации части растения всех генотипов, представленных в подпуле. Предпочтительно, чтобы все фракции теоретически содержали способные к регенерации части всех генотипов, представленных в подпуле. Это может быть достигнуто, например, путем смешивания всех способных к регенерации частей подпула, случайного разделения подпула на фракции, находящиеся в диапазоне от 2 до 10, например, на фракции, находящиеся в диапазоне от 2 до 6, например, на фракции, находящиеся в диапазоне от 2 до 4. Предпочтительно, чтобы указанные по меньшей мере 2 фракции, каждая, содержали в

пределах 10-50%, например, в пределах 15-50%, например, в пределах 20-50% способных к регенерации частей подпула.

Таким образом, подпул может быть разделен по меньшей мере на 2 фракции, например, по меньшей мере на 3 фракции, например, по меньшей мере на 4 фракции или более, причем каждая фракция теоретически содержит способные к регенерации части всех генотипов, представленных в подпуле.

i. Как описано в примере 10, оптимальный размер фракции можно рассчитать на основе среднего количества способных к регенерации частей на растение и количества GECN. Таким образом, в одном варианте осуществления предпочтительно, чтобы образец ДНК был получен из фракции способных к регенерации частей подпула следующим образом: производного деления каждого подпула на фракции, где первая фракция содержит от $\frac{GECN * 100}{RPr} \%$ до $\frac{GECN * 500}{RPr} \%$ способных к регенерации частей подпула; и где указанная первая фракция содержит не более 50% указанной способной к регенерации части;

ii. получение образцов ДНК всей первой фракции каждого подпула, где RPr представляет собой количество способных к регенерации частей на зрелое растение.

Анализ указанной фракции и, следовательно, идентификация подпула может включать следующие стадии:

- a) получение образцов ДНК, каждый из которых содержит ДНК каждого генотипа в пределах одной фракции, что, например, может быть выполнено, как описано в настоящем документе ниже в разделе «Получение образцов ДНК»;
- b) использование средства чувствительного обнаружения для обнаружения наличия мутации(й) в представляющем интерес NOI(s).

Стадия a) может, в частности, включать извлечение ДНК (или РНК) из всей фракции.

Стадия использования чувствительного средства обнаружения может включать выполнение множества ПЦР-амплификаций, каждая из которых включает множество разделенных на компартменты ПЦР-амплификаций, например, осуществляемых, как описано в настоящем документе ниже в разделе «ПЦР-амплификация, включающая множество разделенных на компартменты ПЦР-амплификаций»; и обнаружение продукта(ов) ПЦР-амплификации, содержащего мутацию(и) в представляющем интерес NOI(s), тем самым идентифицируя представляющую интерес фракцию(и) и,

следовательно, представляющий интерес подпул(ы), например, осуществляют, как описано ниже в разделе «Обнаружение продукта(ов) ПЦР-амплификации».

Подпул может быть, например, идентифицирован напрямую при условии, что продукт(ы) ПЦР-амплификации содержит мутацию(и) в NOI(s), которая может быть обнаружена напрямую после или во время ПЦР-амплификации. Это может быть сделано, например, если продукт(ы) ПЦР-амплификации содержит(ат) средства обнаружения, которые вызывают один или несколько обнаруживаемых сигналов, при условии, что продукт(ы) ПЦР-амплификации содержит(ат) целевую(ые) последовательность(и), содержащую мутацию(и) в NOI(s). Такие средства обнаружения более подробно описаны ниже в разделе «Обнаружение продукта(ов) ПЦР» и могут, например, представлять собой зонды для обнаружения мутаций.

Один из примеров способа идентификации фракции, содержащей одну или несколько специфических мутаций, включает следующие стадии:

- обеспечение подпулов и деление каждого подпула на фракции;
- получение образца из всей фракции каждого подпула;
- получение образца ДНК, например, геномной ДНК или мРНК, из указанного образца;
- проведение ПЦР-амплификаций со всеми индивидуальными образцами ДНК из всех подпулов, например, в индивидуальных лунках планшета, т.е. в микротитровальных планшетах;
- отбор тех образцов, которые содержат мутацию(и) в NOI(s).

Получение образцов ДНК

В настоящих способах не требуется, чтобы образцы ДНК были получены индивидуально для каждой из способных к регенерации частей, содержащихся в пределах одной фракции. Фактически, предпочтительно использовать всю фракцию для получения образца ДНК, которая на практике представляет собой объединенный в пул образец ДНК, содержащий ДНК от всех генотипов, содержащихся в пределах фракции. Образец ДНК может представлять собой образец gDNA или cDNA, например, образец cDNA, полученный из образца мРНК, как известно в данной области. Таким образом, предпочтительно, чтобы способы не включали стадию получения индивидуальных образцов ДНК для каждого растения или способной к регенерации части одной фракции или подпула.

Способы по изобретению включают одну или несколько стадий получения образцов ДНК, в частности, образцов gDNA или cDNA, полученных из образцов мРНК, предпочтительно образцов gDNA. В частности, способы могут включать одну стадию получения образцов ДНК из всей фракции.

Способы по изобретению могут также включать стадию получения образцов ДНК из вторичного подпула.

Как правило, указанные образцы ДНК подпулов получают таким образом, что образец ДНК теоретически содержит ДНК от каждого генотипа в пределах подпула. Образец ДНК может быть получен из всей фракции, при этом возможность репродукции способных к регенерации частей каждого генотипа сохраняется в другой фракции(ях) подпула.

Таким образом, каждый подпул, как правило, содержит более одного индивидуального семени или способные к регенерации части каждого генотипа, так что при разделении на фракции каждая фракция включает более одной индивидуальной способной к регенерации части каждого генотипа. В частности, может быть предпочтительно, чтобы каждый подпул содержал достаточное количество способных к регенерации частей каждого генотипа, чтобы иметь возможность произвольным образом разделить подпул или суперпул на фракции, находящиеся в количестве от 2 до 10, например, в количестве от 2 до 6, например, в количестве 2, 3 или 4, таким образом, чтобы каждая часть теоретически включала способные к регенерации части, представляющие каждый генотип подпула или суперпула.

Таким образом, часть способных к регенерации частей из подпула, т.е. в диапазоне от 10 до 90%, предпочтительно в диапазоне от 10 до 50%, например, в диапазоне от 15 до 50%, например, в диапазоне от 20 до 50% способных к регенерации частей каждого подпула может быть использовано для получения образца ДНК. Указанная часть способных к регенерации частей также упоминается в настоящем документе как «образец способных к регенерации частей». Остальные способные к регенерации части подпула могут храниться в условиях, поддерживающих репродуктивный потенциал указанных способных к регенерации частей. Может быть достаточно хранить семена указанных растений. Семена, например, зерна злаков, часто можно хранить в любом сухом и темном месте.

В то время как подпул предпочтительно включает несколько индивидуальных способных к регенерации частей каждого генотипа, как отмечено в настоящем документе в другом месте, вторичный подпул часто может включать только несколько

- иногда даже только одну - индивидуальных способных к регенерации частей каждого генотипа.

Как объяснялось выше, при получении ДНК из подпулов, обычно ДНК выделяют из всех фракций подпулов. Однако при получении ДНК из вторичных подпулов, указанная ДНК также может быть получена из образцов индивидуальных способных к регенерации частей. Когда образцы получают от индивидуальных способных к регенерации частей, может быть предпочтительно, чтобы образцы были получены способом, значительно не повреждающим указанные способные к регенерации части с точки зрения возможности репродукции. Таким образом, образец ДНК может быть получен из образца, содержащего или состоящего из части способной к регенерации части, такой как семя, которая не является существенной для репродукции. Образец может быть получен любым подходящим способом в зависимости от вида, например, с помощью биопсии, разрезания, сверления, растирания, разрыва или применения шприца с иглой. Например, в вариантах осуществления, где видом является злак, указанный образец предпочтительно содержит часть зерна злака, которая не является существенной для репродукции. Образец может быть получен различными способами, например, путем отрезания части(ей) зерна с помощью любого острого инструмента, такого как нож, скальпель или ножницы, путем стирания части зерна или путем просверливания отверстия в зерне. В последнем случае образцом может служить мука, полученная после сверления.

Как только становится доступной либо фракция способных к регенерации частей, либо образец, полученный из способных к регенерации частей, или их части (в данном документе также совместно именуемые как «образец»), образец ДНК может быть получен из указанных образцов любым подходящим способом. Если указанные образцы содержат крупные структуры, например, целые семена, первая стадия получения образца ДНК, такого как образец gDNA, обычно будет включать разделение указанного содержимого указанного образца на более мелкие части, например, с помощью физических средств, например, путем дробления или измельчения. Способы получения образца ДНК обычно включают стадии разрушения клеток и/или тканей, например, детергентом, ферментами (например, литазой), ультразвуком или их комбинациями – тем самым создавая неочищенный лизат. Указанный лизат может быть отделен от любого оставшегося дебриса любым подходящим способом. Неочищенный лизат может представлять собой образец ДНК, содержащий образец gDNA. Альтернативно, ДНК, например, gDNA или мРНК, используемые для

получения образца cDNA, могут быть дополнительно очищены, например, путем отделения указанной ДНК от остатка лизата, например, путем связывания с селективной матрицей, центрифугирования, градиентного центрифугирования и/или осаждения (например, с использованием осаждающего агента, такого как соль, спирт или магнитные шарики). Перед таким разделением другие компоненты лизата, включая белки и/или нуклеопротеины, могут быть денатурированы или разрушены, например, с использованием ферментов и/или денатурирующих агентов. Другие РНК-содержащие молекулы могут быть удалены, например, с помощью ферментов. Полезные способы получения образцов ДНК, таких как образцы cDNA или образцы gDNA, описаны, например, в Sambrook et al., *Molecular Cloning – Laboratory Manual*, ISBN 978-1-936113-42-2.

Чувствительные средства обнаружения

Способы по изобретению включают стадию идентификации фракции подпула и, следовательно, идентификации соответствующего представляющего интерес подпула с использованием средств обнаружения для обнаружения последовательностей нуклеиновых кислот, предпочтительно чувствительных средств обнаружения. Термин «чувствительные средства обнаружения» относится к средствам обнаружения, которые позволяют обнаруживать один мутантный организм в библиотеке, содержащей по меньшей мере 300 организмов, даже в тех случаях, когда генотип мутантного организма отличается от генотипов немутантных организмов максимум на один нуклеотид. Предпочтительно, чтобы средства обнаружения были достаточно чувствительными, чтобы позволить обнаружение одного мутантного генома в библиотеке из 300 геномов, даже если мутантный геном отличается от 299 других геномов только одним нуклеотидом. В некоторых вариантах осуществления средства обнаружения позволяют обнаруживать один мутантный организм, такой как растение, в библиотеке, содержащей по меньшей мере 300 организмов, таких как растения, например, один мутант в библиотеке, содержащей по меньшей мере 400 организмов, например, один мутант в библиотеке, содержащей по меньшей мере 500 организмов, например, один мутант в библиотеке, содержащей по меньшей мере 600 организмов, например, один мутант в библиотеке, содержащей по меньшей мере 700 организмов, например, один мутант в библиотеке, содержащей по меньшей мере 800 организмов, например, один мутант в библиотеке, содержащей по меньшей мере 900 организмов, например, один мутант в библиотеке, содержащей по меньшей мере 1000 организмов, например, один мутант в библиотеке, содержащей по меньшей мере 1100 организмов,

например, один мутант в библиотеке, содержащей по меньшей мере 1200 организмов,
 например, один мутант в библиотеке, содержащей по меньшей мере 1300 организмов,
 например, один мутант в библиотеке, содержащей по меньшей мере 1400 организмов,
 например, один мутант в библиотеке, содержащей по меньшей мере 1500 организмов,
 например, один мутант в библиотеке, содержащей по меньшей мере 1600 организмов,
 например, один мутант в библиотеке, содержащей по меньшей мере 1700 организмов,
 например, один мутант в библиотеке, содержащей по меньшей мере 1800 организмов,
 например, один мутант в библиотеке, содержащей по меньшей мере 1900 организмов,
 например, один мутант в библиотеке, содержащей по меньшей мере 2000 организмов,
 например, один мутант в библиотеке, содержащей по меньшей мере 2500 организмов,
 например, один мутант в библиотеке, содержащей по меньшей мере 3000 организмов,
 например, один мутант в библиотеке, содержащей по меньшей мере 4000 организмов,
 например, один мутант в библиотеке, содержащей по меньшей мере 5000 организмов,
 или более.

Например, чувствительные средства обнаружения имеют чувствительность обнаружения 1 по меньшей мере на 1000, такую как чувствительность обнаружения 1 на диапазон от 1000 до 10000, например, чувствительность обнаружения 1 на диапазон от 1000 до 5000.

Например, способы по изобретению могут включать стадию проведения множества ПЦР-амплификаций, каждая из которых включает образец ДНК из одной фракции, то есть из всех способных к регенерации частей, присутствующих в одной фракции (например, полученных, как описано в разделах выше), тем самым амплифицируя целевую последовательность. Каждая ПЦР-амплификация может включать множество разделенных на компартменты ПЦР-амплификаций, каждая из которых содержит часть указанного образца ДНК, один или несколько наборов праймеров, фланкирующих целевую последовательность, и реагенты для ПЦР.

РНКаза H2-зависимая ПЦР (rhPCR) может улучшить специфичность ПЦР за счет включения заблокированных праймеров, содержащих единичный рибонуклеотидный остаток, которые активируются посредством ферментативного расщепления. В частности, заблокированные праймеры для rhPCR содержат единичный рибонуклеотидный остаток, который индуцирует расщепление РНКазой H2 после идеальной гибридизации с ДНК-мишенью. Расщепление происходит на 5'-стороне основания РНК, высвобождая, таким образом, основание РНК вместе с 3'-блокирующим фрагментом, что позволяет проводить селективную амплификацию

фрагментов ДНК-мишеней. Хвостовая последовательность может быть включена в ампликон для объединения использования репортерной системы на основе анализа 5'-нуклеазы с rhPCR. В некоторых вариантах осуществления реакция ПЦР представляет собой реакцию rhPCR, например, как описано в примере 2.

В другом случае нуклеазы Cas с транс-расщепляющей активностью, такие как Cas12a, могут быть использованы в качестве чувствительных средств обнаружения (Chen et al., 2018). Недавно сообщалось о нескольких нуклеазах Cas с транс-расщепляющей активностью, которые можно использовать для обнаружения нуклеотидных замен в сложных образцах. В этом способе используется способность CRISPR направлять молекулу, в данном случае нуклеазу Cas12a, к определенному участку ДНК или РНК. Молекула CRISPR может быть сконструирована таким образом, чтобы быть специфичной для целевой последовательности, содержащей NOI. При связывании с целевой последовательностью, включающей NOI, Cas12a будет расщеплять специфическую цепь РНК или ДНК. В отличие от обычных нуклеаз, связанных с CRISPR, таких как Cas9, Cas12a также начинает расщеплять одноцепочечные молекулы ДНК или РНК, которые находятся в непосредственной близости от нуклеазы. В данном случае одноцепочечную ДНК или цепь РНК, репортерную молекулу, меченную флуоресцентным маркером, и гаситель используют в качестве субстрата для транс-расщепления Cas12a. В этом примере CRISPR-Cas12a специфически связывается с последовательностью ДНК или РНК, содержащей только NOI, и начинает расщеплять репортерную молекулу, находящуюся в непосредственной близости. Расщепление репортерной молекулы будет высвобождать флуорофор и передавать сигнал о присутствии молекулы ДНК или РНК, содержащей NOI. Если реакция содержит только целевую последовательность без NOI, комплекс CRISPR-Cas12a не связывается с мишенью и, следовательно, не будет инициировать транс-расщепление репортерных молекул, которые остаются погашенными.

Полная ПЦР-реакция, включающая множество разделенных на компартменты ПЦР-амплификаций, может быть проведена несколькими различными способами. В одном варианте осуществления ПЦР-амплификация, включающая множество разделенных на компартменты ПЦР-амплификаций, может быть проведена как цифровая ПЦР-амплификация (dPCR). Любая dPCR-амплификация, известная специалисту в данной области, может быть использована в изобретении. Как правило, по меньшей мере одну dPCR-амплификацию, содержащую множество разделенных на компартменты ПЦР-амплификаций, проводят для каждого полученного образца ДНК.

Таким образом, по меньшей мере одну амплификацию dPCR, включающую множество разделенных на компартменты амплификаций dPCR, будут проводить на фракцию каждого подпула.

Как правило, разделенную на компартменты dPCR-амплификацию проводят способом, включающим следующие стадии:

- проведение dPCR-амплификации, включающей образец ДНК, набор праймеров, фланкирующих целевую последовательность, и реагенты для ПЦР;
- разделение указанной dPCR-амплификации таким образом, чтобы молекулы нуклеиновой кислоты в образце были локализованы и сконцентрированы во множестве пространственно разделенных компартментов;
- проведение dPCR-амплификации;
- обнаружение продуктов амплификации на основе dPCR.

Образец ДНК может представлять собой образец gDNA или образец cDNA, полученный из образца мРНК. В некоторых вариантах осуществления образец ДНК представляет собой образец gDNA.

Указанные разделенные компартменты могут представлять собой любые отдельные компартменты, в которых могут быть проведены PCR-амплификации. Например, это может быть лунка планшета, например, лунка микролуночного планшета или микротитрационного планшета, это могут быть микрожидкостные камеры, это могут быть капилляры, это может быть дисперсная фаза эмульсии, или это могут быть капли или миниатюрные камеры массива миниатюрных камер. Разделенные компартменты также могут представлять собой дискретные точки на твердой подложке, т.е. дискретные поверхности связывания нуклеиновых кислот.

Обычно предпочтительно, чтобы образец ДНК был случайным образом распределен по образцам для разделенной на компартменты dPCR-амплификации. Также предпочтительно, чтобы каждая разделенная на компартменты dPCR-амплификация включала только небольшое количество нуклеиновых кислот, содержащих целевую последовательность. Из-за характера случайного распределения может наблюдаться некоторое колебание в количестве молекул нуклеиновых кислот, включенных в каждую разделенную на компартменты dPCR-амплификацию. В одном варианте осуществления каждая разделенная на компартменты dPCR-амплификация включает в среднем не более 10, такое как не более 5 молекул нуклеиновых кислот, содержащих целевую последовательность.

В одном предпочтительном варианте осуществления изобретения dPCR-амплификация, включающая множество разделенных на компартменты ПЦР-амплификаций, представляет собой капельную цифровую полимеразную цепную реакцию (ddPCR). ddPCR представляет собой способ проведения dPCR, который основан на технологии капли водно-масляной эмульсии. ПЦР-амплификация фракционирована на множество микрокапель, и ПЦР-амплификация целевой последовательности происходит в каждой отдельной капле. В целом, в технологии ddPCR используются реагенты для ПЦР и рабочие процессы, аналогичные тем, которые используются для проведения обычных ПЦР. Таким образом, разделение ПЦР-амплификации является ключевым аспектом метода ddPCR.

Соответственно, разделенные на компартменты ddPCR-амплификации могут содержаться в каплях, которые могут, например, включать эмульсионные композиции, или смеси двух или более несмешивающихся жидкостей (например, как описано в патенте США 7622280 или как описано в приведенных ниже примерах). Капли могут создаваться устройствами, описанными в WO/2010/036352. В частности, капли могут быть изготовлены с использованием генератора капель, например, QX200 Droplet Generator, доступного от Bio-Rad Laboratories, USA (далее сокращенно Bio-Rad). Используемый в настоящем документе термин «эмульсия» может относиться к смеси несмешивающихся жидкостей (таких как масло и вода). Эмульсии могут, например, представлять собой капли воды в масле, т.е. как описано в Hindson et al. (2011). Таким образом, эмульсии могут содержать водные капли в непрерывной масляной фазе. Эмульсии также могут представлять собой эмульсии масло-в-воде, где капли представляют собой капли масла в непрерывной водной фазе. Используемые в настоящем документе капли обычно разработаны для предотвращения смешивания между компартментами, при этом содержимое отдельного компартмента не только защищено от испарения, но также от слияния с содержимым других компартментов. Таким образом, каждую каплю можно рассматривать как пространственно отделенный компартмент.

Каждая капля для ddPCR может иметь любой полезный объем. Однако предпочтительно капли имеют объем в н.л.-диапазоне. Соответственно, предпочтительно, чтобы объем капель в среднем находился в диапазоне от 0,1 до 10 н.л.

Известно, что микрожидкостные способы получения капель эмульсии с использованием микроканальной фокусировки поперечного потока или физического

перемешивания позволяют получать либо монодисперсные, либо полидисперсные эмульсии. Капли могут представлять собой монодисперсные капли. Кроме того, капли могут быть генерированы таким образом, чтобы их размеры не отличались более чем на $\pm 5\%$ от среднего размера капель. В некоторых случаях капли генерируют таким образом, чтобы размер капель изменялся только на $\pm 2\%$ от среднего размера капель.

Более высокая механическая устойчивость может быть полезной для микрожидкостных манипуляций и жидкостной обработки с более высоким усилием сдвига (например, в микрожидкостных капиллярах или поворотах на 90° , таких как клапаны в жидкостных путях). Капли или капсулы до и после термической обработки могут быть механически устойчивы к стандартным манипуляциям с пипеткой и центрифугированию.

Капля может быть образована путем пропускания масляной фазы через водный образец. Водная фаза может содержать или состоять из компонентов ПЦР-амплификации, например, ПЦР-амплификации, включающей образец ДНК, набор праймеров, фланкирующих целевую последовательность, и реагенты для ПЦР, такие как любые реагенты для ПЦР, описанные ниже в настоящем документе в разделе «реагенты для ПЦР».

Масляная фаза может содержать фторированное базовое масло, которое может быть дополнительно стабилизировано комбинацией с фторированным поверхностно-активным веществом, таким как перфторированный полиэфир. В некоторых случаях базовым маслом может служить одно или несколько из HFE 7500, FC-40, FC-43, FC-70 или других распространенных фторированных масел. В некоторых случаях анионным поверхностно-активным веществом является Ammonium Krytox (Krytox-AM), аммониевая соль Krytox FSH или морфолино-производное Krytox-FSH.

Масляная фаза может дополнительно содержать добавку для регулирования свойств масла, таких как давление пара, вязкость или поверхностное натяжение. Неограничивающие примеры включают перфтороктанол и 1H,1H,2H,2H-перфтордеканол. Масляная фаза также может представлять собой масло, образующее капли, например, Droplet Generation Oil, доступное в компании Bio-Rad.

Эмульсия может быть составлена для получения высокомонодисперсных капель, имеющих жидкоподобную межфазную пленку, которые могут быть преобразованы путем нагревания в микрокапсулы, имеющие твердоподобную межфазную пленку; такие микрокапсулы могут вести себя как биореакторы, которые

сохраняют свое содержимое в процессе реакции, такой как ПЦР-амплификация. Превращение в форму микрокапсул может происходить при нагревании. Например, такое превращение может происходить при температуре выше 50, 60, 70, 80, 90 или 95°C. В некоторых случаях этот нагрев происходит с помощью термоциклера. Во время процесса нагрева можно использовать поверхностный слой жидкости или минерального масла для предотвращения испарения.

В некоторых случаях капля создается с помощью имеющегося в продаже генератора капель, такого как генератор капель Bio-Rad QX100™ или генератор капель Bio-Rad QX200™. ddPCR и последующее обнаружение можно проводить с использованием имеющегося в продаже устройства для считывания капель, такого как Bio-Rad QX100 или QX200™ Droplet Reader.

Каждая ПЦР-амплификация может быть разделена на любое подходящее количество компартментов. Однако в одном предпочтительном варианте осуществления каждая ПЦР-амплификация разделена на компартменты в диапазоне от 1000 до 100000 (например, капли). Например, каждая ПЦР-амплификация может быть разделена на компартменты в диапазоне от 10000 до 50000 (например, капли). Например, каждая ПЦР-амплификация может быть разделена на компартменты в диапазоне от 15000 до 25000 (например, капли). Кроме того, каждая ПЦР-амплификация может быть разделена примерно на 20000 компартментов (например, капли).

Организация подпулов

Перед стадией идентификации подпулов способы по изобретению могут также включать стадию организации указанных подпулов таким способом, который облегчает идентификацию подпула, содержащего способные к регенерации части, такие как семена, содержащие представляющую интерес мутацию.

Например, фракции образцов ДНК из нескольких подпулов могут быть объединены в «суперпулы». Суперпулы, содержащие ДНК, содержащую представляющую интерес мутацию, затем могут быть идентифицированы с использованием чрезвычайно чувствительных средств обнаружения. После этого должны быть проверены только подпулы, входящие в суперпулы, содержащие ДНК с представляющей интерес мутацией. Полезные способы получения и тестирования суперпулов описаны в международной патентной заявке WO 2018/001884.

В другом примере подпулы помещают в многомерные сетки, и суперпулы получают из подпулов в каждом измерении. Это проще всего объяснить на примере

двумерной сетки. Каждому подпулу присваивается координата (х, у) в указанной сетке. Фракции образцов ДНК всех подпулов с заданной координатой Х объединяют и тестируют на наличие ДНК, содержащей интересующую мутацию. Точно так же фракции образцов ДНК всех подпулов с заданной координатой Y объединяют и тестируют на наличие ДНК, содержащей интересующую мутацию. Как только суперпулы, содержащие образцы ДНК с заданной координатой Х и заданной координатой Y, идентифицируют как содержащие ДНК с интересующей мутацией, подпул, содержащий указанную ДНК, может быть идентифицирован на основе его координат Х и Y. Специалист в данной области техники поймет, что аналогичная система может быть создана с использованием нескольких измерений, а не только двух.

Таким образом, вместо тестирования каждого подпула на присутствие ДНК, содержащей интересующую мутацию, в способах по изобретению также предусматривается, что подпулы организованы таким образом, что группы подпулов можно тестировать вместе. Тестирование может быть проведено, например, с использованием любого из описанных здесь чувствительных средств обнаружения.

Реагенты для ПЦР

В некоторых вариантах осуществления способ включает проведение нескольких ПЦР-амплификаций, при этом по меньшей мере некоторые из этих ПЦР могут включать множество разделенных на компартменты ПЦР-амплификаций.

Независимо от того, включает ли указанная ПЦР разделенные на компартменты ПЦР-амплификации или нет, ПЦР-амплификации, как правило, будут включать образец ДНК, набор праймеров, фланкирующих целевые последовательности, и реагенты для ПЦР. Указанные реагенты для ПЦР могут представлять собой любые реагенты для ПЦР, описанные в данном разделе.

Реагенты для ПЦР, как правило, содержат по меньшей мере нуклеотиды и полимеразу нуклеиновых кислот. Нуклеотиды могут представлять собой молекулы дезоксирибонуклеотидтрифосфата, и предпочтительно реагенты для ПЦР содержат по меньшей мере dATP, dCTP, dGTP и dTTP. В некоторых случаях реагенты для ПЦР также содержат dUTP.

Полимеразой нуклеиновых кислот может быть любой фермент, способный катализировать зависимую от матрицы полимеризацию нуклеотидов, т.е. репликацию. Полимераза нуклеиновых кислот должна выдерживать температуры, используемые для

ПЦР-амплификации, и должна обладать каталитической активностью при температуре элонгации. Специалисту в данной области известно несколько термостабильных полимераз нуклеиновых кислот.

В некоторых вариантах осуществления изобретения полимераз нуклеиновой кислоты обладает 5'-3' нуклеазной активностью и, таким образом, может быть использована в реакции амплификации с зондом TaqMan®.

Полимеразой нуклеиновой кислоты может быть ДНК-полимераза I *Escherichia coli*. Полимеразой нуклеиновой кислоты также может быть ДНК-полимераза Taq, которая имеет зависимую от синтеза ДНК цепь, замещающую 5'-3'-экзонуклеазную активность. Другие полимеразы, обладающие 5'-3'-нуклеазной активностью, включают, но без ограничения, ДНК-полимеразу *rTth*. ДНК-полимераза Taq, например, полученная от New England Biolabs, может включать ДНК-полимеразу Crimson LongAmp® Taq, ДНК-полимеразу Crimson Taq, Hemo KlenTaq™ или LongAmp® Taq.

В некоторых случаях полимераз нуклеиновой кислоты может представлять собой, например, ДНК-полимеразу *E. coli*, фрагмент Klenow ДНК-полимеразы I *E. coli*, ДНК-полимеразу T7, ДНК-полимеразу T4, полимеразу Taq, ДНК-полимеразу Pfu, ДНК-полимеразу Vent, бактериофаг 29, REDTaq™, геномную ДНК-полимеразу или секвеназу. ДНК-полимеразы описаны, например, в публикации заявки на патент США 20120258501.

Кроме того, реагенты для ПЦР могут включать соли, буферы и средства обнаружения. Буфер может представлять собой любой полезный буфер, например, TRIS. Соль может представлять собой любую полезную соль, например, хлорид калия, хлорид магния или ацетат магния или сульфат магния.

Реагенты для ПЦР могут содержать неспецифический блокирующий агент, такой как BSA, желатин из бычьей кожи, бета-лактоглобулин, казеин, сухое молоко, ДНК спермы лосося или другие обычные блокирующие агенты.

Реагенты для ПЦР могут также содержать биоконсерванты (например, NaN₃), усилители ПЦР (например, бетаин, трегалозу и т.д.) и ингибиторы (например, ингибиторы РНКазы). Другие добавки могут включать диметилсульфоксид (DMSO), глицерин, бетаин(моно)гидрат, трегалозу, 7-деаза-2'-дезоксигуанозинтрифосфат (7-деаза-2'-dGTP), бычий сывороточный альбумин (BSA), формамид (метанамид), хлорид тетраметиламмония (TMAC), другие производные тетраалкиламмония [например, хлорид тетраэтиламмония (TEA-Cl)]; хлорид тетрапропиламмония (TPrA-Cl) или

неионогенное моющее средство, например, Triton X-100, Tween 20, Nonidet P-40 (NP-40) или PREXCEL-Q.

Кроме того, реагенты для ПЦР могут также включать одно или несколько средств для обнаружения продукта(ов) ПЦР-амплификации, содержащего мутацию(и) в NOI(s). Указанные средства могут представлять собой любые обнаруживаемые средства, и они могут быть добавлены в виде индивидуальных соединений или могут быть связаны или даже ковалентно связаны с одним из праймеров. Обнаруживаемые средства включают, но без ограничения, красители, радиоактивные соединения, биOLUMИнесцентные и флуоресцентные соединения. В предпочтительном варианте осуществления средством обнаружения является один или несколько зондов. Таким образом, предпочтительно, чтобы реагент для ПЦР содержал один или несколько зондов для обнаружения, например, любой из зондов, описанных в настоящем документе ниже в разделе «Обнаружение продукта(ов) ПЦР».

Праймеры, фланкирующие целевую последовательность

Способы по изобретению включают использование одного или нескольких наборов праймеров, которые фланкируют целевую последовательность. Дискретный «набор праймеров», который фланкирует целевую последовательность, включает один праймер, содержащий последовательность, идентичную 5'-концу целевой последовательности (также называемый «прямым праймером»), и один праймер, содержащий последовательность, комплементарную 3'-концу целевой последовательности (также называемый «обратным праймером»). Набор праймеров способен амплифицировать целевую последовательность при добавлении к ПЦР вместе с нуклеиновой кислотой, содержащей целевую последовательность и реагенты для ПЦР, в условиях, позволяющих амплифицировать указанную целевую последовательность. Один и тот же набор праймеров может быть использован для всех ПЦР-амплификаций в способах по изобретению, хотя также возможно использование разных наборов праймеров для ПЦР-амплификации на разных стадиях по изобретению.

В дополнение к последовательности, идентичной 5'-концу целевой последовательности, прямой праймер может содержать дополнительные последовательности. Аналогичным образом, в дополнение к последовательности, комплементарной 3'-концу последовательности-мишени, обратный праймер может содержать дополнительные последовательности. Например, праймеры могут на 5'-

конце содержать дополнительную последовательность нуклеиновой кислоты, которая не гибридизуется с целевой нуклеиновой кислотой, но которая облегчает обращение с праймером или продуктом ПЦР-амплификации, т.е. обнаружение указанного продукта.

Длина прямого праймера и обратного праймера может зависеть от последовательности целевой последовательности. Например, длину праймеров можно регулировать для достижения желаемой температуры плавления, T_m , праймеров. Таким образом, длина прямого праймера и обратного праймера по отдельности может составлять от 10 до 100 нуклеотидов, например, в диапазоне от 10 до 50 нуклеотидов, например, в диапазоне от 15 до 20 нуклеотидов, например, в диапазоне от 15 до 25 нуклеотидов, например, в диапазоне от 15 до 30 нуклеотидов, например, в диапазоне от 15 до 40 нуклеотидов, например, в диапазоне от 15 до 45 нуклеотидов, например, в диапазоне от 15 до 50 нуклеотидов в длину. T_m прямого праймера и обратного праймера обычно регулируют в диапазоне от 40 до 70°C.

Концентрация праймеров в водной фазе ПЦР-амплификаций может, например, находиться в диапазоне от 0,05 до 2,0 мкМ, например, в диапазоне от 0,1 до 1,0 мкМ, например, в диапазоне от 0,2 до 1,0 мкМ, например, в диапазоне от 0,3 до 1,0 мкМ, например, в диапазоне от 0,4 до 1,0 мкМ или в диапазоне от 0,5 до 1,0 мкМ.

Прямой праймер и обратный праймер в целом содержат – или даже состоят – из олигонуклеотидов. Однако в некоторых случаях праймеры могут содержать аналоги нуклеотидов. Многочисленные аналоги нуклеотидов известны специалистам в данной области техники и включают производные, в которых сахар модифицирован, как в 2'-О-метильных, 2'-дезоксидных и 2',3'-дидезоксинуклеозидных производных, аналоги нуклеиновой кислоты на основе других сахарных цепей, такие как треоза, закрытые нуклеиновые кислоты (LNA), производные LNA, пептидные нуклеиновые кислоты (PNA), гликолевая нуклеиновая кислота (GNA), треозная нуклеиновая кислота (TNA), бициклическая сахара или гексоза, глицерин и гликолевые сахара, аналоги нуклеиновых кислот на основе неионных остовов или нуклеиновые кислоты и их аналоги в нелинейных топологиях, таких как дендримеры, гребенчатые структуры и наноструктуры.

Праймеры также могут быть связаны с различными метками (например, флуоресцентными метками, функционализированными метками или связывающими метками), которые необязательно могут быть связаны с их концами, сахарами или азотистыми основаниями.

Праймеры могут быть получены различными способами, включая - но не ограничиваясь ими - клонирование соответствующих последовательностей и прямой химический синтез с использованием методов, хорошо известных в данной области [Narang et al., *Methods Enzymol.* 68:90 (1979); Brown et al., *Methods Enzymol.* 68:109 (1979)]. Праймеры также можно получить из коммерческих источников.

Прямой праймер и обратный праймер могут иметь одинаковую температуру плавления или сходную температуру плавления, например, температуру плавления $\pm 5^{\circ}\text{C}$. Длина праймеров может быть увеличена или укорочена на 5'- и/или 3'-конце(концах) с получением набора праймеров с желаемыми температурами плавления. Соответственно, один из праймеров пары праймеров может быть длиннее другого праймера.

Праймеры могут быть разработаны на основе температуры плавления. Уравнение для определения температуры плавления праймеров размером менее 25 пар оснований известно как правило Уоллеса [$T_d = 2 \times (A+T) + 4 \times (G+C)$]. Здесь T_d представляет собой температуру при определенной концентрации соли, при которой 50% олигонуклеотида и его идеально связанный с фильтром комплемент находятся в конформации дуплекса. Обычно T_d определяют в 0,9 М растворе NaCl. Однако при разработке праймера важны и другие соображения, т.е. его предполагаемая вторичная структура. Для разработки праймеров доступно несколько компьютерных программ и онлайн-сервисов.

В одном варианте осуществления набор праймеров может быть сконструирован таким образом, что праймеры специфически способны амплифицировать целевую последовательность, содержащую мутацию в NOI, но не способны амплифицировать целевую последовательность, содержащую эталонный NOI. Это может быть достигнуто, например, путем конструирования прямого праймера таким образом, чтобы он содержал последовательность, идентичную NOI(s), содержащему мутацию(и), и/или это может быть достигнуто путем конструирования обратного праймера таким образом, чтобы он содержал одну или несколько последовательностей, комплементарных NOI, содержащему мутацию(и).

В других вариантах осуществления набор праймеров может быть сконструирован таким образом, что праймеры способны специфически амплифицировать целевую последовательность, содержащую эталонный NOI, но не способны амплифицировать целевую последовательность, содержащую мутацию(и) в

NOI(s). Это может быть достигнуто, например, путем конструирования прямого праймера таким образом, чтобы он содержал последовательность, идентичную эталонному NOI, и/или путем конструирования обратного праймера, который содержит последовательность, комплементарную эталонному NOI.

Однако в некоторых вариантах осуществления изобретения наборы праймеров способны амплифицировать как целевую последовательность, содержащую мутацию(и) в NOI(s), так и целевую последовательность, содержащую эталонный NOI.

Примечательно, что способы по изобретению включают ПЦР-амплификации с более чем одним набором праймеров, например 2 наборами праймеров, например 3 наборами праймеров, например, в диапазоне от 2 до 10 наборов праймеров. Таким образом, каждая ПЦР-амплификация может включать несколько наборов праймеров, фланкирующих разные целевые последовательности. Это позволяет обнаруживать более одной различной мутации во время одной ПЦР-амплификации.

В вариантах осуществления, в которых ПЦР представляет собой rhPCR, ПЦР-реакцию предпочтительно проводят в присутствии блокированных праймеров, содержащих один рибонуклеотидный остаток. Такие блокированные праймеры активируются ферментативным расщеплением РНКазой HE2 после идеальной гибридизации с ДНК-мишенью. Расщепление происходит на 5'-стороне основания РНК, высвобождая, таким образом, основание РНК вместе с 3'-блокирующим фрагментом, что позволяет проводить селективную амплификацию целевых фрагментов ДНК. Хвостовая последовательность может быть включена в ампликон для объединения использования репортерной системы на основе анализа 5'-нуклеазы с rhPCR.

Обнаружение продукта(ов) ПЦР

В некоторых вариантах осуществления способы по изобретению включают стадию обнаружения продукта(ов) ПЦР-амплификации, содержащего целевую(ые) последовательность(и), которая содержит мутацию(и) в представляющем интерес NOI(s). Продукт(ы) ПЦР-амплификации можно обнаружить любым подходящим способом.

Как описано выше, указанная(ые) мутация(и) может представлять собой замену(ы), делецию(и) и/или вставку(и), включающую от только одного или нескольких нуклеотидов до большого количества нуклеотидов. Таким образом, обнаружение может быть адаптировано к конкретной мутации, указанной NOI.

В некоторых вариантах осуществления набор праймеров сконструирован таким образом, что праймеры специфически способны амплифицировать целевую последовательность, содержащую мутацию(и) в NOI(s), но не способны амплифицировать целевую последовательность, содержащую эталонную NOI. В других вариантах осуществления набор праймеров может быть сконструирован таким образом, что они специфически способны амплифицировать целевую последовательность, содержащую эталонную NOI, но не целевую последовательность, содержащую мутацию(и) в NOI(s). В таких вариантах осуществления обнаружение может быть просто основано на обнаружении присутствия или отсутствия продукта ПЦР-амплификации. Это может, в частности, иметь место в вариантах осуществления изобретения, где мутация представляет собой мутацию большего числа NOI(s).

В предпочтительных вариантах осуществления продукты ПЦР-амплификации обнаруживают с помощью зонда(ов) для обнаружения. Это может, в частности, иметь место, когда мутация(и) представляет собой мутацию меньшего числа NOI(s), например, одну или несколько точечных мутаций.

Зонд для обнаружения может представлять собой олигонуклеотид, содержащий последовательность, которая является идентичной NOI, содержащей мутацию. Кроме того, указанный зонд обычно также содержит последовательности, идентичные областям целевой последовательности, которые фланкируют NOI. Точно так же зонд для обнаружения может представлять собой олигонуклеотид, содержащий последовательность, которая является комплементарной NOI, содержащему мутацию, и, кроме того, указанный зонд может также содержать последовательности, комплементарные областям целевой последовательности, фланкирующим NOI. Указанные области, фланкирующие NOI, могут представлять собой последовательность(и), расположенную непосредственно на 5'- и/или 3'-конце NOI(s), соответственно. Такие зонды предпочтительно отжигают с продуктами ПЦР-амплификации, содержащими мутацию в NOI, но не с продуктами ПЦР-амплификации, содержащими эталонный NOI. Такие зонды для обнаружения также называются в настоящем документе как «зонды для обнаружения мутантов». Зонд для обнаружения мутантов обычно содержит от 10 до 30 нуклеотидов. В частности, зонд для обнаружения мутантов может содержать жесткую последовательность в диапазоне от 10 до 30 нуклеотидов, идентичную или комплементарную целевой последовательности, содержащей мутацию в NOI.

Зонд для обнаружения также может представлять собой олигонуклеотид, содержащий последовательность, которая является идентичной эталонному NOI. Кроме того, указанный зонд обычно также содержит последовательности, идентичные областям целевой последовательности, фланкирующим NOI. Точно так же зонд для обнаружения может представлять собой олигонуклеотид, содержащий последовательность, которая является комплементарной эталонному NOI, и указанный зонд может, кроме того, также содержать последовательности, комплементарные областям целевой последовательности, фланкирующим NOI. Такие зонды предпочтительно отжигают с продуктами ПЦР-амплификации, содержащими эталонный NOI, но не с продуктами амплификации, содержащими мутацию в NOI. Такие зонды для обнаружения также называются в настоящем документе как «эталонные зонды для обнаружения». Эталонный зонд для обнаружения обычно содержит от 10 до 30 нуклеотидов. В частности, эталонный зонд для обнаружения может содержать жесткую последовательность в диапазоне от 10 до 30 нуклеотидов, идентичную или комплементарную целевой последовательности, содержащей эталонный NOI.

В некоторых вариантах осуществления реагенты для ПЦР содержат набор зондов, состоящий из зонда для обнаружения мутантов и эталонного зонда для обнаружения. «Набор зондов» предпочтительно способен конкурировать за сайт(ы) связывания в NOI(s) при добавлении к ПЦР-амплификации вместе с нуклеиновой кислотой, содержащей целевую последовательность, ПЦР-реагенты и набор праймеров в условиях, обеспечивающих возможность амплификации указанной целевой последовательности. Как отмечалось выше, зонды для обнаружения обычно представляют собой олигонуклеотиды. В частности, это могут быть короткие нуклеотидные участки одноцепочечной ДНК. Зонды для обнаружения могут быть связаны или даже ковалентно связаны с обнаруживаемыми средствами, включая, но без ограничения, репортеры, красители, радиоактивные соединения, биOLUMИнесцентные соединения, флуоресцентные соединения и пары флуорофор/гаситель. В некоторых вариантах осуществления самый крайний 5'-нуклеотид зонда для обнаружения не представляет собой G. Таким образом, зонд для обнаружения мутации может содержать олигонуклеотид, где самый крайний 5'-нуклеотид не представляет собой G. Аналогично, эталонный зонд для обнаружения может содержать олигонуклеотид, где самый крайний 5'-нуклеотид не представляет собой G.

Способы по изобретению могут включать использование обоих, зонда для обнаружения мутантов и эталонного зонда для обнаружения. В таких случаях предпочтительно, чтобы зонды были помечены по-разному, т.е. таким образом, что один из зондов связан или ковалентно связан с обнаруживаемыми средствами, тогда как другой - нет. Также возможно, что оба зонда связаны или ковалентно связаны с обнаруживаемыми средствами, при этом указанные обнаруживаемые средства являются различными.

В одном варианте осуществления эталонный зонд для обнаружения помечен флуорофором на 5'-конце и гасителем на 3'-конце. Указанные флуорофор и гаситель могут, например, представлять собой HEX и Black-Hole Quencher, соответственно. В одном варианте осуществления зонд для обнаружения мутантов помечен флуорофором на 5'-конце и гасителем на 3'-конце. Указанные флуорофор и гаситель могут, например, представлять собой FAM и Black-Hole Quencher, соответственно.

ПЦР-реакции могут включать одинаковое количество зонда для обнаружения мутантов и зонда для обнаружения дикого типа. Однако в некоторых вариантах осуществления может быть предпочтительным использование избытка зонда для обнаружения мутантов. В частности, это может иметь место при обнаружении ПЦР-продукта(ов) после ПЦР-амплификации в супер-пуле. В таких случаях ПЦР-реакции могут включать по меньшей мере 2-кратный, более предпочтительно по меньшей мере 4-кратный избыток мутантного зонда для обнаружения по сравнению с эталонным зондом для обнаружения.

В некоторых вариантах осуществления зонд для обнаружения представляет собой зонд TaqMan® (Heid et al., 1996), в котором используется 5'-экзонуклеазная активность полимеразы нуклеиновой кислоты. Вот почему реагенты для ПЦР предпочтительно содержат полимеразу нуклеиновых кислот с 5'-экзонуклеазной активностью (например, полимеразу Taq). Как правило, зонд TaqMan® может представлять собой либо зонд для обнаружения мутантов, как описано выше, либо эталонный зонд для обнаружения, как описано выше, ковалентно связанный с парой флуорофор/гаситель. Таким образом, зонд может содержать флуорофор, обычно у 5'-основания или рядом с ним. Кроме того, зонды TaqMan® могут содержать гаситель, который может находиться у 3'-основания или рядом с ним и способен гасить флуоресценцию указанного флуорофора [Tyagi et al. (1998)]. При облучении возбужденный флуорофор передает энергию ближайшему гасителю, а не флуоресцирует [Фёстеровский или резонансный перенос энергии флуоресценции

(FRET)]. Таким образом, непосредственная близость флуорофора и гасителя может предотвратить испускание любой флуоресценции, в то время как зонд не поврежден. Однако, когда зонд TaqMan® отжигается с внутренней областью целевой последовательности и полимераза участвует в репликации матрицы, с которой связан зонд TaqMan®, ее 5'-экзонуклеазная активность может расщеплять зонд. Эта серия событий отменяет функцию гашения, то есть отсутствие FRET, и флуорофор начинает излучать флуоресценцию, которую можно измерить любым подходящим способом.

Примечательно, как подчеркивалось выше, что изобретение также включает в себя то, что способы могут быть использованы для идентификации более чем одной различной мутации. Этого можно достичь, используя множество наборов праймеров, как описано выше. Однако этого также можно достичь, используя несколько различных зондов для обнаружения. Таким образом, в одном варианте осуществления способы можно использовать для идентификации более чем одной различной мутации NOI в целевой последовательности. В таких вариантах осуществления реагенты для ПЦР могут содержать эталонный зонд для обнаружения и несколько зондов для обнаружения мутаций, где каждый зонд для обнаружения мутаций содержит олигонуклеотид, имеющий последовательность, идентичную или комплементарную одной из мутаций. Предпочтительно эталонный зонд обнаружения и зонды для обнаружения мутаций связаны с разными средствами обнаружения. Все зонды для обнаружения мутаций могут быть связаны с одними и теми же средствами обнаружения, или с аналогичными средствами обнаружения, или с разными средствами обнаружения.

В некоторых случаях зонд для обнаружения представляет собой молекулярный маяк (MB), то есть зонд, содержащий комплементарные последовательности, способные к самогибридизации (также называемые «стебель»), что приводит к образованию структуры «петля-шпилька». Петля MB может содержать последовательности, комплементарные или идентичные NOI, либо мутантной, либо эталонной NOI. Кроме того, MB обычно содержит флуорофор и гаситель, расположенные на обоих концах MB, так что они находятся в непосредственной близости друг от друга, когда зонд гибридизуется сам с собой. Молекулярный маяк (MB) может представлять собой либо зонд для обнаружения мутантов, как описано выше, либо эталонный зонд для обнаружения, как описано выше, ковалентно связанный с парой флуорофор/гаситель и со стволочной последовательностью(ями), обеспечивающей комплементарные области MB. Дополнительные сведения о

стандартных методах изготовления и использования МВ хорошо известны в литературе, и МВ доступны от ряда источников коммерческих реагентов.

В некоторых случаях используют праймер/зонд; в некоторых случаях праймер/зонд представляет собой зонд Scorpions™, который может обеспечить основанный на FRET механизм обнаружения стебель-петля, аналогичный МВ, за исключением того, что к зонду также прикреплен сегмент, который служит либо прямым, либо обратным праймером (Whitcombe et al. и др. (1999); патент США 6326145). Зонд Scorpions™ может поддерживать конфигурацию «стебель-петля» в негибридизованном состоянии с погашенным флуорофором. Зонд Scorpions™ может иметь более длинную многокомпонентную структуру, т.е. 5'-флуорофор, затем участок стебель-петля, специфичный для мишени, затем гаситель, затем блокатор [например, гексэтиленгликоль (HEG)] и, наконец, 3' последовательность праймера. Блокатор может предотвратить обратное удлинение продукта на зонд.

В некоторых случаях праймер/зонд представляет собой зонд Sunrise™, который содержит праймер, прикрепленный к зонду-шпильке, который удлиняется во время амплификации. Такое расположение может отделить внутренний гаситель от 5'-концевого флуорофора (Nazarenko et al. (1997)).

Зонд для обнаружения может иметь любую полезную длину, например, в диапазоне от 10 до 60 нуклеотидов в длину. Олигонуклеотидный зонд также может иметь длину в диапазоне от 10 до 30 нуклеотидов. Точная последовательность и длина олигонуклеотидного зонда могут отчасти зависеть от природы целевого полинуклеотида, с которым он связывается. Место связывания и длина могут варьироваться для достижения соответствующих свойств отжига и плавления в конкретной ситуации. Например, зонд для обнаружения может иметь температуру плавления в том же диапазоне, что и прямой праймер и/или обратный праймер, т.е. в пределах $\pm 10^{\circ}\text{C}$, например, в пределах $\pm 5^{\circ}\text{C}$.

3'-Концевой нуклеотид зонда(ов) для обнаружения может быть блокирован или лишен способности к удлинению с помощью полимеразы нуклеиновой кислоты. Такое блокирование может быть удобно осуществлено путем прикрепления обнаруживаемых средств, либо через флуорофор, либо через гаситель к концевому 3'-основанию олигонуклеотидного зонда с помощью связывающего фрагмента.

В литературе имеется большое количество практических руководств, описывающих полезные пары флуорофор-гаситель, включая способы выбора пар флуорофор-гаситель, как показано ниже:

Clegg, *Meth. Enzymol.*, 211: 353-388 (1992); Wo et al., *Anal. Biochem.*, 218: 1-13 (1994); Pesce et al., editors, *Fluorescence Spectroscopy* (Marcel Dekker, New York, 1971); White et al., *Fluorescence Analysis: A Practical Approach* (Marcel Dekker, New York, 1970); и тому подобное;

– литература также включает ссылки, обеспечивающие исчерпывающие списки флуоресцентных и хромогенных молекул, а также их соответствующие оптические свойства для выбора пар репортер-гаситель, например, Berlman, *Handbook of Fluorescence Spectra of Aromatic Molecules*, 2nd Edition (Academic Press, New York, 1971); Griffiths, *Colour and Constitution of Organic Molecules* (Academic Press, New York, 1976); Bishop, editor, *Indicators* (Pergamon Press, Oxford, 1972); Haugland, *Handbook of Fluorescent Probes and Research Chemicals* (Molecular Probes, Eugene, 1992) Pringsheim, *Fluorescence and Phosphorescence* (Interscience Publishers, New York, 1949); и тому подобное;

– кроме того, в литературе имеется обширное руководство по дериватизации репортерных молекул и молекул-гасителей для ковалентного связывания с помощью общих реакционноспособных групп, которые могут быть добавлены к олигонуклеотиду, как показано в следующих ссылках: Haugland (cited above); Ullman et al., патент США 3996345; Khanna et al., патент США 4351760; и тому подобное.

Флуорофоры и гасители могут быть, например, выбраны из флуоресцеиновых и родаминовых красителей. Эти красители в сочетании с подходящими методиками связывания для прикрепления к олигонуклеотидам описаны во многих источниках, например Khanna et al. (цитируется выше); Marshall, *Histochemical J.*, 7:299-303 (1975); Menchen et al., патент США 5188934; Menchen et al., европейская заявка на патент 87310256.0; и Bergot et al., международная заявка PCT/US90/05565. Последние четыре документа включены в настоящий документ посредством ссылки.

Пара флуорофор/гаситель может использовать флуорофор, такой как EDANS или флуоресцеин, например, на 5'-конце и гаситель, такой как Dabcyl, например, на 3'-конце.

Обнаружение продукта(ов) ПЦР-амплификации может включать стадию сравнения сигнала, полученного при ПЦР-амплификации, включающей данный образец ДНК, с сигналом, полученным при контрольной ПЦР-амплификации. Сигнал обычно может относиться к средствам обнаружения, связанным с зондом для обнаружения мутантов и/или зондом для обнаружения дикого типа. Таким образом,

если указанные зонды связаны с флуорофором, сигнал может представлять собой флуоресценцию. Контрольная ПЦР-амплификация может представлять собой ПЦР-амплификацию, проводимую в тех же условиях, но без образца ДНК, т.е. в указанной контрольной ПЦР-амплификации может отсутствовать матрица ДНК, или она может содержать только контрольную ДНК, содержащую только эталонную целевую последовательность, например, ДНК дикого типа. В некоторых вариантах осуществления любой сигнал, который является более сильным, чем сигнал контрольной ПЦР-амплификации, может считаться положительным сигналом, то есть сигналом, указывающим на присутствие целевой последовательности, содержащей одну или несколько мутаций в NOI.

В одном варианте осуществления любой сигнал, превышающий заданное пороговое значение, может считаться положительным сигналом. Пороговое значение может быть определено любым подходящим способом, обычно с использованием подходящего программного обеспечения, например, QX200™ Droplet Reader и Quantasoft™, поставляемые компанией Biorad.

В вариантах осуществления изобретения, в которых используются эталонные зонды для обнаружения мутаций, может быть определена относительная распространенность между сигналом, полученным от зонда для обнаружения мутаций, и сигналом, полученным от эталонного зонда для обнаружения, и использовано для оценки присутствия продукта ПЦР. Относительная распространенность может быть определена как: [Сигнал, полученный от («зонд для обнаружения мутантов»)], разделенный на [Сигналы, полученные от («эталонного зонда для обнаружения» + «зонда для обнаружения мутантов»)].

Например, можно предположить, что образец содержит целевую ДНК, содержащую мутацию в NOI, при условии, что он по сравнению с контрольной ПЦР-амплификацией характеризуется:

- 1) повышенной относительной распространенностью и/или;
- 2) повышенной концентрацией мутантных капель и/или;
- 3) повышенным числом мутантных событий в масштабе 50% или более выше среднего.

Мутантные капли могут представлять собой капли, дающие положительный сигнал от зонда для обнаружения мутантов. В настоящем документе «мутантное событие» равно числу мутантных капель.

Идентификация способных к регенерации частей в подпуле

После идентификации подпула, включающего способные к регенерации части с ДНК, содержащей интересующую(ие) мутацию(и), способы обычно включают стадию идентификации способных к регенерации частей в указанном подпуле, содержащем указанную мутацию. Это может быть сделано любым общепринятым способом. Если подпул достаточно мал, это можно просто сделать путем тестирования каждого семени или его потомства в указанном подпуле на наличие мутации. Например, образец может быть взят из каждой способной к регенерации части, или каждая способная к регенерации часть может быть выращена в растение, и образец от указанного растения может быть использован для обнаружения того, несет ли указанное растение мутацию.

Однако часто могут применяться менее трудоемкие способы идентификации способных к регенерации частей в подпуле, несущем мутацию. Например, подпул может быть разделен на вторичные подпулы, и может быть идентифицирован(ы) вторичный(е) подпул(ы), содержащий способные к регенерации части с ДНК, содержащей интересующую(ие) мутацию(и). Это может быть сделано по существу таким же способом, как описано в настоящем документе выше для идентификации подпула, содержащего интересующую способную к регенерации часть. Однако, поскольку подпул обычно значительно меньше исходного пула, могут быть также использованы менее чувствительные способы.

В одном варианте осуществления интересующие способные к регенерации части идентифицируют с помощью способа, включающего следующие стадии:

- обеспечение подпула, включающего множество способных к регенерации частей растения, при этом указанный подпул содержит способную к регенерации часть(и), содержащую одну или несколько мутаций NOI;
- разделение способных к регенерации частей указанного подпула на вторичные подпулы;
- получение образца из каждой способной к регенерации части вторичных подпулов - таким образом, чтобы способные к регенерации части оставались достаточно неповрежденными для развития растения - и объединение всех образцов от всех способных к регенерации частей каждого вторичного подпула;
- подготовка образца ДНК из указанных объединенных образцов;
- идентификация вторичного подпула(ов), содержащего ДНК, содержащую мутацию(и);
- идентификация способной к регенерации части(ей) в указанном вторичном подпуле, содержащем указанную мутацию.

В одном варианте осуществления стадия идентификации способных к регенерации частей в указанном вторичном подпуле, содержащем указанную мутацию, включает тестирование каждой способной к регенерации части, такой как каждое семя или его потомство, в указанном вторичном подпуле на присутствие мутации. Например, каждая способная к регенерации часть может быть выращена в растение, и образец из указанного растения может быть использован для определения того, несет ли указанное растение мутацию.

Фенотип растения и измененный признак (признаки)

Настоящие методы позволяют использовать для создания библиотек лучшие сорта и самые передовые селекционные линии. Это гарантирует, что желаемые признаки внедряются в желаемый фон, например, адаптированного к региону, высокопродуктивного, климатоустойчивого и/или устойчивого к болезням селекционного материала. Как объяснялось выше, настоящие способы снижают или устраняют необходимость обратного скрещивания. Настоящие способы основаны на случайном мутагенезе. Напротив, технологии GM или редактирования генов обычно требуют стадии трансформации и поэтому ограничены видами, которые могут быть трансформированы, или неэлитными, трансформируемыми разновидностями, которые интенсивно растут в культуре тканей, но плохо адаптированы к местной среде.

Термин фенотип растения представляет собой термин, используемый для описания наблюдаемых характеристик, таких как высота, урожайность, размер зерен, форма листа, биомасса и т.д. растения. Фенотип растения обычно определяется как коллективное выражение генотипа растения в сочетании с влиянием окружающей среды на наблюдаемые характеристики растения.

В ряде случаев растения одного генотипа при взаимодействии со средой могут проявлять разные фенотипы. В других случаях растения могут иметь один и тот же фенотип, но разные генотипы.

После мутагенеза и отбора растения с конкретной желаемой мутацией мутантное растение по настоящему изобретению имеет генотип, отличный от родительского растения. В частности, мутантное растение содержит от 1 до 30 несинонимичных мутаций в кодирующих областях, как подробно описано выше.

В одном варианте осуществления мутантное растение и родительское растение могут иметь одинаковый фенотип. В другом варианте осуществления растения имеют разные фенотипы. Предпочтительно, чтобы мутантное растение имело те же наблюдаемые характеристики, что и родительское растение, за исключением

характеристики/признака, вызванного мутацией(ями) в NOI(s). Другими словами, предпочтительно, чтобы мутагенизированное растение имело тот же фенотип, что и родительское растение, за исключением желаемого мутагенизированного признака.

Таким образом, мутация(и) приводит к одному или нескольким специфическим признакам, которые предпочтительно будут единственными признаками, отличающимися между мутантным растением и растением поколения M0 (т.е. до мутагенеза). Это возможно из-за низких скоростей мутагенеза, используемых в настоящих способах.

В некоторых вариантах осуществления один или несколько из следующих параметров: внешний вид, высота, биомасса, форма листа и вкус мутантного растения являются такими же, как у родительского растения, предпочтительно все из параметров: внешний вид, высота, биомасса, урожайность, размер зерна, форма листа и вкус мутантного растения являются такими же, как и у родительского растения.

Урожайность, качество зерна, стерильность и другие признаки

Высокая скорость мутагенеза часто приводит к снижению урожайности растения. Поскольку настоящие способы основаны на низких скоростях мутагенеза, как описано в настоящем документе выше, могут быть получены мутагенизированные растения, имеющие урожайность, сходную с родительским растением. Мутантное растение может, в частности, давать урожай, который составляет по меньшей мере 90%, например, по меньшей мере 91%, например, по меньшей мере 92%, например, по меньшей мере 93%, например, по меньшей мере 94%, например, по меньшей мере 95%, например, по меньшей мере 96%, например, по меньшей мере 97%, например, по меньшей мере 98%, например, по меньшей мере 99% или, например, 100% от общей урожайности родительского растения, или более. В некоторых вариантах осуществления мутантное растение может иметь урожайность, которая выше, чем общая урожайность родительского растения. Таким образом, в некоторых вариантах осуществления мутантное растение может обеспечивать урожайность, которая составляет 105% от общей урожайности родительского растения или более, например 110% или более, например 120% или более, например 130% или более, например 140% или более, например 150% или более, например 175% или более, например 200% или более.

В другом варианте осуществления изобретения предварительно определено, что NOI влияет на урожайность мутантного растения, таким образом, в этом варианте

осуществления урожайность мутантного растения будет изменена по сравнению с урожайностью родительского растения. В предпочтительном варианте осуществления урожайность мутантного растения является по меньшей мере такой же, как урожайность родительского растения, или превышает ее. Поскольку урожайность может варьироваться в зависимости от условий выращивания или природы почвы, предпочтительно сравнивать урожайности мутантного растения и родительского растения для растений, растущих в сходных условиях.

Мутантные растения, полученные с помощью настоящих способов, предпочтительно имеют качество зерна, которое по существу такое же или выше, чем качество зерна родительского растения. Таким образом, в некоторых вариантах осуществления интересующая мутация не оказывает отрицательного влияния на качество зерна. Соответственно, содержание влаги в зерне, тестовая масса, процент обесцвеченных, сломанных или поврежденных ядер, ломкость зерна, качество помола, содержание белка и масла, и жизнеспособность зерна, являются такими же или больше, чем соответствующие признаки в родительском растении. Кроме того, один или несколько из следующих параметров: цвет зерна, состав, объемная плотность, запах и аромат, размер и форма предпочтительно являются такими же или лучше, чем соответствующие признаки в родительском растении.

Желаемые мутации предпочтительно не повышают или лишь незначительно повышают стерильность растений. Таким образом, в некоторых вариантах осуществления мутантные растения, полученные с помощью настоящих способов, т.е. процент стерильных цветков на колос, имеют стерильность ниже 30%.

ПРИМЕРЫ

Пример 1. Протокол с низкой плотностью мутаций

Для индукции мутаций, ядра, собранные с растений ячменя, инкубировали в растворе с мутагеном, азидом натрия (NaN_3), в соответствии с подробным описанием, предоставленными в работе Kleinhofs et al. (1978) и в US 7838053, со следующими изменениями: концентрация азиды натрия, используемого для мутагенеза, была снижена до 0,3 мМ в течение 2 часов.

Эта процедура вызывает точечные мутации в геномной ДНК (gDNA) зерен ячменя, обычно наделяя случайно распределенными кодонами для аминокислотных замен или остановок трансляции кодирующей белок ДНК, т.е. приводя к белковым изменениям и укорочениям в белках, кодируемых мутагенизированной ДНК,

соответственно. Однако способы по изобретению также применимы для получения злаковых растений с точечными мутациями в gDNA не кодирующих белок областей, например, промоторов, терминаторов и интронах.

15 кг зерна подвергли мутагенезу и высаживали на 30 участках по 7,5 м² каждый. Затем участки разделяли на 17 подучастков, каждый из которых соответствовал подпулу, из которого были созданы банки семян и образцы ДНК.

Пример 2. РНКазы H2-зависимая ПЦР (rhPCR) для обнаружения мутанта(ов) в библиотеке мутантов

РНКазы H2-зависимая ПЦР (rhPCR) (Dobosy et al., 2011) может повысить специфичность ПЦР-амплификации за счет использования так называемых «заблокированных праймеров». Эти праймеры содержат один рибонуклеотидный остаток и 3'-блокирующий фрагмент, который предотвращает удлинение праймера во время ПЦР. Однако праймеры могут быть активированы РНКазой H2, что затем позволяет им удлиняться во время ПЦР-амплификации. Расщепление РНКазой H2 будет происходить только в том случае, если заблокированный праймер связывается с идеальным совпадением. Даже несоответствие только нуклеотидной замены может значительно ингибировать расщепление. Это означает, что заблокированные праймеры могут быть разработаны для специфического блокирования амплификации неспецифических или нежелательных ампликонов, таким образом обеспечивая очень чувствительный инструмент для обнаружения редких специфических нуклеотидных замен. Зонд, меченный флуорофором, комплементарный 5'-концу праймера для rhPCR, позволяет количественно оценить успешное удлинение праймера во время rhPCR определенного ампликона.

Уникальный анализ rhPCR разработан (rhAmp® SNP Genotyping System, Integrated DNA Technologies) для различения мутантного аллеля и аллеля дикого типа гена *Hordeum vulgare*. rhPCR-праймер для обнаружения мутанта комплементарен кодирующей последовательности, включающей NOI. Эталонный rhPCR-праймер для обнаружения комплементарен кодирующей последовательности. Эталонный rhPCR-праймер и праймер для обнаружения мутаций дополнительно мечены, например, флуорофором FAM и Yakima Yellow, соответственно.

Этот анализ используют для скрининга библиотеки подпулов, полученных из пула особей M2. Очищенную gDNA добавляют к смеси для ПЦР, содержащей необходимые реагенты. Образцы подвергают термическому циклированию, используя

стандартные условия ПЦР. Продукты реакции количественно оценивают на флуоресцентном ридере.

Данные, полученные на флуоресцентном ридере, анализируют путем сравнения среднего сигнала флуоресценции всех подвергнутых скринингу подпулов с сигналом флуоресценции отдельных подпулов. Положительный подпул считается кандидатом, когда индивидуальный сигнал флуоресценции является по меньшей мере в 2 раза более высоким, чем средний сигнал флуоресценции всех подпулов.

Подпулы, демонстрирующие флуоресценцию выше среднего сигнала флуоресценции, считаются содержащими мутант. Для подтверждения присутствия мутанта в подпуле часть указанного подпула подвергают неdestructивному выделению ДНК. В этом случае образцы ткани, полученные из отдельных зерен ячменя, берут и подвергают анализу ДНК в пулах. Каждый из этих пулов ДНК анализируют с помощью того же анализа *rhPCR*, который описан выше. Идентифицируют пулы ДНК, имеющие сигнал флуоресценции, который по меньшей мере в 2 раза выше, чем средний сигнал флуоресценции остальных пулов ДНК. Сигнал мутации в этих пулах подтверждает наличие желаемого мутанта.

Пример 3. Прямое секвенирование низкомуtagenизированных растений ячменя

Авторы изобретения секвенировали библиотеку из 6000 растений ячменя (сорт Quench). Зерна подвергали мутации с использованием 0,3 мМ азида натрия в течение 2 часов. Была построена библиотека, состоящая из 6000 индивидуальных колосьев МЗ. Для выделения ДНК использовали одно зерно на колос.

Авторы изобретения амплифицировали фрагмент интересующего гена длиной 790 bp (пар нуклеотидов) для анализа 645 bp экзона 3 этого гена. Ампликоны для всех 6000 линий подвергали анализу на полиморфизмы. Было идентифицировано 3 отдельных мутанта в пределах 645 bp для 6000 линий. Это соответствует $3/(645 \times 6000) = 1$ мутанту на каждые 1,29 миллиона bp, что соответствует примерно 24 несинонимичным мутациям в кодирующих областях.

Три мутанта соответствовали следующим сайтам мутаций:

1. Нуклеотид G989>A (cDNA), соответствующий аминокислотной замене S330>N (белок);
2. Нуклеотид C1091>T (cDNA), соответствующий аминокислотной замене P364>L (белок);
3. Нуклеотид C858>T (cDNA) не вызвал аминокислотной замены.

Аmplифицировали фрагмент гена AAP3 (HvAAP3) длиной 860 bp, чтобы проанализировать 648 bp экзона 6 этого гена. Ампликоны для всех 6000 линий подвергали анализу на полиморфизмы. Было идентифицировано 3 отдельных мутанта в пределах 648 bp для 6000 линий. Это соответствует 1 мутанту на каждые 1,296 миллиона bp, что соответствует примерно 24 несинонимичным мутациям в кодирующих областях, что соответствует 99,94% генов, не содержащих несинонимичных мутаций.

Три мутанта соответствовали следующим сайтам мутаций:

4. Нуклеотид C982>T (cDNA), соответствующий аминокислотной замене H328>Y (белок);
5. Нуклеотид G820>A (cDNA), соответствующий аминокислотной замене A274>T (белок);
6. Нуклеотид C851>T (cDNA), соответствующий аминокислотной замене A284>V (белок).

Пример 4 - скрининг на мутанты ячменя с преждевременным стоп-кодоном в аминокислотной пермеазе 3 (HvAAP3)

Мутагенез ячменя сорта. Planet выполняли, как описано в примере 1, для достижения плотности мутаций, равной около 1 мутации на 1,3 миллиона bp, что соответствует примерно 24 несинонимичным мутациям в кодирующих областях, что соответствует 99,94% генов, не содержащих несинонимичных мутаций.

Что касается приведенного ниже примера 6, оптимальный размер библиотеки составляет приблизительно 520,000, если желательна вероятность успеха >94% и выполняется поиск 7 различных мутаций. Таким образом, оптимальное количество способных к регенерации частей для культивирования составляет 99,640 при расчете в соответствии с приведенным ниже примером 7.

Скрининг проводили с использованием протокола ddPCR с чувствительностью обнаружения 1 в 2000, ячмень имеет количество GECN, равное 5-6. На этой основе теоретический максимальный размер подпула, рассчитанный, как показано в примере 8 ниже, составил 363.

Таким образом, подвергали скринингу библиотеку из 374 подпулов, полученных из пула из 110000 растений, для идентификации одного положительного подпула. Таким образом, каждый подпул содержал зерна поколения M1 от приблизительно 294 растений. Затем часть положительного подпула снова подвергали скринингу для идентификации мутантного зерна. Скрининг проводили с помощью ddPCR, как

описано в PCT/EP2017/065516. Искомая мутация представляла собой стоп-кодон и, соответственно, приемлемыми были 7 различных мутаций. Один из желаемых мутантов имеет нуклеотидную замену в положении 820 с гуанина на аденин в кодирующей последовательности HvAAP3, которая доступна под номером HORVU.MOREX.r2.7HG0553990.1 в базе данных IPK Barley Blast Server Международного консорциума по секвенированию ячменя (29 сентября 2019 г.), что вызывает аминокислотную замену в положении 274, что приводит к аминокислотной замене аланина на треонин в белковой последовательности, которая доступна под номером HORVU.MOREX.r2.7HG0553990.1 в базе данных IPK Barley Blast Server Международного консорциума по секвенированию ячменя (29 сентября 2019 г.). Для идентификации мутанта разрабатывали специфический для мишени прямой праймер (CCGGCCGAGAACAAG), специфический для мишени обратный праймер (CGTGGTGGTGGAGAC), специфический для мутанта зонд для обнаружения (AGGAAGACAAACCTACTGG), меченный 6-карбоксифлуоресцеином (FAM), и специфический для эталона зонд для обнаружения (AGGAAGGCAAACCTACTG), меченный *гексахлорфлуоресцеином* (HEX).

Пример 5. Скрининг на мутанты ячменя с преждевременным стоп-кодоном в глицин-богатом РНК-связывающем белке (HvGR-RBP1)

Мутагенез ячменя сорта Paustian выполняли, как описано в примере 1, для достижения частоты мутаций примерно 1 мутация на 1,3 миллиона bp, что соответствует примерно 25 несинонимичным мутациям в кодирующих областях. Была подвергнута скринингу библиотека из 374 подпулов, полученных из пула 110000 особей M1, чтобы идентифицировать один положительный подпул. Затем подфракцию положительного подпула снова подвергали скринингу для идентификации мутантного зерна. Скрининг проводили с помощью ddPCR, как описано в WO 2018/001884, с использованием праймеров и зондов, разработанных специально для идентификации желаемого мутанта в локусе HvGR-RBP1.

Мутация соответствовала следующему сайту мутации: Нуклеотид G54>A (сDNA), соответствующий аминокислотной замене W18 (белок).

Мутагенез ячменя сорта Planet выполняли, как описано в примере 1, для достижения плотности мутаций около 1 мутации на 1,3 миллиона bp, что соответствует примерно 24 несинонимичным мутациям в кодирующих областях, что соответствует 99,94% генов, не содержащих несинонимичных мутаций.

При обращении к приведенному ниже примеру 6, оптимальный размер библиотеки составляет 520620, если желательна вероятность успеха $>33\%$ и выполняется поиск одной мутации. Таким образом, оптимальное количество способных к регенерации частей для начала составляет 99640 при расчете в соответствии с приведенным ниже примером 7.

Скрининг проводили с использованием протокола ddPCR с чувствительностью обнаружения 1 в 2000, ячмень имеет количество GECN, равное 5-6. Исходя из этого, максимальный размер подпула, рассчитанный, как показано в приведенном ниже примере 8, составил 363.

Таким образом, была подвергнута скринингу библиотека из 374 подпулов, полученных из пула 110000 растений, для идентификации одного положительного подпула. Таким образом, каждый подпул содержал зерна поколения M1 от приблизительно 294 растений M1. Затем часть положительного подпула снова подвергали скринингу для идентификации мутантного зерна. Скрининг проводили с помощью ddPCR, как описано в PCT/EP2017/065516. Желаемый мутант имеет нуклеотидную замену в положении 54 гуанина на аденин кодирующей последовательности HvGR-RBP1, которая доступна под номером HORVU.MOREX.r2.6HG0491380.1 в базе данных IPK Barley Blast Server Международного консорциума по секвенированию ячменя (29 сентября 2019 г.), что вызывает аминокислотную замену в положении 18, что приводит к преждевременному стоп-кодону в белковой последовательности, которая доступна под номером HORVU.MOREX.r2.6HG0491380.1 на сервере IPK Barley Blast Server Международного консорциума по секвенированию ячменя (29 сентября 2019 г.). Для идентификации мутанта разрабатывали специфический для мишени прямой праймер (CGCTGCTTCGTCGG), специфический для мишени обратный праймер (GAGCTAAAAGCAGCCTC), специфический для мутанта зонд для обнаружения (CTGTCCTGAAACACCGA), меченный 6-карбоксифлуоресцеином (FAM), и специфический для эталона зонд для обнаружения (CTGTCCTGGAACACCG), меченный гексахлорфлуоресцеином (HEX).

Пример 6. Определение оптимального размера библиотеки

Чтобы повысить вероятность успеха в идентификации интересующего мутанта, можно учитывать различные факторы. Прежде всего, размер библиотеки может быть оптимизирован. Оптимальный размер библиотеки (OLS) зависит от ряда факторов. Если желаемая мутация представляет собой мутацию остановки трансляции, в качестве

мишенной могут быть выбраны несколько потенциальных кодонов. Это означает, что библиотека меньшего размера может привести к более высокой вероятности успеха. Если желаемая мутация представляет собой замену конкретного нуклеотида или кодон для аминокислотной замены, то только одно событие, иногда несколько событий, дадут желаемый результат, и для повышения вероятности успеха предпочтительно иметь библиотеку большего размера.

Как правило, оптимальный размер библиотеки предпочтительно отражает общее количество мутантных геномов, которые необходимо подвергнуть скринингу, чтобы идентифицировать хотя бы один желаемый мутант. Вероятность идентификации конкретной мутации в определенной библиотеке мутантных геномов можно рассчитать с помощью частоты мутации (M_f). Если одновременно проводится скрининг нескольких мутаций (например, если имеется несколько потенциальных стоп-кодонов, которые одинаково желательны), то вероятность увеличивается. Количество подвергнутых скринингу мутаций также обозначается в настоящем документе как n .

Таким образом, вероятность идентифицировать хотя бы одну желаемую мутацию с использованием библиотеки с размером библиотеки LS и частоты мутаций M_f можно рассчитать следующим образом:

$$PS = 1 - M_f^{LS}$$

Например, в библиотеке семян M1 из 1000000 мутагенизированных растений рапса (каждое с частотой мутаций 1/500 kbp (тысяч пар нуклеотидов) существует вероятность, равная 0,86, идентифицировать по меньшей мере один конкретный нуклеотид в целевом локусе, как рассчитано ниже:

$$PS = 1 - (499999/500000)^{1000000} = 0,86$$

Эту формулу также можно использовать для определения оптимального размера библиотеки, необходимого для достижения определенной вероятности. Оптимальный размер библиотеки равен максимальному размеру библиотеки, который в приведенном выше примере равен 1000000 уникальных геномов.

Таким образом, оптимальный размер библиотеки можно рассчитать следующим образом:

$$\text{Оптимальный размер библиотеки (OLS)} = \frac{\log \left(1 - \frac{PS}{100} \right)}{\log \left(1 - (1 - (M_f \times n)) \right)}$$

где

OLS представляет собой оптимальный размер библиотеки;

PS представляет собой вероятность успеха в %;

Mf представляет собой частоту мутаций и

n представляет собой количество подвергнутых скринингу мутаций.

Например, по меньшей мере одна замена конкретного нуклеотида в библиотеке растений, каждого с частотой мутаций $1/500$ kbp, может быть идентифицирована с вероятностью 95%, если библиотека содержит 1497864 мутантных генома.

$$\text{Оптимальный размер библиотеки (OLS)} = \frac{\log(1 - \frac{95}{100})}{\log(1 - (1 - (1 - (1/500000 \times 1))))} = 1.497.864$$

Частота мутаций может быть определена с использованием общего количества подвергнутых скринингу нуклеотидов в геноме и общего количества генетических вариантов, идентифицированных из числа подвергнутых скринингу нуклеотидов. Таким образом, частоту мутаций можно рассчитать следующим образом:

$$Mf = Nnc/Nna$$

где

Nnc представляет собой общее количество нуклеотидных замен;

Nna представляет собой количество подвергнутых анализу нуклеотидов.

Например, если проанализировать 1000000 нуклеотидов в мутантном геноме и немутантном геноме, и идентифицировать 2 нуклеотида, отличающихся друг от друга, то частота мутаций составит одну на 500000.

$$Mf = 2/1000000 = 1/500000$$

Пример 7. Определение оптимального количества способных к регенерации частей (Np), для начала

Оптимальный размер библиотеки может быть определен, как описано в примере 1.

Затем может быть определено оптимальное количество способных к регенерации частей (ONp), подлежащих случайному мутагенезу для способов по изобретению. Другими словами, ONp равно оптимальному количеству способных к регенерации частей (например, семян), которые необходимо мутировать для достижения заданной вероятности успеха. Оптимальное ONp можно рассчитать, используя оптимальный размер библиотеки, количество клеток зародышевой линии ($GECN$) и скорость прорастания материала следующим образом:

$$ONp = ((OLS \times 100)/Gr)/GECN, \text{ где}$$

OLS представляет собой оптимальный размер библиотеки, рассчитанный, как описано в примере 6;

G_r представляет собой ожидаемую скорость прорастания в % после случайного мутагенеза; и

GECN представляет собой количество клеток зародышевой линии.

Для ячменя уровень сбора урожая и скорость прорастания являются примерно одинаковыми, и поэтому скорость прорастания может быть использована вместо скорости сбора для расчета ON_p.

Например, если предположить, что средняя скорость прорастания составляет 95%, и количество GECN равно 2, необходимо мутировать 526315 семян, чтобы создать библиотеку из 1000000 геномов.

$$\text{ON}_p = ((1000000 \times 100)/95)/2 = 526314$$

Пример 8.

Определение количества и размера подпулов

Для наиболее эффективного скрининга больших библиотек в настоящем изобретении предлагается многостадийный способ, включающий по меньшей мере стадии разделения исходного пула на подпулы, стадию скрининга для идентификации подпулов, содержащих способную(ые) к регенерации часть(и), содержащую мутацию, на наличие которой проводили скрининг, с последующей стадией идентификации конкретных способных к регенерации частей, содержащих мутацию, на наличие которой проводили скрининг, в пределах идентифицированного подпула. Способ по изобретению позволяет проводить скрининг очень большого количества способных к регенерации частей в разумные сроки и с разумными трудозатратами.

Одним из основных ограничивающих факторов в отношении определения размера подпулов является предел обнаружения используемого способа скрининга. Если чувствительность обнаружения бесконечна, то общий размер библиотеки является самым большим возможным подпулом. Однако даже в этом сценарии будет выгоднее разбить общую библиотеку (общий пул) на подпулы, в ином случае каждый скрининг общего пула будет приводить к положительному результату и, как правило, будет очень сложно и трудоемко идентифицировать отдельные способные к регенерации части, ответственные за положительный результат.

Если чувствительность обнаружения составляет 1 геном в Y геномах, то максимальный размер подпула в принципе равен Y. В качестве примера, если

чувствительность обнаружения составляет 1 в 2000 геномах, то максимальный размер подпула в принципе представляет собой подпул, содержащий 2000 различных геномов. Если общее количество различных геномов в подпуле является таким же, как чувствительность обнаружения, то каждый геном должен быть представлен примерно в равных количествах в указанном подпуле. Это означает, что образцы предпочтительно собирать таким образом, чтобы экстракция объединенной ДНК приводила к равному вкладу ДНК каждого генотипа. Поместив всю способную к регенерации часть любого заданного растения в один и тот же подпул, подпулы будут содержать примерно одинаковое количество способной к регенерации части каждого растения.

Подпул, размер которого меньше максимального размера подпула (определяемого на основе чувствительности обнаружения метода скрининга), имеет то преимущество, что он может быть менее однородным, что значительно упрощает подготовку подпулов. Например, если создан пул из 200 особей, каждый из которых внесет 10 семян, общий пул будет состоять из 2000 семян. Чувствительность обнаружения 1/2000 позволит 1999 семенам быть дикого типа и 1 мутировать для его идентификации. Соответственно, с помощью этого пула можно было бы собрать на 90% меньше материала от одной особи, при этом собирая такое же или большее количество семян для оставшихся 199 особей, при этом имея возможность обнаружить мутант в конечном пуле.

Таким образом, подпул должен иметь достаточный размер для конкретной чувствительности обнаружения. Окончательное количество подпулов может быть определено размером общего пула (т.е. N_p), деленным на количество растений в подпуле, исходя из чувствительности обнаружения и степени гибкости, необходимой для неоднородного пула ДНК.

Способы по изобретению, как правило, включают стадию выращивания способной к регенерации части в зрелые растения и разделения способной к регенерации части от указанных зрелых растений на подпулы таким образом, чтобы вся способная к регенерации часть от одного зрелого растения помещалась в один и тот же подпул. Подготовка подпулов таким образом гарантирует, что подпулы содержат примерно одинаковое количество способных к регенерации частей каждого генотипа. Если растения содержат несколько клеток зародышевой линии на растение, каждое зрелое растение представляет способную к регенерации часть более чем одного другого генотипа.

Таким образом, максимальное количество зрелых растений, обеспечивающих способную к регенерации часть заданного подпула (далее именуемое в настоящем документе максимальным размером подпула (SP_m)) может быть рассчитано следующим образом:

$$SP_m = Y/GECN$$

где

чувствительность обнаружения составляет 1 в Y; и

GECN представляет собой количество клеток зародышевой линии.

Например, если чувствительность обнаружения равна 1 в 2000, и если предположить, что GECN равно 5, максимальный размер подпула составляет

$$400 = 2000/5$$

Если количество способных к регенерации частей, подвергнутых случайному мутагенезу, равно 526315, и максимальный размер подпула равен 400, то пул способных к регенерации частей M1 следует разделить на по меньшей мере $526315/400=1315$ подпулов, где каждый подпул содержит способные к регенерации части от 400 растений. Может быть предпочтительным уменьшить размер подпула и, соответственно, фактический размер подпула (N_p) может быть меньше, чем необязательный размер подпула (ON_p).

Пример 9. Определение размера оптимального количества способных к регенерации частей (N_p), для начала, и оптимального размера подпула

Существует множество специфических для растений факторов, которые могут ограничивать размер пула для скрининга. Этими факторами могут являться:

- стоимость хранения пула определенного размера;
- легкость, с которой отдельное растение или его способная к регенерации часть может быть идентифицирована из пула определенного размера;
- скорость, с которой одно растение может быть идентифицировано из определенного размера пула;
- стоимость идентификации одного растения.

Все эти факторы являются высоко вариабельными и растение-специфическими. Например, размножение отдельных растений некоторых культур обходится дороже, чем скрининг, для других растений наоборот. Человек, знакомый с конкретной культурой, сможет оценить все факторы и определить удобный размер общего пула.

Этот размер является максимальным размером пула по практическим, а не по техническим причинам.

Например, большие растения (например, банановые растения или деревья) или растения с большими плодами (например, арбузы) занимают много места и, таким образом, фактором, ограничивающим скорость, является общее количество растений, которые можно вырастить. Соответственно, если растение представляет собой арбуз, даже несмотря на то, что количество семян для достижения оптимального размера библиотеки может быть большим, для способа можно использовать меньшее количество.

Если растение выращивают в поле, обычно можно выращивать большое количество растений с низкими затратами. Таким образом, количество размножаемых семян не является ограничивающим фактором.

Заметный ограничивающий фактор относится к количеству зерен, которые можно собрать с одного растения. Если растение производит только 10 зерен, а количество клеток зародышевой линии равно 5, то каждая клетка зародышевой линии будет представлена только 2 зёрнами. В самоопыляемой культуре только 75% зерен одной зародышевой линии будут нести желаемую мутацию и; в культуре с перекрестным опылением только 50%. Если на каком-то этапе необходимо разделить пул, то это число становится важным.

Пример 10. Подготовка образца ДНК

После создания подпулов способ по изобретению включает стадию подготовки образца ДНК. Как правило, образец ДНК получают путем деления всего подпула на случайное количество фракций, а затем получения ДНК из всей фракции. В зависимости от растения каждый подпул будет содержать некоторое количество способных к регенерации частей с одним и тем же генотипом. Количество будет зависеть от общего количества способных к регенерации частей на взрослое растение (RPr), а также от числа GECN. Если количество GECN равно 1, подпул будет содержать способные к регенерации части RPr с тем же генотипом.

Оптимальный размер фракции представляет собой размер фракции, где - теоретически - каждый генотип представлен одной способной к регенерации частью. Таким образом, необязательный размер фракции может быть рассчитан следующим образом:

$$\text{Оптимальный размер фракции (\%)} = (\text{GECN} * 100) / (\text{RPr})$$

Например, следующее применимо к сорту ячменя, который в среднем содержит 20 семян на растение, и который имеет количество GECN, равное 5. Таким образом, каждое растение ячменя M1 будет содержать 20 семян каждого генотипа. Подпулы содержат все семена зрелых растений ячменя. Таким образом, в принципе, случайная фракция оптимального размера фракции (=25%) семян ячменя заданного подпула будет содержать 1 семя, представляющее каждый генотип.

$$\text{Оптимальный размер фракции (\%)} = (5 * 100) / (20) = 25\%$$

Пример 11. Низкая скорость мутагенеза

«Низкая скорость мутагенеза» варьируется от культуры к культуре. В частности, плоидность культуры играет важную роль в том, насколько серьезно культура реагирует на обработку конкретной мутацией (Tsai et al 2013). Это также можно наблюдать по количеству допустимых мутаций (частоте мутаций) в гексаплоидном, тетраплоидном или диплоидном геноме (Kugowska et al 2013). Как поясняется выше, частота мутаций в настоящем документе представлена как $Mf = 1/(\text{количество нуклеотидов на мутацию})$.

Культура	Общее название	Плоидность	Общепринятое число нуклеотидов на мутацию в M2
Arabidopsis thaliana	Резушка Таля	диплоидное	300000
Triticum durum	пшеница твердая	тетраплоидное	40000
Triticum aestivum	пшеница	гексаплоидное	24000

Количество нуклеотидов на мутацию у растений M2 обычно на 25% ниже, чем количество нуклеотидов на мутацию у соответствующих растений M1, из-за самоопыления, которое обычно приводит к потере 25% всех гетерозиготных мутаций. Чтобы определить низкие и высокие скорости мутагенеза, авторы сравнивают количество генов, свободных от каких-либо кодирующих мутаций у растений M1. Количество генов, свободных от каких-либо кодирующих мутаций, основано на количестве генов в заданном геноме, частоте мутаций и средней длине растительного

белка (Tiessen et al 2012). В диплоидных культурах несколько сотен генов могут содержать кодирующие мутации после обработки высокими, общепринятыми дозами мутаций. Напротив, несколько тысяч генов обычно содержат кодирующие мутации в традиционно мутагенизированном гексаплоидном геноме. С целью создания библиотеки с низким уровнем мутагенеза низкая скорость мутагенеза в диплоидной культуре сопоставима с <0,1% генов с несинонимичными мутациями в поколении M1. У полиплоидных культур, таких как гексаплоидная пшеница, низкой частотой мутаций считается <1% генов с несинонимичными мутациями в поколении M1.

Процент генов, свободных от несинонимичных мутаций (GFoNSM), можно рассчитать по следующей формуле:

$$GFoNSM = Gn \times Gl \times Mf \times Nsnm / 100$$

где Gn представляет собой количество генов, Gl представляет собой среднюю длину гена, Mf представляет собой частоту мутаций и Nsnm представляет собой среднее число несинонимичных мутаций, вызванных мутагенезом, в %. В этом примере Nsnm составляет 66%.

					Общепринятый (высокий) мутагенез			Низкий мутагенез	
Культура	Общее название	Плюидность	Кол-во генов	Средняя длина генов	Нуклеотидов на мутацию M1 (25% меньше M2)	Нуклеотидов на мутацию M2	Мутации в кодирующих областях в M1	% генов, не имеющих несинонимичных мутаций в M1	% генов, не имеющих несинонимичных мутаций в M1
Arabidopsis thaliana	Резушка Таля	Диплоидное	27655	1176	225000	300000	95	99,7	>99,9
Triticum durum	Твердая пшеница	Тетраплоидное	36434	1176	30000	40000	943	97,4	>99
Triticum aestivum	Пшеница	Гексаплоидное	107891	1176	18000	24000	4652	95,7	>99

Пример 12. Разработка популяции мутагенизированных Brassica napus и скрининг на мутанты со специфической мутацией в гене синапальдегиддегидрогеназы/кониферальдегиддегидрогеназы (REF1)

Оптимальный размер библиотеки и оптимальное количество растений

Ген REF1 принимает участие в биосинтезе синапина, основного антипитательного соединения в семенах рапса (Emrani et al., 2015). Авторы стремились идентифицировать нокаут REF1 в библиотеке слегка мутагенизированных растений Brassica napus (т.е. с низкой скоростью мутагенеза). В общепринятой мутагенизированной библиотеке этой тетраплоидной культуры >3,5% всех генов

содержат кодирующие мутации. В слегка мутагенизированной библиотеке менее 1% всех генов содержат мутацию в кодирующей области. Мы использовали подход секвенирования GBS, чтобы определить количество мутаций и количество источников мутаций в нашей популяции. Коммерчески доступный метод секвенирования GBS был использован для секвенирования 12 растений, полученных от мутировавшей особи *Brassica napus*. Для анализа данных мы использовали следующие фильтры:

- удаление из анализа локусов, которые не охвачены по меньшей мере 8 считываниями последовательностей во всех анализируемых образцах;
- удаление локусов с изменениями из анализа, которые имеют дополнительные изменения в пределах 50 пар оснований, как интерпретировано на эталонном геноме;
- удаление образцов из анализа, которые определены как выбросы;
- игнорирование изменений локусов при анализе, когда локус показывает изменения в любом количестве контрольных образцов;
- игнорирование изменений локусов при анализе для локусов, которые не показывают изменений в 25% или более прочтений последовательности;
- игнорирование изменений локусов при анализе для локусов, которые показывают изменения в группе образцов, а также в контрольных образцах;
- игнорирование изменений локусов при анализе для локусов, которые не показывают изменений более чем в одном образце в группе образцов.

Всего в анализ было включено 6018138 локусов. В среднем на одно растение было идентифицировано 65,75 мутаций, что соответствует средней частоте мутаций, равной 1 мутации в 86607 пар оснований. Было идентифицировано 4 источника мутаций, или генетически эффективных клеток, в этой группе растений, т.е. количество GECN равно 4. Это число можно рассчитать следующим образом. Все уникальные мутации, идентифицированные во всем наборе из 12 образцов, представляют собой мутационные события, произошедшие на родительском растении, из которого произошли все образцы. Сумма уникальных мутаций в наборе образцов соответствует сумме идентифицированных наследственных мутаций, которые присутствовали у родительского растения в исследованных областях генома (Sum_{par}).

Если все мутации произошли в одном наследуемом источнике в родительском растении, то во всех образцах будут идентифицированы одни и те же мутации, а отношение суммы уникальных мутаций во всем наборе образцов к среднему количеству мутаций в одном образце (avg_{sam}) будет равно 1. Если предположить, что

все 12 образцов в наборе образцов происходят из разных источников мутаций, это отношение будет равно 12, поскольку все идентифицированные мутации будут уникальными. Затем количество различных источников мутаций в родительских растениях можно вывести как отношение между суммой уникальных родительских мутаций и средним числом мутаций в одном образце.

$$\text{Источники мутаций} = \text{Sum}_{\text{par}} / \text{Avg}_{\text{sam}}$$

$$\text{Источники мутаций} = 272 / (65,8 = 4,12)$$

Было идентифицировано 272 уникальные мутации в 12 образцах со средним числом мутаций на образец 65,8, что соответствует номеру источника мутации (или GECN) 4,14.

Эта информация может быть использована для определения оптимального размера библиотеки, который дает 95%-ную вероятность идентификации по меньшей мере одной нуклеотидной замены, приводящей к преждевременной остановке трансляции в REF1.

$$OLS = \frac{\log(1 - (\frac{PS}{100}))}{\log(1 - (1 - (1 - (M^f \times n))))}$$

$$OLS = \frac{\log(1 - (\frac{95}{100}))}{\log(1 - (1 - (1 - (1/86607 \times 4))))} = 129,724$$

Чтобы с вероятностью 95% идентифицировать один из двух стоп-кодонов в REF1, требуется библиотека из 129724 особей. Чтобы рассчитать, сколько способных к регенерации частей требуется для получения библиотеки из 129724 генотипов, можно использовать следующую формулу:

$$\text{ONp} = \frac{\frac{OLS \times 100}{Hr}}{GECN}$$

$$\text{ONp} = \frac{\frac{129724 \times 100}{17}}{4} = 190,771$$

При выращивании рапса, как правило, высевается намного больше растений, чем собирается, и, таким образом, урожай рапса довольно низок и составляет приблизительно 17%. При уровне сбора 17% и числе GECN 4 190771 являются оптимальным количеством семян для мутагенеза для создания библиотеки, которая с вероятностью 95% идентифицирует хотя бы одну нуклеотидную замену, которая приводит к преждевременной остановке трансляции в REF1.

Создание библиотеки и идентификация мутантов

Получали мутагенизированную библиотеку *Brassica napus* cv. *Euphr*. Библиотека была организована с использованием метода объединения в пул и разделения, описанного в PCT/EP2017/065516. Пул из 200000 семян обрабатывали 0,25% EMS в течение 16 часов при комнатной температуре и размножали в поле, используя стандартные практики. Было собрано 34992 растения в пулах примерно из 300 растений. ДНК извлекали из 25% собранного материала и использовали для скрининга библиотеки.

Идентификацию мутантов проводили в соответствии с методом скрининга ddPCR, описанным в PCT/EP2017/065516. Праймеры и зонды были разработаны для идентификации конкретного мутанта в REF1-локусе. Желаемый мутант имеет нуклеотидную замену в положении 504 с гуанина на аденин кодирующей последовательности REF-1, которая доступна под номером доступа в GenBank NCBI: FN995990.1, которая вызывает аминокислотную замену в положении 168, что приводит к преждевременной остановке трансляции в белковой последовательности, которая доступна под номером доступа в GenBank NCBI: FN995990.1. Для идентификации мутанта разрабатывали специфический для мишени прямой праймер (TATTGGAGTGGTCGGTC), специфический для мишени обратный праймер (CACTTTCATGGCAAACATAAT), специфический для мутанта зонд для обнаружения (ATAATCCCTTGAAATTTTCCAAG), меченный 6-карбоксифлуоресцеином (FAM), и специфический для эталона зонд для обнаружения (ATAATCCCTTGGAATTTTCCAA), меченный *гексахлорфлуоресцеином* (HEX).

Пункты формулы изобретения

1. Способ идентификации мутантного растения предварительно определенного вида, несущего одну или более мутаций в представляющем интерес нуклеотиде(ах) [NOI(s)] в заданной целевой последовательности, при этом указанный способ включает следующие стадии:

a) обеспечение способных к регенерации частей N_p родительского растения указанного вида, где N_p представляет собой целое число >5000 , такое как >10000 , такое как >15000 , такое как >20000 , такое как >25000 , такое как >30000 ;

b) подвергание указанных способных к регенерации частей стадии мутагенеза, предпочтительно случайного мутагенеза, тем самым создавая пул способных к регенерации частей поколения M_0 , представляющих множество генотипов;

c) выращивание указанных способных к регенерации частей поколения M_0 в зрелые растения поколения M_1 и получение способных к регенерации частей поколения M_1 из указанных зрелых растений;

d) необязательное повторение предыдущей стадии X раз для получения растений поколения $M(1+X)$, содержащих способные к регенерации части поколения $M(1+X)$;

e) разделение способных к регенерации частей множества растений поколения M_1 или $M(1+X)$ на подпулы, причем каждый подпул содержит способные к регенерации части, представляющие множество генотипов, при этом все способные к регенерации части от одного данного зрелого растения помещают в один и тот же подпул;

f) получение образцов ДНК, таких как образцы gDNA или образцы cDNA, полученные из образцов мРНК, из каждого подпула;

g) идентификацию подпула(ов), включающего ДНК, содержащую мутацию(и), с использованием чувствительных средств обнаружения;

h) идентификацию способных к регенерации частей в указанном подпуле, содержащем указанную мутацию, тем самым идентифицируя указанное мутантное растение,

при этом указанная стадия мутагенеза приводит к скорости мутагенеза от 1 до 30 несинонимичных мутаций на клетку способной к регенерации части поколения M_1 .

2. Способ идентификации мутантного растения заранее определенного вида, несущего одну или более мутаций в представляющем интерес нуклеотиде(ах) [NOI(s)] в заданной целевой последовательности, при этом указанный способ включает следующие стадии:

a) обеспечение способных к регенерации частей N_p родительского растения указанного вида, где N_p представляет собой целое число, равное по меньшей мере 5000, такое как по меньшей мере 10000, такое как по меньшей мере 15000, такое как по меньшей мере 20000, такое как по меньшей мере 25000, такое как по меньшей мере 30000;

b) подвергание указанных способных к регенерации частей стадии мутагенеза, предпочтительно случайного мутагенеза, приводящего к низкой скорости мутагенеза, тем самым создавая пул способных к регенерации частей поколения M_0 , представляющих множество генотипов;

c) выращивание указанных способных к регенерации частей поколения M_0 в зрелые растения поколения M_1 и получение способных к регенерации частей поколения M_1 из указанных зрелых растений;

d) необязательное повторение предыдущей стадии X раз для получения растений поколения $M(1+X)$, включающих способные к регенерации части поколения $M(1+X)$;

e) разделение способных к регенерации частей множества растений поколения M_1 или $M(1+X)$ на подпулы, причем каждый подпул содержит способные к регенерации части, представляющие множество генотипов, при этом все способные к регенерации части от одного данного зрелого растения помещают в один и тот же подпул;

f) получение образцов ДНК, таких как образцы gDNA или образцы cDNA, полученные из образцов мРНК, из каждого подпула;

g) идентификацию подпула(ов), включающего ДНК, содержащую мутацию(и), с использованием чувствительных средств обнаружения;

h) идентификацию способных к регенерации частей в указанном подпуле, содержащем указанную мутацию, тем самым идентифицируя указанное мутантное растение,

где низкая скорость мутагенеза у диплоидных растений заключается в том, что в среднем по меньшей мере 99,8%, предпочтительно по меньшей мере 99,9% всех генов не содержат несинонимичные мутации, а у растений с более высокой плоидностью в

среднем по меньшей мере 98,0%, предпочтительно по меньшей мере 99,0% всех генов не содержат несинонимичные мутации в способной к регенерации части поколения M1, предпочтительно где способная к регенерации часть представляет собой вегетативную ткань, семя, зерно, зародышевую клетку в семени, зерне или зародыше указанного растения, пыльцу или зародыш, предпочтительно способная к регенерации часть представляет собой семя.

3. Способ по любому из предыдущих пунктов, где стадию е) выполняют с помощью способа, включающего следующие стадии:

i. определение оптимального размера подпула, при этом максимальный размер подпула (SPm) $SPm = Y/GECN$;

ii. разделение указанных способных к регенерации частей на подпулы, где каждый подпул содержит способные к регенерации части от некоторого количества зрелых растений, при этом указанное количество находится в диапазоне от $0,5 * SPm$ до $1,1 * SPm$, при этом все способные к регенерации части от одного данного зрелого растения помещают в один и тот же подпул;

при этом чувствительность обнаружения чувствительных средств обнаружения составляет 1 в Y, и при этом растения имеют генетически эффективное число клеток GECN.

4. Способ идентификации растения, или его регенеративных частей, предварительно определенного вида растения, имеющего генетически эффективное число клеток (GECN), где растение несет по меньшей мере одну из n заданных мутаций в представляющем интерес нуклеотиде(ах) [NOI(s)] в целевой последовательности, причем в указанном способе используют средства обнаружения для обнаружения последовательностей нуклеиновых кислот с чувствительностью обнаружения 1 в Y, включающий следующие стадии:

a) обеспечение Np способных к регенерации частей растений указанного вида, где Np представляет собой целое число >1000 , например, по меньшей мере 5000, такое как по меньшей мере 10000, такое как по меньшей мере 15000, такое как по меньшей мере 20000, такое как по меньшей мере 25000, такое как по меньшей мере 30 000,

b) подвергание указанных способных к регенерации частей стадии мутагенеза, предпочтительно случайного мутагенеза, тем самым создавая пул

способных к регенерации частей поколения M_0 , представляющих множество генотипов;

с) выращивание указанных способных к регенерации частей поколения M_0 в зрелые растения и получение способных к регенерации частей поколения M_1 из указанных зрелых растений;

d) необязательно повторение предыдущей стадии X раз для получения растений, содержащих способные к регенерации части поколения $M(1+X)$;

e) разделение указанных способных к регенерации частей поколения M_1 или $M(1+X)$ на подпулы с помощью способа, включающего следующие стадии:

i) определение оптимального размера подпулы, при этом максимальный размер подпулы (SPm) $SPm = Y/GECN$;

ii) разделение указанных способных к регенерации частей на подпулы, при этом каждый подпул содержит способные к регенерации части от некоторого количества зрелых растений, причем указанное количество находится в диапазоне от $0,5*SPm$ до $1,1*SPm$, при этом все способные к регенерации части от одного данного зрелого растения помещают в один и тот же подпул;

f) получение образца ДНК, такого как образец gDNA или образец cDNA, из каждого подпула;

g) идентификацию подпула(ов), включающего способные к регенерации части, содержащие мутацию(и), с использованием указанных средств детекции;

h) идентификацию способных к регенерации частей в указанном подпуле, содержащем указанную мутацию, тем самым идентифицируя указанное мутантное растение.

5. Способ по пунктам 3-4, где средства обнаружения для обнаружения последовательностей нуклеиновых кислот являются чувствительными средствами обнаружения.

6. Способ по любому из предшествующих пунктов, где каждый подпул содержит способную к регенерации часть зрелых растений в диапазоне от $0,5*SPm$ до SPm .

7. Способ по любому из предшествующих пунктов, где каждый подпул содержит способную к регенерации часть зрелых растений в диапазоне от $0,7 * SPm$ до SPm .

8. Способ по любому из предшествующих пунктов, где Np находится в диапазоне от $0,7 * ONp$ до $1,3 * ONp$, где $ONp = ((OLS \times 100) / 4) / GECN$; и где Hr представляет собой ожидаемый средний уровень сбора указанных видов растений после мутагенеза в %; и

OLS представляет собой оптимальный размер библиотеки, где

$$OLS = \frac{\log \left(1 - \frac{PS}{100} \right)}{\log \left(1 - (1 - (1 - (Mf \times n))) \right)}$$

где PS представляет собой вероятность успеха в %; и

Mf представляет собой частоту мутаций; и

n представляет собой количество мутаций, на наличие которых проводили скрининг; и

где стадия б) включает стадию случайного мутагенеза, приводящую к частоте мутагенеза Mf .

9. Способ по п. 8, где уровень сбора представляет собой скорость прорастания указанных видов растений после мутагенеза в %.

10. Способ по любому из предшествующих пунктов, где Np находится в диапазоне от $0,8 * ONp$ до $1,2 * ONp$.

11. Способ по любому из предшествующих пунктов, где Np находится в диапазоне от $0,9 * ONp$ до $1,1 * ONp$.

12. Способ по любому из предшествующих пунктов, где способные к регенерации части представляют собой семена.

13. Способ по любому из предшествующих пунктов, где вероятность успеха составляет по меньшей мере 85%, предпочтительно по меньшей мере 90%, более предпочтительно по меньшей мере 95%, такую как по меньшей мере 98%.

14. Способ по любому из предшествующих пунктов, где по меньшей мере 30000 способных к регенерации частей обеспечивают на стадии а), такое как по меньшей мере 40000 способных к регенерации частей, такое как по меньшей мере 50000 способных к регенерации частей, такое как по меньшей мере 60000 способных к регенерации частей, такое как по меньшей мере 70000 способных к регенерации частей, такое как по меньшей мере 80000 способных к регенерации частей, такое как по меньшей мере 90000 способных к регенерации частей, такое как по меньшей мере 100000 способных к регенерации частей, такое как по меньшей мере 200000 способных к регенерации частей, такое как по меньшей мере 300000 способных к регенерации частей, такое как по меньшей мере 400000 способных к регенерации частей, такое как по меньшей мере 500000 способных к регенерации частей, такое как по меньшей мере 600000 способных к регенерации частей, такое как по меньшей мере 700000 способных к регенерации частей, такое как по меньшей мере 800000 способных к регенерации частей, такое как по меньшей мере 900000 способных к регенерации частей, такое как по меньшей мере 10^6 способных к регенерации частей, $2 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $3 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $4 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $5 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $6 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $7 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $8 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $9 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере 10^7 способных к регенерации частей, или более.

15. Способ по любому из предшествующих пунктов, где скорость мутагенеза составляет не более 1 мутации на каждые 1000 генов, такую как не более 1 мутации на каждые 5000 генов, такую как не более 1 мутации на каждые 10000 генов, такую как не более 1 мутации на каждые 20000 генов, такую как не более 1 мутации на каждые 30000 генов.

16. Способ по любому из предшествующих пунктов, где указанный мутагенез приводит к не более чем 25 несинонимичным мутациям в кодирующих областях, например не более чем к 20 несинонимичным мутациям в кодирующих областях, например не более чем к 15 несинонимичным мутациям в кодирующих областях, например не более чем к 10 несинонимичным мутациям в кодирующих областях.

областях, например не более чем к 5 несинонимичным мутациям в кодирующих областях, например 1 немолчающей мутации в кодирующей области у растений поколения M1.

17. Способ по любому из предшествующих пунктов, где растение представляет собой диплоидное растение, и указанная низкая скорость мутагенеза представляет собой частоту мутаций, равную 1 по меньшей мере на 500000, например 1 по меньшей мере на 750000, например 1 по меньшей мере на 1000000.

18. Способ по любому из предшествующих пунктов, где растение представляет собой тетраплоидное растение, и указанная низкая скорость мутагенеза представляет собой частоту мутаций, равную 1 по меньшей мере на 70000, например 1 по меньшей мере на 100000, например 1 по меньшей мере на 300000.

19. Способ по любому из предшествующих пунктов, где растение представляет собой гексаплоидное растение, и указанная низкая скорость мутагенеза представляет собой частоту мутаций, равную 1 по меньшей мере на 40000, например 1 по меньшей мере на 80000, например 1 по меньшей мере на 200000.

20. Способ по любому из предшествующих пунктов, где мутагенез представляет собой химический мутагенез или мутагенез, индуцированный облучением.

21. Способ по любому из предшествующих пунктов, где мутагенез представляет собой химический мутагенез, осуществляемый с помощью мутагенизирующего химического вещества, такого как мутагенизирующее химическое вещество, выбранное из группы, состоящей из азида натрия (NaN_3), этиленимина, гидроксиламина (NH_2OH), N-метил-N'-нитро-N-нитрозогуанидина (MNNG), диазومتана, метилнитрозогуанидина (MNNG) и этилметансульфоната (EMS), и алкилирующие агенты, такие как сульфонаты, например этилметансульфонат (EMS), диэтилсульфонат (DES) или N-этил-N-нитрозомочевина (ENU), или такие как сернистые иприты, например этил-2-хлорэтилсульфид, или такие как азотистые иприты, например 2-хлорэтилдиметиламин, или такие как эпоксиды, например этиленоксид.

22. Способ по любому из предшествующих пунктов, где стадию случайного мутагенеза осуществляют с помощью азида натрия, предпочтительно при концентрации 1 мМ или менее в течение 2 часов или менее, такой как 0,9 мМ или менее, такой как 0,8 мМ или менее, такой как 0,7 мМ или менее, такой как 0,6 мМ или менее, такой как 0,5 мМ или менее, такой как 0,4 мМ или менее, такой как 0,3 мМ или менее.

23. Способ по любому из пп. 1-21, в котором стадию случайного мутагенеза осуществляют с помощью EMS, предпочтительно при концентрации 4% или менее в течение 2 часов или менее, такой как 3% или менее, такой как 2% или менее, такой как 1,5% или менее, такой как 1% или менее, такой как 0,75% или менее, такой как 0,5% или менее, такой как 0,4% или менее, такой как 0,3% или менее, такой как 0,2% или менее, такой как 0,1% или менее, такой как 0,1% в течение 24 часов или менее.

24. Способ по любому из предшествующих пунктов, в котором мутагенез осуществляют в таких условиях, что по меньшей мере 90% способных к регенерации частей выживают на стадии мутагенеза, например, по меньшей мере 91% способных к регенерации частей, например, по меньшей мере 92% способных к регенерации частей, например, по меньшей мере 93% способных к регенерации частей, например, по меньшей мере 94% способных к регенерации частей, например, по меньшей мере 95% способных к регенерации частей, например, по меньшей мере 96% способных к регенерации частей, например, по меньшей мере 97% способных к регенерации частей, например, по меньшей мере 98% способных к регенерации частей, или например, по меньшей мере 99% способных к регенерации частей переживают стадию мутагенеза.

25. Способ по любому из предшествующих пунктов, где указанное мутантное растение имеет тот же фенотип, что и родительское растение, за исключением фенотипического признака(ов), вызванного мутацией(ями) в целевой последовательности.

26. Способ по любому из предшествующих пунктов, где одно или несколько из внешнего вида, высоты, биомассы, урожайности, размера зерна, приспособляемости, формы листа и вкуса указанного мутантного растения являются такими же, как у

родительского растения, предпочтительно все из внешнего вида, высоты, биомассы, формы листа и вкуса указанного мутантного растения являются такими же, как у родительского растения.

27. Способ по любому из предшествующих пунктов, в котором урожайность (культур) указанного мутантного растения соответствует по меньшей мере 90%, например, по меньшей мере 91%, например, по меньшей мере 92%, например, по меньшей мере 93%, например, по меньшей мере 94%, например, по меньшей мере 95%, например, по меньшей мере 96%, например, по меньшей мере 97%, например, по меньшей мере 98%, например, по меньшей мере 99%, например, по меньшей мере 100%, например, по меньшей мере 105%, например, по меньшей мере 110%, например, по меньшей мере 125%, например, по меньшей мере 150%, например, по меньшей мере 175%, например, по меньшей мере 200% от общей урожайности (культур) материнского растения, или более.

28. Способ по любому из предшествующих пунктов, в котором качество зерна указанного мутантного растения является по существу таким же, как качество зерна исходного растения, или выше.

29. Способ по любому из предшествующих пунктов, в котором представляющие интерес мутации происходят с частотой, соответствующей 1 мутации на 1,5 миллиона растений.

30. Способ по любому из предшествующих пунктов, в котором по меньшей мере 5000 способных к регенерации частей, например, по меньшей мере 10000 способных к регенерации частей, например, по меньшей мере 15000 способных к регенерации частей, например, по меньшей мере 20000 способных к регенерации частей, например, по меньшей мере 25000 способных к регенерации частей, например, по меньшей мере 30000 способных к регенерации частей, например, по меньшей мере 40000 способных к регенерации частей, например, по меньшей мере 50000 способных к регенерации частей, например, по меньшей мере 60000 способных к регенерации частей, например, по меньшей мере 70000 способных к регенерации частей, например, по меньшей мере 80000 способных к регенерации частей, например, по меньшей мере 90000 способных к регенерации частей, например, по меньшей мере 100000 способных

к регенерации частей, например, по меньшей мере 200000 способных к регенерации частей, например, по меньшей мере 300000 способных к регенерации частей, например, по меньшей мере 400000 способных к регенерации частей, например, по меньшей мере 500000 способных к регенерации частей, например, по меньшей мере 600000 способных к регенерации частей, например, по меньшей мере 700000 способных к регенерации частей, например, по меньшей мере 800000 способных к регенерации частей, например, по меньшей мере 900000 способных к регенерации частей, например, по меньшей мере 10^6 способных к регенерации частей, такое как по меньшей мере 10^6 зрелых растений, такое как по меньшей мере $2 \cdot 10^6$ зрелых растений, такое как по меньшей мере $3 \cdot 10^6$ зрелых растений, такое как по меньшей мере $4 \cdot 10^6$ зрелых растений, такое как по меньшей мере $5 \cdot 10^6$ зрелых растений, такое как по меньшей мере $6 \cdot 10^6$ зрелых растений, такое как по меньшей мере $7 \cdot 10^6$ зрелых растений, такое как по меньшей мере $8 \cdot 10^6$ зрелых растений, такое как по меньшей мере $9 \cdot 10^6$ зрелых растений, такое как по меньшей мере 10^7 зрелых растений, получают на стадии с).

31. Способ по любому из предшествующих пунктов, в котором жизнеспособность способных к регенерации частей поколения M0 является такой же или выше, чем жизнеспособность способных к регенерации частей, обеспеченных на стадии а).

32. Способ по любому из предшествующих пунктов, в котором стерильность указанного мутантного растения составляет менее 30%.

33. Способ по любому из предшествующих пунктов, в котором растение представляет собой культурное растение или цветущее растение, такое как растение, принадлежащее к семейству Brassicaceae, семейству Fabaceae, семейству Poaceae, семейству Amaranthaceae, семейству Asteraceae, семейству Solanaceae, семейству Rubiaceae или семейству Orchidaceae.

34. Способ по любому из предшествующих пунктов, в котором растения выбирают из группы, состоящей из злаков, бобовых и *Brassicaceae*.

35. Способ по любому из предшествующих пунктов, в котором растение представляет собой ячмень.

36. Способ по любому из предшествующих пунктов, где вид растения имеет среднее количество способных к регенерации частей на растение RPr, и где стадия f) включает:

- i. случайное разделение каждого подпула на фракции, где первая фракция содержит в диапазоне от $(GECN*100)/(RPr)\%$ до $(GECN*500)/(RPr)\%$ способную к регенерации часть подпула; и где указанная первая фракция содержит не более 50% указанной способной к регенерации части;
- ii. получение образцов ДНК всей первой фракции каждого подпула.

37. Способ по любому из предшествующих пунктов, где стадия f) включает следующие стадии:

- разделение каждого подпула на фракции;
- получение образцов ДНК, таких как образцы gDNA или образцы cDNA, полученные из образцов мРНК, всей фракции каждого подпула.

38. Способ по любому из предшествующих пунктов, где стадию f) выполняют путем разделения каждого подпула по меньшей мере на 2 фракции, например, по меньшей мере на 3 фракции, например, по меньшей мере на 4 фракции, и получение образцов ДНК из всей фракции каждого подпула.

39. Способ по любому из предшествующих пунктов, где стадия f) включает получение фракции, содержащей в диапазоне от 10 до 50% способные к регенерации части каждого подпула, и получение образцов ДНК из всех фракций каждого подпула.

40. Способ по любому из предшествующих пунктов, где средства обнаружения или чувствительные средства обнаружения имеют чувствительность обнаружения 1, приходящуюся по меньшей мере на 1000, такую как чувствительность обнаружения 1, приходящуюся на диапазон от 1000 до 10000, например, чувствительность обнаружения 1, приходящуюся на диапазон от 1000 до 5000.

41. Способ по любому из предшествующих пунктов, где стадия g) для каждого подпула включает следующие стадии:

- выполнение одной или нескольких ПЦР-амплификаций, включающих образец ДНК из указанной фракции подпула, один набор или несколько наборов праймеров, при этом каждый набор фланкирует целевую последовательность, и реагенты для ПЦР, таким образом амплифицируя целевую последовательность(и);
- обнаружение продукта(ов) ПЦР-амплификации, содержащего целевую последовательность(и), содержащую одну мутацию или несколько мутаций в NOI(s).

42. Способ по п. 41, где одну или несколько ПЦР-амплификаций разделяют на множество пространственно разделенных компартментов.

43. Способ по п. 42, где указанные пространственно разделенные компартменты представляют собой капли, такие как капли водно-масляной эмульсии.

44. Способ по п. 43, где каждая капля имеет средний объем в диапазоне от 0,1 до 10 нл.

45. Способ по любому из пп. 42-44, где каждая ПЦР разделена на компартменты в диапазоне от 1000 до 100000 пространственно разделенных компартментов.

46. Способ по любому из пп. 41-45, где реагенты для ПЦР содержат один или несколько зондов для обнаружения мутаций, причем каждый зонд(ы) для обнаружения мутаций содержит(ат) олигонуклеотид, необязательно связанный с детектируемым средством, где олигонуклеотид является идентичным – или комплементарным – целевой последовательности, включая заданную мутацию NOI.

47. Способ по любому из пп. 41-46, где реагенты для ПЦР содержат один эталонный зонд или несколько эталонных зондов для обнаружения, где каждый эталонный зонд(ы) для обнаружения содержит(ат) олигонуклеотид, необязательно связанный с детектируемым средством, где олигонуклеотид является идентичным – или комплементарным – целевой последовательности, включая эталонный NOI.

48. Способ по п. 47, в котором зонд(ы) для обнаружения мутантов связан(ы) с флуорофором и гасителем, а эталонный зонд для обнаружения связан с другим флуорофором и гасителем.

49. Способ по любому из пунктов 47-48, в котором реагенты для ПЦР содержат по меньшей мере 2-кратный избыток зонда(ов) для обнаружения мутанта по сравнению с эталонным зондом(ами) для обнаружения.

50. Способ по любому из пунктов 40-49, в котором ПЦР представляет собой РНКазы Н-зависимую ПЦР.

51. Способ по любому из предшествующих пунктов, где мутагенез на стадии а) представляет собой случайный мутагенез.

52. Способ по любому из предшествующих пунктов, в котором 10^6 способных к регенерации частей поколения М0 с различными генотипами получают на стадии б), такое как по меньшей мере $2 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $3 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $4 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $5 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $6 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $7 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $8 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $9 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере 10^7 способных к регенерации частей, или более.

53. Способ по любому из предшествующих пунктов, отличающийся тем, что указанный способ включает стадию размножения организмов или их способных к регенерации частей в пуле, и при этом указанная стадия размножения может выполняться одновременно со стадией е) разделения организмов на подпулы или после нее.

54. Способ по любому из предшествующих пунктов, отличающийся тем, что способ включает стадию получения растений поколения М(1+Х), где все способные к

регенерации части всего потомства данного растения поколения M0 помещают в один и тот же подпул.

55. Способ по любому из предшествующих пунктов, в котором стадия идентификации способной к регенерации части в идентифицированном подпуле включает следующие стадии:

- i) выращивание способной к регенерации части в растения;
- ii) получение образца ДНК из каждого из указанных растений;
- iii) идентификацию растений, содержащих мутацию.

56. Способ по любому из предшествующих пунктов, в котором способные к регенерации части на стадии е) относятся к поколению M2.

57. Способ по любому из предшествующих пунктов, в котором стадия h) включает следующие стадии:

- обеспечение подпула, содержащего множество способных к регенерации частей растения, где указанный подпул содержит способные к регенерации части, содержащие одну мутацию или несколько мутаций NOI(s);
- разделение способных к регенерации частей указанного подпула на вторичные подпулы;
- получение образца из каждой способной к регенерации части вторичных подпулов - таким способом, чтобы способные к регенерации части оставались в достаточной степени неповрежденными для развития в растение - и объединение всех образцов из всех способных к регенерации частей каждого вторичного подпула;
- получение образца ДНК из указанных объединенных образцов;
- идентификация вторичного подпула(ов), включающего ДНК, содержащую мутацию(и).

58. Способ по п. 57, в котором указанный образец получают путем просверливания отверстия в способной к регенерации части, предпочтительно семени, с последующим сбором полученной муки.

59. Способ по любому из пунктов 57-58, дополнительно включающий следующие стадии:

- обеспечение идентифицированного вторичного подпула, содержащего способные к регенерации части, содержащие мутацию в NOI, где регенеративные части представляют собой семена;

- выращивание всех семян в указанном вторичном подпуле с обеспечением возможности прорастания и необязательно развитие растений из каждого семени;

- получение образца от каждого проросшего семени;

- тестирование указанного образца на наличие указанной мутации в NOI, тем самым выявляя растение, несущее мутацию в NOI;

- необязательное выращивание указанного растения до зрелости.

60. Способ по пункту 59, отличающийся тем, что способ дополнительно включает следующие стадии, выполняемые после стадии f):

- получение фракции каждого образца ДНК из каждого подпула;

- объединение множества фракций в суперпулы, тем самым получая суперпулы ДНК, содержащие образцы ДНК из множества подпулов, при этом ДНК из каждого подпула присутствует только в одном суперпуле;

- проведение множества ПЦР-амплификаций, каждая из которых включает суперпул образцов ДНК, при этом каждая ПЦР-амплификация включает множество разделенных на компартменты ПЦР-амплификаций, каждая из которых включает часть указанного образца ДНК, один или несколько наборов праймеров, при этом каждый набор фланкирует целевую последовательность, и реагенты для ПЦР, тем самым амплифицируя целевую последовательность(и);

- обнаружение продукта(ов) амплификации ПЦР, содержащего одну или несколько целевых последовательностей, содержащих мутацию(и) в NOI(s), тем самым выявляя суперпул(ы), содержащий указанную мутацию.

61. Способ по п. 60, в котором ПЦР-амплификацию проводят только на образцах ДНК из подпула(ов), содержащего одну из мутаций.

62. Способ по любому из пп. 60-61, в котором готовят от 5 до 100 суперпулов.

63. Способ по любому из пунктов 60-62, в котором ПЦР-амплификацию(и), включающую суперпул образцов ДНК, проводят с помощью способа, включающего следующие стадии:

- подготовка ПЦР-амплификации, включающей образец ДНК, один или несколько наборов праймеров, при этом каждый набор фланкирует целевую последовательность, и реагенты для ПЦР;
- разделение указанной ПЦР-амплификации(й) на множество пространственно разделенных компартментов, где каждый компартмент имеет средний объем в диапазоне от 0,1 до 10 пл;
- проведение ПЦР-амплификации;
- обнаружение продуктов ПЦР-амплификации.

64. Способ по любому из пунктов 60-63, в котором ПЦР представляет собой РНКазы Н-зависимую ПЦР.

65. Способ по любому из пп. 60-64, в котором каждая ПЦР разделена по меньшей мере на 1000000 компартментов, таких как пространственно разделенные компартменты.

66. Способ по любому из пп. 60-65, в котором указанные пространственно разделенные компартменты представляют собой капли, такие как капли водно-масляной эмульсии.

67. Способ по любому из предшествующих пунктов, в котором мутация представляет собой мутацию с ожидаемой потерей функции в заданном полипептиде, и ожидаемое количество мутаций, на наличие которых был проведен скрининг, составляет 5% от количества аминокислот указанного полипептида.

68. Способ по любому из предшествующих пунктов, в котором количество мутаций, на наличие которых был проведен скрининг, (n) равно 1.

69. Способ по любому из предшествующих пунктов, в котором мутация представляет собой замену одного нуклеотида.

70. Способ по любому из предшествующих пунктов, который представляет собой способ идентификации одной заданной мутации в целевой последовательности, и в способе используется только один набор праймеров, фланкирующих указанную целевую последовательность.

71. Способ по любому из пп. 41-70, в котором реагенты для ПЦР содержат только один зонд для обнаружения мутации.

72. Способ по любому из предшествующих пунктов, где способная к регенерации часть представляет собой вегетативную ткань, семя, зерно, клетку зародышевой линии в семени, зерне или зародыше указанного растения, пыльцу или зародыш, предпочтительно способная к регенерации часть представляет собой семя.

Формула изобретения

1. Способ идентификации мутантного растения заранее определенного вида, несущего одну или более мутаций в представляющем интерес нуклеотиде(ах) [NOI(s)] в заданной целевой последовательности, при этом указанный способ включает следующие стадии:

a) обеспечение способных к регенерации частей N_p родительского растения указанного вида, где N_p представляет собой целое число, равное по меньшей мере 5000, такое как по меньшей мере 10000, такое как по меньшей мере 15000, такое как по меньшей мере 20000, такое как по меньшей мере 25000, такое как по меньшей мере 30000;

b) подвергание указанных способных к регенерации частей стадии мутагенеза, предпочтительно случайного мутагенеза, приводящего к низкой скорости мутагенеза, тем самым создавая пул способных к регенерации частей поколения M_0 , представляющих множество генотипов;

c) выращивание указанных способных к регенерации частей поколения M_0 в зрелые растения поколения M_1 и получение способных к регенерации частей поколения M_1 из указанных зрелых растений;

d) необязательное повторение предыдущей стадии X раз для получения растений поколения $M(1+X)$, включающих способные к регенерации части поколения $M(1+X)$;

e) разделение способных к регенерации частей множества растений поколения M_1 или $M(1+X)$ на подпулы, причем каждый подпул содержит способные к регенерации части, представляющие множество генотипов, при этом все способные к регенерации части от одного данного зрелого растения помещают в один и тот же подпул;

f) получение образцов ДНК, таких как образцы gDNA или образцы cDNA, полученные из образцов мРНК, из каждого подпула;

g) идентификацию подпула(ов), включающего ДНК, содержащую мутацию(и), с использованием чувствительных средств обнаружения;

h) идентификацию способных к регенерации частей в указанном подпуле, содержащем указанную мутацию, тем самым идентифицируя указанное мутантное растение,

где низкая скорость мутагенеза у диплоидных растений заключается в том, что в среднем по меньшей мере 99,8%, предпочтительно по меньшей мере 99,9% всех генов не содержат несинонимичные мутации, а у растений с более высокой плоидностью в среднем по меньшей мере 98,0%, предпочтительно по меньшей мере 99,0% всех генов не содержат несинонимичные мутации в способной к регенерации части поколения M1, предпочтительно где способная к регенерации часть представляет собой вегетативную ткань, семя, зерно, зародышевую клетку в семени, зерне или зародыше указанного растения, пыльцу или зародыш, предпочтительно способная к регенерации часть представляет собой семя.

2. Способ по п. 1, где скорость мутагенеза составляет не более 1 мутации на каждые 1000 генов, такую как не более 1 мутации на каждые 5000 генов, такую как не более 1 мутации на каждые 10000 генов, такую как не более 1 мутации на каждые 20000 генов, такую как не более 1 мутации на каждые 30000 генов.

3. Способ по любому из предшествующих пунктов, где указанный мутагенез приводит к не более чем 25 несинонимичным мутациям в кодирующих областях, например не более чем к 20 несинонимичным мутациям в кодирующих областях, например не более чем к 15 несинонимичным мутациям в кодирующих областях, например не более чем к 10 несинонимичным мутациям в кодирующих областях, например не более чем к 5 несинонимичным мутациям в кодирующих областях, например 1 немолчающей мутации в кодирующей области у растений поколения M1.

4. Способ по любому из предшествующих пунктов, где мутагенез представляет собой мутагенез, индуцированный облучением или химический мутагенез, осуществляемый с помощью мутагенизирующего химического вещества, такого как мутагенизирующее химическое вещество, выбранное из группы, состоящей из азиды натрия (NaN_3), этиленimina, гидроксилamina (NH_2OH), N-метил-N'-нитро-N-нитрозогуанидина (MNNG), диазометана, метилнитрозогуанидина (MNNG) и этилметансульфоната (EMS), и алкилирующие агенты, такие как сульфонаты, например этилметансульфонат (EMS), диэтилсульфонат (DES) или N-этил-N-нитрозомочевина (ENU), или такие как сернистые иприты, например этил-2-хлорэтилсульфид, или такие как азотистые иприты, например 2-хлорэтилдиметиламин,

или такие как эпоксиды, например этиленоксид, при этом предпочтительно стадию случайного мутагенеза осуществляют с помощью азиды натрия, предпочтительно при концентрации 1 мМ или менее в течение 2 часов или менее, такой как 0,9 мМ или менее, такой как 0,8 мМ или менее, такой как 0,7 мМ или менее, такой как 0,6 мМ или менее, такой как 0,5 мМ или менее, такой как 0,4 мМ или менее, такой как 0,3 мМ или менее, или стадию случайного мутагенеза осуществляют с помощью EMS, предпочтительно при концентрации 4% или менее в течение 2 часов или менее, такой как 3% или менее, такой как 2% или менее, такой как 1,5% или менее, такой как 1% или менее, такой как 0,75% или менее, такой как 0,5% или менее, такой как 0,4% или менее, такой как 0,3% или менее, такой как 0,2% или менее, такой как 0,1% или менее, такой как 0,1% в течение 24 часов или менее.

5. Способ по любому из предшествующих пунктов, где указанное мутантное растение имеет тот же фенотип, что и родительское растение, за исключением фенотипического признака(ов), вызванного мутацией(ями) в целевой последовательности.

6. Способ по любому из предшествующих пунктов, в котором представляющие интерес мутации происходят с частотой, соответствующей 1 мутации на 1,5 миллиона растений.

7. Способ по любому из предшествующих пунктов, где N_p находится в диапазоне от $0,7 * ON_p$ до $1,3 * ON_p$, где $ON_p = ((OLS \times 100) / 4) / GECN$; и где N_r представляет собой ожидаемый средний уровень сбора указанных видов растений после мутагенеза в %; и

OLS представляет собой оптимальный размер библиотеки, где

$$OLS = \frac{\log \left(1 - \frac{PS}{100} \right)}{\log \left(1 - (1 - (1 - (Mf \times n))) \right)}$$

где PS представляет собой вероятность успеха в %; и

Mf представляет собой частоту мутаций; и

n представляет собой количество мутаций, на наличие которых проводили скрининг; и

где стадия b) включает стадию случайного мутагенеза, приводящую к частоте мутагенеза Mf.

8. Способ по любому из предшествующих пунктов, где вероятность успеха составляет по меньшей мере 85%, предпочтительно по меньшей мере 90%, более предпочтительно по меньшей мере 95%, такую как по меньшей мере 98%.

9. Способ по любому из предшествующих пунктов, в котором по меньшей мере 5000 способных к регенерации частей, например, по меньшей мере 10000 способных к регенерации частей, например, по меньшей мере 15000 способных к регенерации частей, например, по меньшей мере 20000 способных к регенерации частей, например, по меньшей мере 25000 способных к регенерации частей, например, по меньшей мере 30000 способных к регенерации частей, например, по меньшей мере 40000 способных к регенерации частей, например, по меньшей мере 50000 способных к регенерации частей, например, по меньшей мере 60000 способных к регенерации частей, например, по меньшей мере 70000 способных к регенерации частей, например, по меньшей мере 80000 способных к регенерации частей, например, по меньшей мере 90000 способных к регенерации частей, например, по меньшей мере 100000 способных к регенерации частей, например, по меньшей мере 200000 способных к регенерации частей, например, по меньшей мере 300000 способных к регенерации частей, например, по меньшей мере 400000 способных к регенерации частей, например, по меньшей мере 500000 способных к регенерации частей, например, по меньшей мере 600000 способных к регенерации частей, например, по меньшей мере 700000 способных к регенерации частей, например, по меньшей мере 800000 способных к регенерации частей, например, по меньшей мере 900000 способных к регенерации частей, например, по меньшей мере 10^6 способных к регенерации частей, такое как по меньшей мере 10^6 зрелых растений, такое как по меньшей мере $2 \cdot 10^6$ зрелых растений, такое как по меньшей мере $3 \cdot 10^6$ зрелых растений, такое как по меньшей мере $4 \cdot 10^6$ зрелых растений, такое как по меньшей мере $5 \cdot 10^6$ зрелых растений, такое как по меньшей мере $6 \cdot 10^6$ зрелых растений, такое как по меньшей мере $7 \cdot 10^6$ зрелых растений, такое как по меньшей мере $8 \cdot 10^6$ зрелых растений, такое как по меньшей мере $9 \cdot 10^6$ зрелых растений, такое как по меньшей мере 10^7 зрелых растений, получают на стадии с).

10. Способ по любому из предшествующих пунктов, в котором растение представляет собой культурное растение или цветущее растение, такое как растение, принадлежащее к семейству Brassicaceae, семейству Fabaceae, семейству Poaceae,

семейству Amaranthaceae, семейству Asteraceae, семейству Solanaceae, семейству Rubiaceae или семейству Orchidaceae, предпочтительно где растение представляет собой ячмень.

11. Способ по любому из предыдущих пунктов, где стадию е) выполняют с помощью способа, включающего следующие стадии:

i. определение оптимального размера подпула, при этом максимальный размер подпула (SPm) $SPm = Y/GECN$;

ii. разделение указанных способных к регенерации частей на подпулы, где каждый подпул содержит способные к регенерации части от некоторого количества зрелых растений, при этом указанное количество находится в диапазоне от $0,5*SPm$ до $1,1*SPm$, при этом все способные к регенерации части от одного данного зрелого растения помещают в один и тот же подпул;

при этом чувствительность обнаружения чувствительных средств обнаружения составляет 1 в Y , и при этом растения имеют генетически эффективное число клеток $GECN$.

12. Способ по любому из предшествующих пунктов, где стадия f) включает следующие стадии:

- разделение каждого подпула на фракции;
- получение образцов ДНК, таких как образцы gDNA или образцы cDNA, полученные из образцов мРНК, всей фракции каждого подпула.

13. Способ по любому из предшествующих пунктов, где вид растения имеет среднее количество способных к регенерации частей на растение RPr , и где стадия f) включает:

i. случайное разделение каждого подпула на фракции, где первая фракция содержит в диапазоне от $(GECN*100)/(RPr)\%$ до $(GECN*500)/(RPr)\%$ способную к регенерации часть подпула; и где указанная первая фракция содержит не более 50% указанной способной к регенерации части;

ii. получение образцов ДНК всей первой фракции каждого подпула.

14. Способ по любому из предшествующих пунктов, где стадию f) выполняют путем разделения каждого подпула по меньшей мере на 2 фракции,

например, по меньшей мере на 3 фракции, например, по меньшей мере на 4 фракции, и получение образцов ДНК из всей фракции каждого подпула, и/или стадия f) включает получение фракции, содержащей в диапазоне от 10 до 50% способные к регенерации части каждого подпула, и получение образцов ДНК из всех фракций каждого подпула, и/или

где стадия g) для каждого подпула включает следующие стадии:

- выполнение одной или нескольких ПЦР-амплификаций, включающих образец ДНК из указанной фракции подпула, один набор или несколько наборов праймеров, при этом каждый набор фланкирует целевую последовательность, и реагенты для ПЦР, таким образом амплифицируя целевую последовательность(и);
- обнаружение продукта(ов) ПЦР-амплификации, содержащего целевую последовательность(и), содержащую одну мутацию или несколько мутаций в NOI(s).

15. Способ по любому из предшествующих пунктов, в котором 10^6 способных к регенерации частей поколения M0 с различными генотипами получают на стадии b), такое как по меньшей мере $2 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $3 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $4 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $5 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $6 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $7 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $8 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $9 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере 10^7 способных к регенерации частей, или более.

16. Способ по любому из предшествующих пунктов, отличающийся тем, что указанный способ включает стадию размножения организмов или их способных к регенерации частей в пуле, и при этом указанная стадия размножения может выполняться одновременно со стадией e) разделения организмов на подпулы или после нее, и/или где способные к регенерации части на стадии e) относятся к поколению M2.

17. Способ по любому из предшествующих пунктов, в котором стадия h) включает следующие стадии:

- обеспечение подпула, содержащего множество способных к регенерации частей растения, где указанный подпул содержит способные к регенерации части, содержащие одну мутацию или несколько мутаций NOI(s);
- разделение способных к регенерации частей указанного подпула на вторичные подпулы;
- получение образца из каждой способной к регенерации части вторичных подпулов - таким способом, чтобы способные к регенерации части оставались в достаточной степени неповрежденными для развития в растение - и объединение всех образцов из всех способных к регенерации частей каждого вторичного подпула;
- получение образца ДНК из указанных объединенных образцов;
- идентификация вторичного подпула(ов), включающего ДНК, содержащую мутацию(и).

18. Способ по п. 17, дополнительно включающий следующие стадии:

- обеспечение идентифицированного вторичного подпула, содержащего способные к регенерации части, содержащие мутацию в NOI, где регенеративные части представляют собой семена;
- выращивание всех семян в указанном вторичном подпуле с обеспечением возможности прорастания и необязательно развитие растений из каждого семени;
- получение образца от каждого проросшего семени;
- тестирование указанного образца на наличие указанной мутации в NOI, тем самым выявляя растение, несущее мутацию в NOI;
- необязательное выращивание указанного растения до зрелости.

19. Способ по пункту 18, отличающийся тем, что способ дополнительно включает следующие стадии, выполняемые после стадии f):

- получение фракции каждого образца ДНК из каждого подпула;
- объединение множества фракций в суперпулы, тем самым получая суперпулы ДНК, содержащие образцы ДНК из множества подпулов, при этом ДНК из каждого подпула присутствует только в одном суперпуле;
- проведение множества ПЦР-амплификаций, каждая из которых включает суперпул образцов ДНК, при этом каждая ПЦР-амплификация включает множество разделенных на компартменты ПЦР-амплификаций, каждая из которых включает часть указанного образца ДНК, один или несколько наборов праймеров, при этом каждый

набор фланкирует целевую последовательность, и реагенты для ПЦР, тем самым амплифицируя целевую последовательность(и);

- обнаружение продукта(ов) амплификации ПЦР, содержащего одну или несколько целевых последовательностей, содержащих мутацию(и) в NOI(s), тем самым выявляя суперпул(ы), содержащий указанную мутацию,

предпочтительно где ПЦР-амплификацию проводят только на образцах ДНК из подпула(ов), содержащего одну из мутаций.

Измененная формула изобретения

1. Способ идентификации мутантного растения заранее определенного вида, несущего одну или более мутаций в представляющем интерес нуклеотиде(ах) [NOI(s)] в заданной целевой последовательности, при этом указанный способ включает следующие стадии:

а) обеспечение способных к регенерации частей N_p родительского растения указанного вида, где N_p представляет собой целое число, равное по меньшей мере 5000, где N_p находится в диапазоне от $0,7 * ON_p$ до $1,3 * ON_p$, где $ON_p = ((OLS \times 100) / 4) / GECN$; и где H_g представляет собой ожидаемый средний уровень сбора указанных видов растений после мутагенеза в %; и

OLS представляет собой оптимальный размер библиотеки, где

$$OLS = \frac{\log \left(1 - \frac{PS}{100} \right)}{\log \left(1 - (1 - (1 - (Mf \times n))) \right)}$$

где PS представляет собой вероятность успеха в %; и

Mf представляет собой частоту мутаций; и

n представляет собой количество мутаций, на наличие которых проводили скрининг; и

Mf представляет собой частоту случайного мутагенеза стадии b):

b) подвергание указанных способных к регенерации частей стадии мутагенеза, предпочтительно случайного мутагенеза, приводящего к низкой скорости мутагенеза, тем самым создавая пул способных к регенерации частей поколения M_0 , представляющих множество генотипов;

c) выращивание указанных способных к регенерации частей поколения M_0 в зрелые растения поколения M_1 и получение способных к регенерации частей поколения M_1 из указанных зрелых растений;

d) необязательное повторение предыдущей стадии X раз для получения растений поколения $M(1+X)$, включающих способные к регенерации части поколения $M(1+X)$;

e) разделение способных к регенерации частей множества растений поколения M_1 или $M(1+X)$ на подпулы, причем каждый подпул содержит способные к регенерации части, представляющие множество генотипов, при этом все способные к

регенерации части от одного данного зрелого растения помещают в один и тот же подпул; и где данная стадия дополнительно включает:

i. определение оптимального размера подпула, при этом максимальный размер подпула (SPm) $SPm = Y/GECN$;

ii. разделение указанных способных к регенерации частей на подпулы, где каждый подпул содержит способные к регенерации части от некоторого количества зрелых растений, при этом указанное количество находится в диапазоне от $0,5*SPm$ до $1,1*SPm$, при этом все способные к регенерации части от одного данного зрелого растения помещают в один и тот же подпул;

при этом чувствительность обнаружения чувствительных средств обнаружения составляет 1 в Y , и при этом растения имеют генетически эффективное число клеток $GECN$.

f) получение образцов ДНК, таких как образцы gDNA или образцы cDNA, полученные из образцов мРНК, из каждого подпула; где стадия получения образцов ДНК включает:

- разделение каждого подпула на фракции;
- получение образцов ДНК, таких как образцы gDNA или образцы cDNA, полученные из образцов мРНК, всей фракции каждого подпула.

g) идентификацию подпула(ов), включающего ДНК, содержащую мутацию(и), с использованием чувствительных средств обнаружения;

h) идентификацию способных к регенерации частей в указанном подпуле, содержащем указанную мутацию, тем самым идентифицируя указанное мутантное растение,

где низкая скорость мутагенеза у диплоидных растений заключается в том, что в среднем по меньшей мере 99,8%, предпочтительно по меньшей мере 99,9% всех генов не содержат несинонимичные мутации, а у растений с более высокой плоидностью в среднем по меньшей мере 98,0%, предпочтительно по меньшей мере 99,0% всех генов не содержат несинонимичные мутации в способной к регенерации части поколения $M1$, предпочтительно где способная к регенерации часть представляет собой вегетативную ткань, семя, зерно, зародышевую клетку в семени, зерне или зародыше указанного растения, пыльцу или зародыш, предпочтительно способная к регенерации часть представляет собой семя.

2. Способ по п. 1, где указанная способная к регенерации часть представляет собой семя.

3. Способ по любому из предшествующих пунктов, где скорость мутагенеза составляет не более 1 мутации на каждые 1000 генов, такую как не более 1 мутации на каждые 5000 генов, такую как не более 1 мутации на каждые 10000 генов, такую как не более 1 мутации на каждые 20000 генов, такую как не более 1 мутации на каждые 30000 генов.

4. Способ по любому из предшествующих пунктов, где указанный мутагенез приводит к не более чем 25 несинонимичным мутациям в кодирующих областях, например не более чем к 20 несинонимичным мутациям в кодирующих областях, например не более чем к 15 несинонимичным мутациям в кодирующих областях, например не более чем к 10 несинонимичным мутациям в кодирующих областях, например не более чем к 5 несинонимичным мутациям в кодирующих областях, например 1 немолчающей мутации в кодирующей области у растений поколения M1.

5. Способ по любому из предшествующих пунктов, где мутагенез представляет собой мутагенез, индуцированный облучением или химический мутагенез, осуществляемый с помощью мутагенизирующего химического вещества, такого как мутагенизирующее химическое вещество, выбранное из группы, состоящей из азиды натрия (NaN_3), этиленimina, гидроксилamina (NH_2OH), N-метил-N'-нитро-N-нитрозогуанидина (MNNG), диазометана, метилнитрозогуанидина (MNNG) и этилметансульфоната (EMS), и алкилирующие агенты, такие как сульфонаты, например этилметансульфонат (EMS), диэтилсульфонат (DES) или N-этил-N-нитрозомочевина (ENU), или такие как сернистые иприты, например этил-2-хлорэтилсульфид, или такие как азотистые иприты, например 2-хлорэтилдиметиламин, или такие как эпоксиды, например этиленоксид, при этом предпочтительно стадию случайного мутагенеза осуществляют с помощью азиды натрия, предпочтительно при концентрации 1 мМ или менее в течение 2 часов или менее, такой как 0,9 мМ или менее, такой как 0,8 мМ или менее, такой как 0,7 мМ или менее, такой как 0,6 мМ или менее, такой как 0,5 мМ или менее, такой как 0,4 мМ или менее, такой как 0,3 мМ или менее, или стадию случайного мутагенеза осуществляют с помощью EMS,

предпочтительно при концентрации 4% или менее в течение 2 часов или менее, такой как 3% или менее, такой как 2% или менее, такой как 1,5% или менее, такой как 1% или менее, такой как 0,75% или менее, такой как 0,5% или менее, такой как 0,4% или менее, такой как 0,3% или менее, такой как 0,2% или менее, такой как 0,1% или менее, такой как 0,1% в течение 24 часов или менее.

6. Способ по любому из предшествующих пунктов, где указанное мутантное растение имеет тот же фенотип, что и родительское растение, за исключением фенотипического признака(ов), вызванного мутацией(ями) в целевой последовательности.

7. Способ по любому из предшествующих пунктов, в котором представляющие интерес мутации происходят с частотой, соответствующей 1 мутации на 1,5 миллиона растений.

8. Способ по любому из предшествующих пунктов, где N_p находится в диапазоне от $0,8 * ON_p$ до $1,2 * ON_p$.

9. Способ по любому из предшествующих пунктов, где вероятность успеха составляет по меньшей мере 85%, предпочтительно по меньшей мере 90%, более предпочтительно по меньшей мере 95%, такую как по меньшей мере 98%.

10. Способ по любому из предшествующих пунктов, в котором по меньшей мере 5000 способных к регенерации частей, например, по меньшей мере 10000 способных к регенерации частей, например, по меньшей мере 15000 способных к регенерации частей, например, по меньшей мере 20000 способных к регенерации частей, например, по меньшей мере 25000 способных к регенерации частей, например, по меньшей мере 30000 способных к регенерации частей, например, по меньшей мере 40000 способных к регенерации частей, например, по меньшей мере 50000 способных к регенерации частей, например, по меньшей мере 60000 способных к регенерации частей, например, по меньшей мере 70000 способных к регенерации частей, например, по меньшей мере 80000 способных к регенерации частей, например, по меньшей мере 90000 способных к регенерации частей, например, по меньшей мере 100000 способных к регенерации частей, например, по меньшей мере 200000 способных к регенерации

частей, например, по меньшей мере 300000 способных к регенерации частей, например, по меньшей мере 400000 способных к регенерации частей, например, по меньшей мере 500000 способных к регенерации частей, например, по меньшей мере 600000 способных к регенерации частей, например, по меньшей мере 700000 способных к регенерации частей, например, по меньшей мере 800000 способных к регенерации частей, например, по меньшей мере 900000 способных к регенерации частей, например, по меньшей мере 10^6 способных к регенерации частей, такое как по меньшей мере 10^6 зрелых растений, такое как по меньшей мере $2 \cdot 10^6$ зрелых растений, такое как по меньшей мере $3 \cdot 10^6$ зрелых растений, такое как по меньшей мере $4 \cdot 10^6$ зрелых растений, такое как по меньшей мере $5 \cdot 10^6$ зрелых растений, такое как по меньшей мере $6 \cdot 10^6$ зрелых растений, такое как по меньшей мере $7 \cdot 10^6$ зрелых растений, такое как по меньшей мере $8 \cdot 10^6$ зрелых растений, такое как по меньшей мере $9 \cdot 10^6$ зрелых растений, такое как по меньшей мере 10^7 зрелых растений, получают на стадии с).

11. Способ по любому из предшествующих пунктов, в котором растение представляет собой культурное растение или цветущее растение, такое как растение, принадлежащее к семейству Brassicaceae, семейству Fabaceae, семейству Poaceae, семейству Amaranthaceae, семейству Asteraceae, семейству Solanaceae, семейству Rubiaceae или семейству Orchidaceae, предпочтительно где растение представляет собой ячмень.

12. Способ по любому из предшествующих пунктов, где вид растения имеет среднее количество способных к регенерации частей на растение RPr, и где стадия f) включает:

i. случайное разделение каждого подпула на фракции, где первая фракция содержит в диапазоне от $(GECN \cdot 100)/(RPr)\%$ до $(GECN \cdot 500)/(RPr)\%$ способную к регенерации часть подпула; и где указанная первая фракция содержит не более 50% указанной способной к регенерации части;

ii. получение образцов ДНК всей первой фракции каждого подпула.

13. Способ по любому из предшествующих пунктов, где стадию f) выполняют путем разделения каждого подпула по меньшей мере на 2 фракции, например, по меньшей мере на 3 фракции, например, по меньшей мере на 4 фракции, и получение образцов ДНК из всей фракции каждого подпула, и/или стадия f) включает

получение фракции, содержащей в диапазоне от 10 до 50% способные к регенерации части каждого подпула, и получение образцов ДНК из всех фракций каждого подпула, и/или

где стадия g) для каждого подпула включает следующие стадии:

- выполнение одной или нескольких ПЦР-амплификаций, включающих образец ДНК из указанной фракции подпула, один набор или несколько наборов праймеров, при этом каждый набор фланкирует целевую последовательность, и реагенты для ПЦР, таким образом амплифицируя целевую последовательность(и);
- обнаружение продукта(ов) ПЦР-амплификации, содержащего целевую последовательность(и), содержащую одну мутацию или несколько мутаций в NOI(s).

14. Способ по любому из предшествующих пунктов, в котором 10^6 способных к регенерации частей поколения M0 с различными генотипами получают на стадии b), такое как по меньшей мере $2 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $3 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $4 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $5 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $6 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $7 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $8 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере $9 \cdot 10^6$ способных к регенерации частей, такое как по меньшей мере 10^7 способных к регенерации частей, или более.

15. Способ по любому из предшествующих пунктов, отличающийся тем, что указанный способ включает стадию размножения организмов или их способных к регенерации частей в пуле, и при этом указанная стадия размножения может выполняться одновременно со стадией e) разделения организмов на подпулы или после нее, и/или где способные к регенерации части на стадии e) относятся к поколению M2.

16. Способ по любому из предшествующих пунктов, в котором стадия h) включает следующие стадии:

- обеспечение подпула, содержащего множество способных к регенерации частей растения, где указанный подпул содержит способные к регенерации части, содержащие одну мутацию или несколько мутаций NOI(s);

- разделение способных к регенерации частей указанного подпула на вторичные подпулы;
- получение образца из каждой способной к регенерации части вторичных подпулов - таким способом, чтобы способные к регенерации части оставались в достаточной степени неповрежденными для развития в растение - и объединение всех образцов из всех способных к регенерации частей каждого вторичного подпула;
- получение образца ДНК из указанных объединенных образцов;
- идентификация вторичного подпула(ов), включающего ДНК, содержащую мутацию(и).

17. Способ по п. 16, дополнительно включающий следующие стадии:

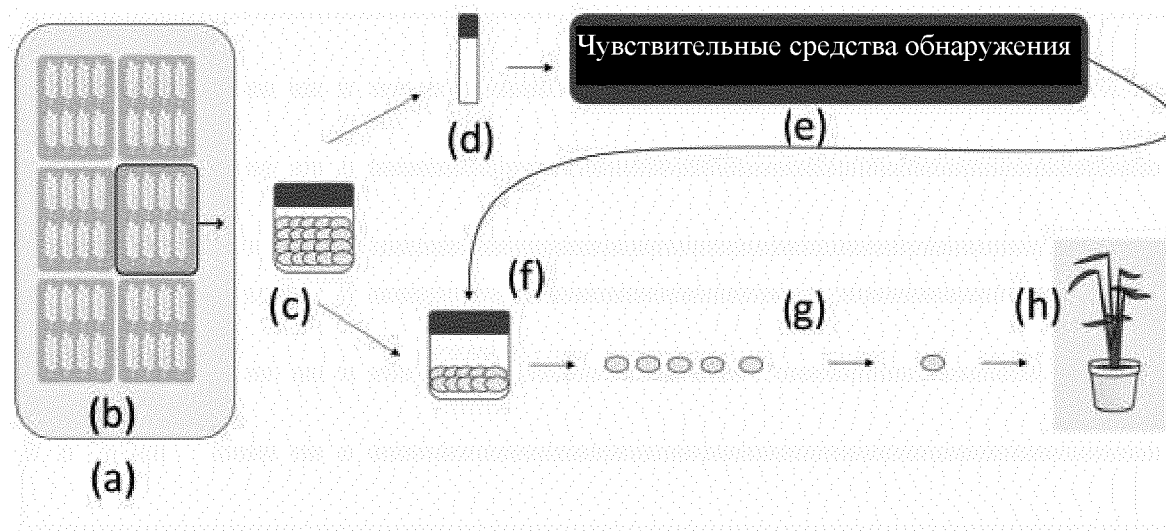
- обеспечение идентифицированного вторичного подпула, содержащего способные к регенерации части, содержащие мутацию в NOI, где регенеративные части представляют собой семена;
- выращивание всех семян в указанном вторичном подпуле с обеспечением возможности прорастания и необязательно развитие растений из каждого семени;
- получение образца от каждого проросшего семени;
- тестирование указанного образца на наличие указанной мутации в NOI, тем самым выявляя растение, несущее мутацию в NOI;
- необязательное выращивание указанного растения до зрелости.

18. Способ по пункту 17, отличающийся тем, что способ дополнительно включает следующие стадии, выполняемые после стадии f):

- получение фракции каждого образца ДНК из каждого подпула;
- объединение множества фракций в суперпулы, тем самым получая суперпулы ДНК, содержащие образцы ДНК из множества подпулов, при этом ДНК из каждого подпула присутствует только в одном суперпуле;
- проведение множества ПЦР-амплификаций, каждая из которых включает суперпул образцов ДНК, при этом каждая ПЦР-амплификация включает множество разделенных на компартменты ПЦР-амплификаций, каждая из которых включает часть указанного образца ДНК, один или несколько наборов праймеров, при этом каждый набор фланкирует целевую последовательность, и реагенты для ПЦР, тем самым амплифицируя целевую последовательность(и);

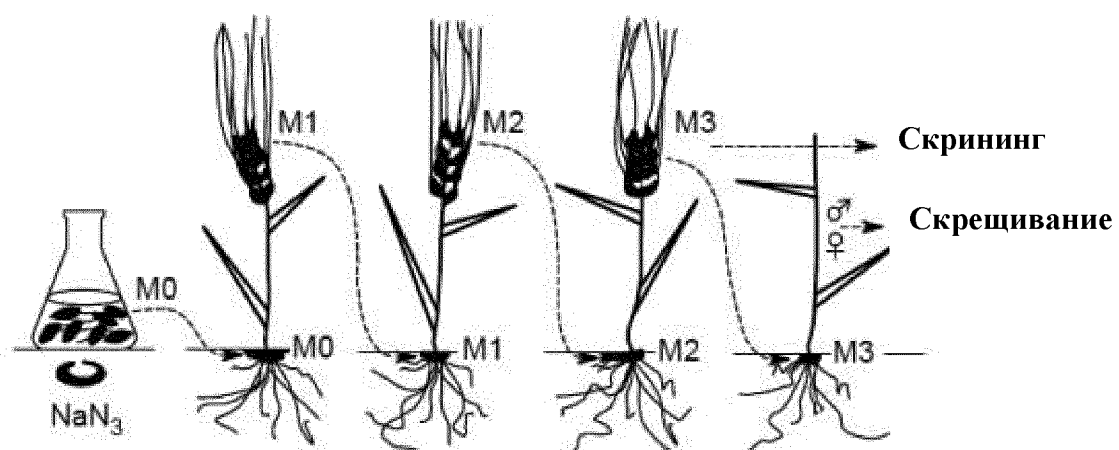
- обнаружение продукта(ов) амплификации ПЦР, содержащего одну или несколько целевых последовательностей, содержащих мутацию(и) в NOI(s), тем самым выявляя суперпул(ы), содержащий указанную мутацию,

предпочтительно где ПЦР-амплификацию проводят только на образцах ДНК из подпула(ов), содержащего одну из мутаций.



ФИГ. 1

ФИГ. 2



ФИГ. 3

