

(19)



**Евразийское
патентное
ведомство**

(21) **202291586** (13) **A1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2022.10.11

(22) Дата подачи заявки
2020.11.26

(51) Int. Cl. **C12Q 1/6876** (2018.01)
C12Q 1/68 (2018.01)
C12Q 1/6809 (2018.01)
C12Q 1/686 (2018.01)
G16B 30/00 (2019.01)
G16B 50/00 (2019.01)

(54) СПОСОБЫ И КОМПОЗИЦИИ ДЛЯ ОБЕСПЕЧЕНИЯ ИДЕНТИФИКАЦИИ И/ИЛИ ВОЗМОЖНОСТИ ОТСЛЕЖИВАНИЯ БИОЛОГИЧЕСКОГО МАТЕРИАЛА

(31) **62/940,587**

(32) **2019.11.26**

(33) **US**

(86) **PCT/CA2020/051622**

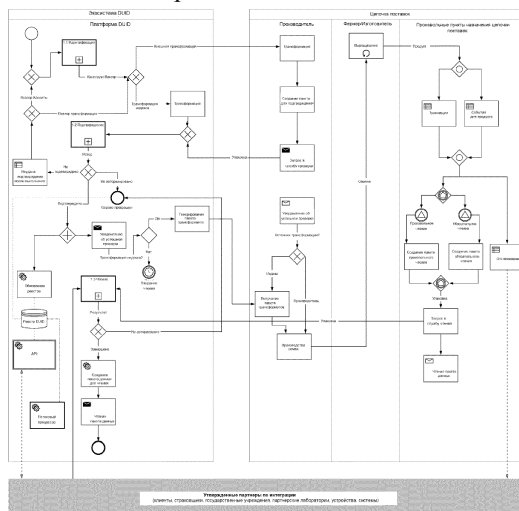
(87) **WO 2021/102579 2021.06.03**

(71) Заявитель:
ИНДЕКС БАЙОСИСТЕМС ИНК.
(CA)

(72) Изобретатель:
Борг Майкл, Фридберг Джереми Н.
(CA)

(74) Представитель:
Строкова О.В., Гизатуллин Ш.Ф.,
Гизатуллина Е.М., Прищепный С.В.,
Джермакян Р.В., Христофоров А.А.,
Угрюмов В.М., Костюшенкова М.Ю.
(RU)

(57) В настоящем документе предложены способы и композиции для обеспечения идентификации и/или отслеживания биологических материалов. В некоторых вариантах осуществления изобретения предложены способы, предусматривающие стадии: определения последовательности по меньшей мере одной последовательности уникального идентификатора в геномной ДНК биологического объекта; подтверждение идентификации биологического объекта путем проверки наличия последовательности уникального идентификатора в геномной ДНК и сравнения последовательности уникального идентификатора с базой данных для подтверждения ее уникальности; предоставление указания о соответствии получения биологического материала из биологического объекта; и ввод последовательности уникального идентификатора в запись базы данных, и связывание последовательности уникального идентификатора с информацией об идентификации и/или по отслеживанию; обеспечивая тем самым возможность отслеживания путем считывания последовательности уникального идентификатора и извлечения соответствующей записи из базы данных для получения информации об идентификации и/или по отслеживанию. Также предложены олигонуклеотиды, кассеты и композиции для обеспечения идентификации и/или отслеживания биологических материалов.



202291586
A1

202291586
A1

СПОСОБЫ И КОМПОЗИЦИИ ДЛЯ ОБЕСПЕЧЕНИЯ ИДЕНТИФИКАЦИИ И/ИЛИ ВОЗМОЖНОСТИ ОТСЛЕЖИВАНИЯ БИОЛОГИЧЕСКОГО МАТЕРИАЛА

Описание

Область техники, к которой относится настоящее изобретение

Настоящее изобретение в целом относится к идентификации и/или отслеживанию биологических материалов. Более конкретно, настоящее изобретение относится к способам и средствам для идентификации и/или отслеживания биологических материалов с использованием нуклеиновой кислоты.

Предшествующий уровень техники настоящего изобретения

Продовольственная система достигла беспрецедентного уровня эффективности распределения и объема производства. Эта эволюция принесла обществу большие преимущества в виде снижения затрат и разнообразия; однако остаются серьезные недостатки, которые подвергают риску здоровье населения, промышленность и инновации. Возможность отслеживания является одним из основных методов эффективного управления и решения этих проблем.

Ограничения существующих систем отслеживания пищевых продуктов и напитков в первую очередь проявляются в случаях контаминации. Когда происходят такие события, могут потребоваться месяцы, чтобы отследить произведенные продукты до источника их происхождения. Продукты, получаемые клонально, могут создать дополнительные проблемы для идентификации источника их происхождения, поскольку у них отсутствует генетическая изменчивость. Трансформированные и смешанные продукты также могут быть проблематичными для идентификации источника их происхождения, поскольку они требуют применения передовых методов отслеживания по всей цепочке поставок. Проблемы, связанные с возможностью быстрого и недорогого отслеживания этих продуктов, приводят к значительному риску для безопасности потребителей; ведут к значительным финансовым потерям для заинтересованных сторон; и наносят серьезный ущерб репутации производителей.

В 2015 году Всемирная организация здравоохранения (ВОЗ) завершила 10-летнюю инициативу по оценке глобального воздействия заболеваний пищевого происхождения. В рамках этой инициативы было установлено, что «... глобальное воздействие заболеваний пищевого происхождения ... в 2010 г. составило 33 (95% UI 25–46) миллионов DALY; при этом 40% воздействия заболеваний пищевого происхождения приходилось на детей в возрасте до 5 лет» (стр. 11). DALY – это количество лет жизни с поправкой на инвалидность. Этот показатель можно рассматривать как один потерянный год здоровой жизни. Оценки, сделанные в этом исследовании, были ограничены пробелами в исходных данных. Было отмечено, что для более точной оценки требуется улучшение эпиднадзора и лабораторных мощностей. Рабочей группой по установлению источника (SATF) были дополнительно определены потребности в наблюдениях.

SATF представляла одну из многочисленных специальных групп, задействованных в этой инициативе. Ее задача заключалась в оценке влияния конкретных источников на передачу болезни. На фиг. 1 (адаптировано из публикации ВОЗ, 2015 г., стр. 101) показаны основные источники. FERG, референтная группа для этого исследования, определила, что для целей исследования наиболее простой точкой для установления источника является конец цепочки передачи, то есть контакт с человеком. Эта упрощение является характерным ограничением для существующих методов отслеживания. FERG также отметила (стр. 100), что для управления рисками могут быть более подходящими другие точки установления источника,

например, первичное производство. и при этом FERG считает желательным осуществлять контроль по установлению масштабов источника.

Современные методы отслеживания продуктов в цепочке поставок пищевых продуктов и напитков обычно начинаются со сбора урожая на фермере или с производственного объекта. Продукты часто отслеживаются на уровне упаковки, поскольку упаковка содержит много единиц товара. Иногда на каждый товар наносится физический штрих-код. Глобальный номер предмета торговли (GTIN) и глобальный идентификационный номер (GLN) товара идеально связаны с продуктом. Для паллеты с множеством упаковок может быть создан серийный код транспортной упаковки (SSCC). Эти методы отслеживания обычно предписываются стандартом, и для свежих продуктов это часто стандарт GS1. По мере того, как паллеты проходят через цепочку поставок, указанные идентификаторы, содержащиеся в штрих-кодах, используют в сочетании с элементами основных данных (KDE), записанных для критических событий отслеживания (СТЕ). СТЕ может описывать передачу продукта от производителя к упаковщику/грузоотправителю. Существует широко используемый афоризм, который предполагает, что каждый участник цепочки поставок должен иметь возможность отслеживать продукт «на один шаг вперед и на один шаг назад». К сожалению, это требование во многих отношениях оказалось недостаточным.

Как только пищевой продукт попадает в точку продажи, он может быть переработан или смешан с другими продуктами от разных производителей, например, в случае фруктового салата. Часто, как только продукт отделяется от своей первоначальной упаковки или от идентификатора продукта, то зачастую невозможно отследить продукт обратно до его производителя. Как видно из недавней вспышки заболеваний, связанных с потреблением салата романа, следователям потребовалось более месяца, чтобы точно определить источник заражения, поскольку у них не было информации об источнике происхождения салата (FDA, 2019, стр. 1), хотя подавляющее большинство его производителей находится на юго-западе США. В результате Управление по санитарному надзору за качеством пищевых продуктов и медикаментов (FDA) призвало «... во всю цепочку поставок листовой зелени внедрить передовые методы отслеживания и самые современные технологии, чтобы обеспечить быстрый, точный и легкий доступ к основным ключевым элементам данных по цепи от фермы до прилавка, когда листовая зелень вовлечена в потенциальный отзыв товара или возможную вспышку заболевания». (FDA, 2019, стр. 8). Затраты, связанные с этой вспышкой, все еще не покрыты. Однако хорошо известны и другие случаи контаминаций.

Отзыв шпината в 2006 году был связан с пятью смертельными случаями и приблизительно 200 опасными для жизни заболеваниями в 26 штатах. Это нанесло финансовый ущерб приблизительно в 500 миллионов долларов (GS1, 2013, стр. 3). В более общем плане «... правительственные учреждения также выразили озабоченность по поводу последствий недавних отзывов пищевых продуктов для здоровья и финансовой сферы, поскольку заболевания пищевого происхождения поражают 48 миллионов человек в год и обходятся Соединенным Штатам каждый год в 152 миллиарда долларов в виде расходов на здравоохранение» (GS1, 2013, стр. 2). Установлено, что возможность отслеживания всей цепочки, что можно понимать как отслеживание продукта от семян до его продажи, снижает общее количество отзываемой продукции до 12%, как в случае отзыва кинзы от Frontera Produce. McKinsey показал, что повышение точности отзыва продукта на 25% может сэкономить отрасли свежих продуктов 250–275 миллионов долларов в год (GS1, 2013, стр. 10).

В отслеживании всей цепочки продукта отсутствовала эффективная форма идентификации на уровне единицы продукта, и не было гарантий в отношении источника его происхождения. Существующие методы идентификации на уровне единицы продукта обычно основаны на физической маркировке (лазером),

радиочастотном идентификаторе (RFID) и штриховом кодировании, т.е. на внешних физических идентификаторах. Существуют также проблемы масштабирования, связанные с этими методами. Каждой единице продукта требуется физический идентификатор, а его производство связано с затратами. Кроме того, существует риск ошибочного считывания и/или риск злонамеренного вмешательства, связанного с их использованием: например, наклейки иногда отваливаются или их удаляют.

Контаминация пищевых продуктов, такое как заражение *E. coli* и/или *Salmonella*, влияющее на пищевые продукты, представляет собой угрозу для здоровья населения, и поэтому крайне желательны быстрые действия по выявлению и устранению источника(ов) заражения. В этой области давно ощущается неудовлетворенная потребность в надежных, экономически эффективных и/или быстрых стратегиях повышения возможности отслеживания продуктов в системе снабжения продовольствием. Отслеживание биологических объектов и/или биологических материалов необходимо не только в сельском хозяйстве и пищевой промышленности, но оно также востребовано в самых разных отраслях и областях, связанных с биологическими объектами и/или биологическими материалами, содержащими их или полученными из них.

Поэтому необходимы альтернативные, дополнительные и/или улучшенные способы и/или композиции для обеспечения идентификации и/или отслеживания биологических объектов и/или биологических материалов.

Сущность настоящего изобретения

В настоящем документе предложены способы и композиции для обеспечения идентификации и/или отслеживания биологических материалов. В некоторых вариантах осуществления изобретения в способах, описанных в настоящем документе, может быть использована последовательность уникального идентификатора (также называемая здесь ДНК последовательностью уникального идентификатора), которую экзогенно вводят в геном биологического объекта, чтобы обеспечить идентификацию и/или возможность отслеживания биологического объекта и/или биологических материалов, содержащих биологический объект и/или биологические материалы, полученные из биологического объекта, и содержащих его геномную ДНК. В некоторых вариантах осуществления изобретения последовательность уникального идентификатора может быть из рандомизированного пула последовательностей. В некоторых вариантах осуществления изобретения может поддерживаться база данных, связывающая последовательности уникальных идентификаторов с соответствующей информацией об идентификации и/или по отслеживанию. Также в настоящем документе предложены олигонуклеотидные конструкторы и кассеты, содержащие одну или более последовательностей уникальных идентификаторов для использования при обеспечении идентификации и/или отслеживания биологических материалов. В некоторых вариантах осуществления изобретения олигонуклеотидные конструкторы и/или кассеты могут содержать специфические конфигурации последовательности праймера (праймеров) отжига, которые могут быть предназначены для амплификации последовательностей уникальных идентификаторов, секвенирования последовательностей уникальных идентификаторов или того и другого. В некоторых вариантах осуществления изобретения способы и композиции, описанные в настоящем документе, могут использоваться для обеспечения возможности отслеживания пищевых продуктов, и они могут обеспечивать быстрое реагирование и/или отзыв пищевых продуктов, например, в случае контаминации.

В одном варианте осуществления изобретения в настоящем документе предложен способ идентификации биологического материала, включающий:

получение или предоставление образца, содержащего геномную ДНК, из биологического материала;

амплификацию по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК из биологического материала и секвенирование ДНК последовательности уникального идентификатора; и

поиск ДНК последовательности уникального идентификатора в базе данных, и извлечение записи из базы данных, соответствующей ДНК последовательности уникального идентификатора, где запись в базе данных предоставляет информацию об идентификации и/или по отслеживанию биологического материала.

В другом варианте осуществления вышеуказанного способа биологический материал может включать материал растительного происхождения, материал на основе гриба, материал животного происхождения, материал на основе вируса или материал на основе бактерии.

В некоторых вариантах осуществления изобретения биологический материал может содержать материал на основе гриба. В некоторых вариантах осуществления изобретения биологический материал может содержать дрожжи. В некоторых вариантах осуществления дрожжи необязательно могут быть спорообразующими (т.е. биологический материал может содержать споры дрожжей). В некоторых вариантах осуществления изобретения дрожжи могут быть добавлены, смешаны или иным образом связаны с продуктом, для которого желательна идентификация и/или отслеживание, например, с пищевым ингредиентом или пищевым продуктом.

В другом варианте осуществления изобретения в настоящем документе предложен способ обеспечения возможности отслеживания биологического материала, включающий:

определение последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК биологического объекта;

подтверждение идентификации биологического объекта посредством: проверки наличия ДНК последовательности уникального идентификатора в геномной ДНК; и сравнение ДНК последовательности уникального идентификатора с базой данных, для подтверждения того, что ДНК последовательность уникального идентификатора еще не используется в базе данных;

предоставление указания о соответствии получения биологического материала из биологического объекта, где биологический материал содержит геномную ДНК из биологического объекта; и

ввод последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в запись базы данных, и связывание ДНК последовательности уникального идентификатора с информацией об идентификации и/или по отслеживанию биологического материала;

обеспечивая тем самым возможность отслеживания биологического материала путем считывания ДНК последовательности уникального идентификатора в биологическом материале, и извлечения соответствующей записи из базы данных, предоставляющей информацию об идентификации и/или по отслеживанию биологического материала.

В другом варианте осуществления вышеуказанного способа способ может дополнительно предусматривать вставку по меньшей мере одной ДНК последовательности уникального идентификатора в геномную ДНК биологического объекта или модификацию уже существующей последовательности идентификатора в геномной ДНК биологического объекта путем редактирования гена для создания ДНК последовательности уникального идентификатора в геномной ДНК биологического объекта, обеспечивая тем самым его идентификацию.

В еще одном варианте осуществления любого из вышеуказанного способа или способов, способ может дополнительно предусматривать предоставление по меньшей мере одной ДНК последовательности уникального идентификатора для вставки в геномную ДНК биологического объекта.

В еще одном варианте осуществления любого из вышеуказанного способа или способов биологический материал может содержать материал растительного происхождения, материал на основе гриба, материал животного происхождения, материал на основе вируса или материал на основе бактерии.

В еще одном варианте осуществления любого из вышеуказанного способа или способов биологический объект может включать клетку растения, клетку гриба, клетку животного, вирус или бактериальную клетку.

В другом варианте осуществления любого из вышеуказанного способа или способов биологический материал, биологический объект или оба могут включать материал на основе гриба или клетку гриба. В некоторых вариантах осуществления биологический материал, биологический объект или оба могут содержать дрожжи. В некоторых вариантах осуществления дрожжи необязательно могут быть спорообразующими (т.е. могут содержать споры дрожжей).

В еще одном варианте осуществления любого из вышеуказанного способа или способов получение биологического материала из биологического объекта может включать размножение биологического объекта.

В другом варианте осуществления любого из вышеуказанного способа или способов ДНК последовательность уникального идентификатора может быть из рандомизированного пула ДНК последовательностей уникального идентификатора.

В еще одном варианте осуществления любого из вышеуказанного способа или способов считывание ДНК последовательности уникального идентификатора в биологическом материале и поиск соответствующей записи в базе данных может предусматривать:

получение или предоставление образца, содержащего геномную ДНК, из биологического материала; амплификацию по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК из биологического материала, и секвенирование ДНК последовательности уникального идентификатора; и

сравнение ДНК последовательности уникального идентификатора с базой данных, и извлечение записи из базы данных, соответствующей ДНК последовательности уникального идентификатора, где запись в базе данных предоставляет информацию об идентификации и/или по отслеживанию биологического материала.

В еще одном варианте осуществления любого из вышеуказанного способа или способов ДНК последовательность уникального идентификатора может содержать уникальную последовательность нуклеотидов, встроенную в межгенную область геномной ДНК.

В еще одном варианте осуществления любого из вышеуказанного способа или способов ДНК последовательность уникального идентификатора может содержать последовательность размером до около 1500 нуклеотидов (нт) в длину; до около 1000 нт в длину; от около 200 до около 600 нт в длину; от около 200 нт до около 400 нт в длину; или от около 400 нт до около 600 нт в длину.

В другом варианте осуществления любого из вышеуказанного способа или способов ДНК последовательность уникального идентификатора может быть фланкирована одной или более последовательностями праймеров отжига для ПЦР-амплификации ДНК последовательности уникального идентификатора, секвенирования ДНК последовательности уникального идентификатора или для того и другого.

В еще одном варианте осуществления любого из вышеуказанного способа или способов биологический материал может включать пищевой продукт.

В еще одном варианте осуществления любого из вышеуказанного способа или способов информация об идентификации и/или по отслеживанию в записи в базе данных может содержать информацию о цепочке поставок биологического материала. В некоторых вариантах осуществления информация о цепочке поставок может содержать информацию о цепочке поставок пищевых продуктов, сельскохозяйственных продуктов, фармацевтических препаратов, розничных продуктов, текстильных продуктов, товаров широкого потребления, химических продуктов, или информацию о других элементах цепочки поставок продуктов, с которыми может быть связан биологический материал.

В другом варианте осуществления любого из вышеуказанного способа или способов информация об идентификации и/или по отслеживанию в записи в базе данных может содержать информацию об источнике происхождения биологического материала.

В еще одном варианте осуществления любого из вышеуказанного способа или способов информация об идентификации и/или по отслеживанию в записи в базе данных может включать информацию о производителе, регионе, партии, лоте, дате или другую соответствующую информацию о цепочке поставок или любые их комбинации.

В еще одном варианте осуществления любого из вышеуказанного способа или способов кассета может быть встроена в геномную ДНК, при этом кассета может содержать ДНК последовательность уникального идентификатора, фланкированную одной или более последовательностями праймеров отжига для ПЦР-амплификации ДНК последовательности уникального идентификатора, для секвенирования ДНК последовательности уникального идентификатора или для того и другого.

В другом варианте осуществления любого из вышеуказанного способа или способов ДНК последовательность уникального идентификатора может представлять собой случайную последовательность, полученную из рандомизированного пула последовательностей нуклеиновых кислот размером до около 1500 нуклеотидов (нт) в длину; до около 1000 нт в длину; от около 200 до около 600 нт в длину; от около 200 нт до около 400 нт в длину; или от около 400 нт до около 600 нт в длину.

В другом варианте осуществления изобретения в настоящем документе предложен олигонуклеотид, содержащий ДНК последовательность уникального идентификатора, фланкированную одной или более последовательностями праймеров отжига для ПЦР-амплификации ДНК последовательности уникального идентификатора, секвенирования ДНК последовательности уникального идентификатора или для того и другого.

В другом варианте осуществления вышеуказанный олигонуклеотид с ДНК последовательностью уникального идентификатора может содержать случайную последовательность размером до около 1500 нуклеотидов (нт) в длину; до около 1000 нт в длину; от около 200 до около 600 нт в длину; от около 200 нт до около 400 нт в длину; или от около 400 нт до около 600 нт в длину.

В другом варианте осуществления изобретения в настоящем документе предложена кассета, содержащая любой олигонуклеотид или олигонуклеотиды, описанные в настоящем документе.

В еще одном варианте осуществления изобретения в настоящем документе предложена клетка или вирус, содержащие любой олигонуклеотид или олигонуклеотиды, описанные в настоящем документе, или любую кассету или кассеты, описанные в настоящем документе, встроенные в геном клетки или вируса.

В другом варианте осуществления изобретения в настоящем документе предложена клетка или вирус, содержащие ДНК последовательность уникального идентификатора, встроенную в геном клетки или вируса.

В другом варианте осуществления любых вышеуказанных клеток или вирусов ДНК последовательность уникального идентификатора может быть встроена в межгенную область геномной ДНК клетки или вируса.

В еще одном варианте осуществления любых вышеуказанных клеток или вирусов клетка может быть клеткой растения, клеткой гриба, клеткой животного или клеткой бактерии.

В другом варианте осуществления изобретения клетка может представлять собой клетку гриба, такую как дрожжевая клетка.

В другом варианте осуществления изобретения в настоящем документе предложен набор, содержащий любой один или более из следующего:

- ДНК последовательность уникального идентификатора;
- рандомизированный пул ДНК последовательностей уникальных идентификаторов;
- любой из олигонуклеотида или олигонуклеотидов, описанных в настоящем документе;
- любую из кассеты или кассет, описанных в настоящем документе;
- одну или более пар праймеров для амплификации и/или секвенирования ДНК последовательности уникального идентификатора;
- буфер;
- полимеразу; или
- инструкции по выполнению любого из описанного здесь способа или способов.

В другом варианте осуществления изобретения в настоящем документе предложен способ идентификации биологического материала, включающий:

- получение на вычислительном устройстве ДНК последовательности уникального идентификатора (DUID), извлеченной из известного биологического материала;
- поиск в вычислительном устройстве базы данных DUID, в которой хранится множество DUID вместе с соответствующей информацией о биологическом материале, для совпадения с полученной DUID;
- если поиск в базе данных DUID не дает совпадения с полученной DUID, сохранение в базе данных DUID полученной DUID вместе с информацией о биологическом материале, связанной с известным биологическим материалом;
- после сохранения полученной DUID и информации, связанной с известным биологическим материалом в базе данных DUID, прием на вычислительном устройстве запроса DUID, извлеченной из неизвестного биологического материала;
- поиск в вычислительном устройстве базы данных DUID для совпадения с полученным запросом DUID; и
- если поиск DUID обеспечивает совпадение с полученным запросом DUID, возврат в ответ на полученный запрос DUID биологической информации, сохраненной в ассоциации с DUID, совпадающей с запросом DUID.

В другом варианте осуществления вышеуказанного способа поиск в базе данных DUID совпадения с полученной DUID может предусматривать:

- поиск в базе данных DUID точного совпадения с полученной DUID; и
- если точное совпадение не найдено, выполнение выравнивания/поиска идентичности для DUID, хранящихся в базе данных DUID, которые являются близким соответствием с полученной DUID.

В еще одном варианте осуществления любого из вышеуказанных способа или способов поиск в базе данных DUID совпадения с запросом DUID может предусматривать:

поиск в базе данных DUID точного совпадения с запросом DUID; и
если точное совпадение не найдено, выполнение выравнивания/поиска идентичности для DUID, хранящихся в базе данных DUID, которые близко соответствуют запросу DUID.

В еще одном варианте осуществления любого из вышеуказанного способа или способов, способ может дополнительно предусматривать:

если поиск обеспечивает близкое соответствие с DUID запроса, сохранение DUID запроса в ассоциации с DUID, который является близким соответствием DUID запроса.

В другом варианте осуществления в настоящем документе предложена вычислительная система для идентификации биологического материала, включающая:

блок обработки, способный выполнять инструкции; и

блок памяти, хранящий инструкции, которые при выполнении их блоком обработки конфигурируют вычислительную систему для выполнения любого из описанного здесь способа или способов.

В другом варианте осуществления изобретения в настоящем документе предложена машиночитаемая память, содержащая хранящиеся в ней инструкции, которые при выполнении их блоком обработки вычислительной системы конфигурируют систему для выполнения любого из описанного здесь способа или способов

В другом варианте осуществления изобретения в настоящем документе предложен способ идентификации биологического материала, включающий:

получение или предоставление образца, содержащего геномную ДНК, из биологического материала; амплификацию по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК из биологического материала, и секвенирование ДНК последовательности уникального идентификатора; и

декодирование или расшифровку информации об идентификации и/или по отслеживанию биологического материала, хранящегося в ДНК последовательности уникального идентификатора.

В другом варианте осуществления изобретения в настоящем документе предложен способ обеспечения возможности отслеживания биологического материала, включающий:

определение последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК биологического объекта;

подтверждение идентификации биологического объекта посредством: проверки наличия ДНК последовательности уникального идентификатора в геномной ДНК; и декодирование или расшифровку информации об идентификации и/или по отслеживанию, хранящейся в ДНК последовательности уникального идентификатора, для проверки ДНК последовательности уникального идентификатора; и

предоставление указания о соответствии получения биологического материала из биологического объекта, где биологический материал содержит геномную ДНК из биологического объекта;

обеспечивая тем самым возможность отслеживания биологического материала путем считывания ДНК последовательности уникального идентификатора в биологическом материале и декодирования или расшифровки информации, хранящейся в ДНК последовательности уникального идентификатора, предоставляя информацию об идентификации и/или по отслеживанию биологического материала.

В еще одном варианте осуществления изобретения в настоящем документе предложен способ идентификации биологического материала, включающий:

получение на вычислительном устройстве ДНК последовательности уникального идентификатора (DUID), извлеченной из неизвестного биологического материала; и

декодирование или расшифровку информации об идентификации и/или по отслеживанию неизвестного биологического материала, хранящейся в ДНК последовательности уникального идентификатора.

В другом варианте осуществления изобретения в настоящем документе предложена кассета, содержащая ДНК последовательность уникального идентификатора, где ДНК последовательность уникального идентификатора фланкирована по меньшей мере одной последовательностью 5'-праймера отжига и по меньшей мере одной последовательностью 3'-праймера отжига для амплификации ДНК последовательности уникального идентификатора, секвенирования ДНК последовательности уникального идентификатора или для того и другого.

В другом варианте осуществления вышеуказанной кассеты ДНК последовательность уникального идентификатора может быть фланкирована двумя последовательностями 5'-праймеров отжига и двумя последовательностями 3'-праймеров отжига, чтобы обеспечить амплификацию ДНК последовательности уникального идентификатора с помощью вложенной ПЦР.

В еще одном варианте осуществления любой вышеуказанной кассеты или кассет две последовательности 5'-праймеров отжига могут частично перекрываться; последовательности двух 3'-праймеров отжига могут частично перекрываться; или имеют место оба случая.

В еще одном варианте осуществления любой вышеуказанной кассеты или кассет кассета может дополнительно содержать последовательность праймера отжига для секвенирования, расположенную в положении 5' ДНК последовательности уникального идентификатора, для секвенирования ДНК последовательности уникального идентификатора.

В еще одном варианте осуществления любой вышеуказанной кассеты или кассет последовательность праймера отжига для секвенирования может быть расположена между двумя последовательностями 5'-праймеров отжига.

В другом варианте осуществления любой вышеуказанной кассеты или кассет последовательность праймеров отжига для секвенирования может по меньшей мере частично перекрываться с одной или с обеими последовательностями 5'-праймеров отжига.

В еще одном варианте осуществления любой вышеуказанной кассеты или кассет две последовательности 5'-праймеров отжига могут частично перекрываться, и по меньшей мере часть последовательности секвенирующих праймеров отжига может располагаться в месте перекрытия.

В другом варианте осуществления любой вышеуказанной кассеты или кассет последовательность кассеты может иметь размер до около 1500 нуклеотидов (нт) в длину; до около 1000 нт в длину; от около 200 до около 600 нт в длину; от около 200 нт до около 400 нт в длину; или от около 400 нт до около 600 нт в длину.

В еще одном варианте осуществления любой вышеуказанной кассеты или кассет последовательности праймеров отжига могут не встречаться в природе в геноме целевого биологического объекта.

В другом варианте осуществления изобретения в настоящем документе предложена композиция, содержащая множество из любой кассеты или кассет, как описано в настоящем документе, причем каждая кассета содержит вышеуказанные последовательности праймеров отжига, и каждая кассета содержит рандомизированную ДНК последовательность уникального идентификатора.

В еще одном варианте осуществления изобретения в настоящем документе предложена композиция, содержащая множество из любой кассеты или кассет, как описано в настоящем документе, причем каждая кассета содержит вышеуказанные последовательности праймеров отжига и вышеуказанную

последовательность праймера отжига для секвенирования, и каждая кассета содержит рандомизированную ДНК последовательность уникального идентификатора.

В еще одном варианте осуществления изобретения в настоящем документе предложен способ обеспечения возможности отслеживания биологического материала, причем указанный способ включает:

встраивание по меньшей мере одной ДНК последовательности уникального идентификатора в геномную ДНК биологического объекта для использования при получении биологического материала.

В другом варианте осуществления вышеуказанного способа ДНК последовательность уникального идентификатора может быть встроена в виде любой кассеты или кассет, как описано в настоящем документе.

В другом варианте осуществления любого из вышеуказанных способов способ может дополнительно предусматривать стадию определения последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК биологического объекта.

В другом варианте осуществления любого из вышеуказанного способа или способов, способ может дополнительно предусматривать стадию проверки подтверждения подлинности идентификации биологического объекта путем: проверки наличия ДНК последовательности уникального идентификатора в геномной ДНК; и сравнения ДНК последовательности уникального идентификатора с базой данных для подтверждения того, что ДНК последовательность уникального идентификатора еще не используется в базе данных.

В еще одном варианте осуществления любого из вышеуказанного способа или способов, способ может дополнительно предусматривать стадию:

получения биологического материала из биологического объекта, где биологический материал содержит геномную ДНК из биологического объекта; и/или

предоставления указания о соответствии получения биологического материала из биологического объекта, где биологический материал содержит геномную ДНК из биологического объекта.

В еще одном варианте осуществления любого из вышеуказанных способа или способов, способ может дополнительно предусматривать стадию ввода последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в запись базы данных, и связывания ДНК последовательности уникального идентификатора с информацией об идентификации и/или по отслеживанию биологического объекта и/или биологического материала

В еще одном варианте осуществления любого из вышеуказанного способа или способов, способ может дополнительно предусматривать стадию:

обеспечения возможности отслеживания биологического объекта и/или биологического материала путем считывания ДНК последовательности уникального идентификатора в биологическом объекте и/или биологическом материале и извлечения соответствующей записи из базы данных, предоставляющей информацию об идентификации и/или по отслеживанию биологического объекта и/или биологического материала.

В другом варианте осуществления изобретения в настоящем документе предложена плазида или вектор экспрессии, содержащий любой олигонуклеотид или олигонуклеотиды, или кассету, или кассеты, как описано в настоящем документе.

В еще одном варианте осуществления изобретения в настоящем документе предложен способ обеспечения возможности отслеживания представляющего интерес продукта, где указанный способ предусматривает:

получение или предоставление образца представляющего интерес продукта, где образец содержит геномную ДНК из части биологического материала, смешанного или иным образом связанного с представляющим интерес продуктом;

амплификацию по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК из биологического материала, и секвенирование ДНК последовательности уникального идентификатора; и

поиск ДНК последовательности уникального идентификатора в базе данных и извлечение записи из базы данных, соответствующей ДНК последовательности уникального идентификатора, где запись в базе данных предоставляет информацию об идентификации и/или по отслеживанию представляющего интерес продукта.

В другом варианте осуществления вышеуказанного способа способ может включать введение или добавление любого биологического материала или биологических материалов, или биологического объекта, или биологических объектов, как описано в настоящем документе, к представляющему интерес продукту, биологическому материалу или объекту, содержащему по меньшей мере одну ДНК последовательность уникального идентификатора, как описано в настоящем документе, как часть ее геномного материала.

В еще одном варианте осуществления любого из вышеуказанного способа или способов информация об идентификации и/или по отслеживанию в записи в базе данных может содержать информацию о цепочке поставок представляющего интерес продукта.

В еще одном варианте осуществления любого из вышеуказанного способа или способов представляющий интерес продукт может включать пищевой продукт, сельскохозяйственный продукт, фармацевтический препарат, розничный продукт, текстильный продукт, товар широкого потребления, химический продукт или другой элемент цепочки поставок.

Краткое описание фигур

Эти и другие особенности изобретения станут более понятными с учетом следующего описания и прилагаемых фигур, где показано следующее:

На фиг. 1 показаны пути передачи, указанные Всемирной организацией здравоохранения (ВОЗ) в их отчете за 2015 г. (адаптировано из публикации ВОЗ, 2015 г., стр. 101);

На фиг. 2 показан пример кассеты, описанной в настоящем документе, включающей последовательность DUID, и ее получение, как описано в Примере 1. Изображенная последовательность представляет собой SEQ ID NO: 1;

На фиг. 3 показан общий вид приведенного в качестве примера процесса для системы DUID, описанного в Примере 1;

На фиг. 4 показан пример стадии идентификации системного процесса DUID, как описано в Примере 1;

На фиг. 5 показан пример стадии проверки системного процесса DUID, как описано в Примере 1;

На фиг. 6 показан пример стадии чтения системного процесса DUID, как описано в Примере 1;

На фиг. 7 показан другой пример системы DUID и процесса, как описано в настоящем документе;

На фиг. 8 показан другой пример системы и процесса DUID, как описано в настоящем документе, где возможность отслеживания биологического объекта обеспечивается с использованием DUID и базы данных/реестра;

На фиг. 9 показан еще один пример системы и процесса DUID, как описано в настоящем документе, где информацию об идентификации и/или по отслеживанию биологического материала получают из базы данных с использованием последовательности DUID и базы данных/реестра;

На фиг. 10 показан другой пример системы и процесса DUID, как описано в настоящем документе, где возможность отслеживания биологического объекта обеспечивается с использованием DUID, хранящей информацию об идентификации и/или по отслеживанию;

На фиг. 11 показан другой пример системы и процесса DUID, как описано в настоящем документе, где информация об идентификации и/или по отслеживанию биологического материала получают с использованием последовательности DUID, хранящей информацию об идентификации и/или по отслеживанию;

На фиг. 12 показан другой пример системы DUID и процесса, как описано в настоящем документе, где информация об идентификации и/или по отслеживанию биологического материала получают с использованием последовательности DUID, хранящей информацию об идентификации и/или по отслеживанию;

На фиг. 13 показаны дополнительные примеры конструкций касет, как описано в настоящем документе, включающих последовательность уникального идентификатора (UID). На фиг. 13(a) показана конструкция с двойным праймером, на фиг. 13(b) показана конструкция с одним праймером, и на фиг. 13(c) показана конструкция без праймера;

На фиг. 14 показаны карты двух конструкторов DUID размером 370 п.о., как описано в Примере 2. А) Конструктор DUID для ПЦР-амплификации и количественной ПЦР. Конструктор имеет 370 п.о. Этот конструктор DUID содержит 2 прямых праймера и два обратных праймера. В нем есть два идентификатора (ID1 и ID2). ID1 идеально подходит для ПЦР-амплификации. ID2 идеально подходит для амплификации qPCR. В) Конструктор DUID для петлевой изотермической амплификации (LAMP) и ПЦР. Эта карта включает праймеры как для ПЦР, так и для LAMP;

На фиг. 15 показано обнаружение YCr-DUID в геномной ДНК дрожжей с помощью ПЦР конечной точки, как описано в Примере 2. Амплификацию ПЦР проводили с использованием (А) вектора YCr-DUID, (В) гДНК, экстрагированной из BY4743, и (С) штамма дрожжей BY4743, трансформированного вектором YCr-DUID, в качестве матриц с праймерами отзыва DUID. Реакции проводили с использованием последовательно разбавленной ДНК-матрицы с исходными количествами (1) 100 нг, (2) 10 нг, (3) 1 нг, (4) 100 мкг, (5) 10 мкг, (6) 1 мкг, (7) 100 мкг и (8) 10 фг, и разделение проводили на 1% агарозном геле с использованием в качестве стандарта GeneRuler™ 100 bp Plus Ready-to-use Ladder;

На фиг. 16 показано обнаружение DUID в экстрактах общей ДНК дрожжей, как описано в Примере 2. Количественную ПЦР в реальном времени проводили с серийными 10-кратными разбавлениями вектора YCr в диапазоне от 50 нг до 500 мкг для получения стандартной кривой (синяя линия) с помощью MS Excel. Результаты аналогичного эксперимента с количественной ПЦР с использованием ДНК, полученной из BY4743, трансформированной вектором YCr-DUID, наносили на график (оранжевая полоса) и сопоставляли со значениями стандартной кривой для количественного определения обнаружения DUID в биомассе дрожжей; и

На фиг. 17 показан пример гомологии последовательностей идентификаторов, которые действуют как средство для идентификации варианта DUID, его происхождения и субпоследовательностей протоколов для взаимодействия с DUID, как дополнительно описано в Примере 2.

Подробное описание настоящего изобретения

В настоящем документе описаны способы и композиции для обеспечения идентификации и/или отслеживания биологического материала. Следует понимать, что варианты осуществления и примеры представлены для иллюстративных целей, и они предназначены для специалистов в данной области техники, и никоим образом не предназначены для ограничения притязаний.

В настоящем документе предложены способы и композиции для обеспечения идентификации и/или отслеживания биологических материалов. В некоторых вариантах осуществления способов, описанных в настоящем документе, может использоваться последовательность уникального идентификатора (также называемая здесь ДНК последовательностью уникального идентификатора), которая может быть экзогенно введена (т.е. вставлена/интегрирована) в геном биологического объекта, чтобы обеспечить идентификацию и/или отслеживание биологического объекта и/или биологических материалов, содержащих биологический объект, и/или биологических материалов, полученных из биологического объекта, и содержащих его геномную ДНК. В некоторых вариантах осуществления изобретения стратегии, описанные в настоящем документе, могут обеспечить эффект за счет устойчивости и способности к репликации нуклеиновой кислоты, такой как ДНК, для обеспечения идентификации и/или возможности отслеживания. В некоторых вариантах осуществления изобретения последовательность уникального идентификатора может быть получена из рандомизированного пула последовательностей. В некоторых вариантах осуществления изобретения может поддерживаться база данных, связывающая последовательности уникальных идентификаторов с соответствующей информацией об идентификации и/или по отслеживанию.

Также в настоящем документе предложены олигонуклеотидные конструкции и кассеты, содержащие одну или более уникальных последовательностей-идентификаторов для использования при обеспечении идентификации и/или отслеживания биологических материалов. В некоторых вариантах осуществления изобретения олигонуклеотидные конструкции и/или кассеты могут содержать специфические конфигурации последовательностей праймеров отжига, которые могут быть предназначены для амплификации последовательностей уникальных идентификаторов, секвенирования последовательностей уникальных идентификаторов или того и другого. В некоторых вариантах осуществления расположение последовательности (последовательностей) праймеров отжига может быть представлено, как описано в настоящем документе, для уменьшения непреднамеренной и/или нецелевой амплификации и/или нежелательных событий секвенирования, что может, например, обеспечить повышенную точность и/или уменьшить ошибки результатов идентификации.

В некоторых вариантах осуществления изобретения способы и композиции, описанные в настоящем документе, могут использоваться для обеспечения возможности отслеживания пищевых продуктов, и они могут обеспечить быстрое реагирование и/или отзыв пищевых продуктов, например, в случае их контаминации. Контаминация пищевых продуктов, такое как заражение *E. coli* и/или *Salmonella*, влияющее на пищевые продукты, представляет собой угрозу для здоровья населения, и в этой связи крайне желательны быстрые действия по выявлению и устранению источника(ов) контаминации. В этой области давно ощущается неудовлетворенная потребность в надежных, экономически эффективных и/или быстрых стратегиях повышения возможности отслеживания продуктов в системе снабжения продовольствием. Стратегии, описанные в настоящем документе, могут обеспечить возможность отслеживания в пищевой системе от источника происхождения до пищеварения и далее. Отслеживание биологических объектов и/или биологических материалов желательно не только в сельском хозяйстве и пищевой промышленности, но оно также востребовано в самых разных отраслях и областях, связанных с биологическими объектами и/или

биологическими материалами, содержащими их или полученными из них. Соответственно, в дополнение к безопасности пищевых продуктов, здесь также рассматриваются применения в области безопасности пищевых продуктов/семян, отслеживания интеллектуальной собственности (ИС), сертификации (например, ассоциации семян, кошерность, халяльность и т.п.), для идентификации и/или характеристики генетически модифицированного объекта (ГМО) и/или для снижения рисков финансирования торговли.

В некоторых вариантах осуществления изобретения для обеспечения идентификации и/или отслеживания пищевые продукты или ингредиенты (такие как, например, фрукты и овощи или другие подобные пищевые продукты, содержащие клетки) могут содержать последовательность(и) уникального идентификатора, как описано здесь, в виде части генома по меньшей мере в некоторых их клетках. В других вариантах осуществления изобретения последовательность(и) уникального идентификатора, как описано здесь, может быть частью генома одного или более биологических объектов или биологических материалов, содержащих клетки, и биологические объекты или биологические материалы могут быть добавлены, смешаны или иным образом объединены с одним или более продуктами, для которых требуется идентификация и/или отслеживание. Например, в некоторых вариантах осуществления изобретения безопасные для пищевых продуктов дрожжевые клетки, содержащие одну или более последовательностей уникальных идентификаторов, как описано здесь, как часть одной или более стабильно введенных искусственных хромосом, могут быть добавлены или смешаны с одним или более пищевыми продуктами или пищевыми ингредиентами для обеспечения их идентификации и/или отслеживания.

Методы идентификации и/или обеспечения возможности отслеживания

В некоторых вариантах осуществления изобретения в настоящем документе представлены способы идентификации и/или обеспечения возможности отслеживания биологического материала или биологического объекта. В таких способах может использоваться последовательность уникального идентификатора для достижения такой идентификации и/или возможности отслеживания. Как правило, представляющий интерес биологический объект, такой как сельскохозяйственная культура (например, шпинат), может быть генетически модифицирована для включения в его геном последовательности уникального идентификатора. В качестве неограничивающего и иллюстративного примера, клетка растения шпината может быть генетически модифицирована для включения кассеты, содержащей последовательность уникального идентификатора, окруженную одной или более последовательностями праймеров отжига для последующей амплификации и/или секвенирования последовательности уникального идентификатора, в геном клетки шпината в межгенном или другом безвредном участке генома. Последовательность уникального идентификатора может быть известна или может быть получена из рандомизированного пула, и впоследствии она может быть определена после интеграции, а также она может быть введена и затем записана в базе данных или реестре. Затем клетку можно использовать для выращивания/размножения с получением одного или более урожаев шпината, и вместе с соответствующей последовательностью уникальных идентификаторов, соответствующей информации об идентификации и/или по отслеживанию урожая шпината (например, источник происхождения, информация о партии/лоте, фермере/производителе, местонахождении, дате, поставщике и/или любая другая представляющая интерес информация о цепочке поставок), могут быть записаны в базу данных или реестр. При необходимости запись в базе данных может обновляться по мере развития событий в цепочке поставок (например, сбор урожая, отгрузка поставщику, продажа и т.п.). Урожай шпината можно использовать для производства биологического материала, такого как пакет шпината, или готового салата для продажи в продуктовом магазине. В случае или при подозрении на контаминацию или на

возникновение заболевания пищевого происхождения может быть получен образец подозреваемого шпината или салата, из него может быть получена геномная ДНК, и геномная ДНК может быть проанализирована для определения наличия или отсутствия последовательности уникального идентификатора (т.е. является ли этот шпинат шпинатом, отслеживаемым настоящей системой), и, если да, то последовательность уникального идентификатора может быть секвенирована для определения последовательности нуклеотидов, и эта последовательность нуклеотидов может использоваться для представления запроса в базу данных или в реестр, чтобы получить соответствующую запись из базы данных, содержащую информацию об идентификации и/или по отслеживанию, чтобы облегчить отзыв контаминированного шпината или салата. Понятно, что вышеприведенный пример со шпинатом приведен в иллюстративных целях, и способы, описанные в настоящем документе, могут быть использованы для предоставления широкого спектра вариантов идентификации и/или отслеживания в отношении широкого спектра биологических объектов и/или биологических материалов для многих различных применений.

В одном варианте осуществления изобретения в настоящем документе предложен способ идентификации биологического материала, включающий:

получение или предоставление образца, содержащего геномную ДНК, из биологического материала; амплификацию по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК из биологического материала, и секвенирование ДНК последовательности уникального идентификатора; и

поиск ДНК последовательности уникального идентификатора в базе данных и извлечение записи из базы данных, соответствующей ДНК последовательности уникального идентификатора, где запись в базе данных предоставляет информацию об идентификации и/или по отслеживанию биологического материала.

Блок-схема, изображающая вариант осуществления такого способа, показана на фиг. 9.

Понятно, что биологический материал может включать, как правило, любой подходящий биологический материал, представляющий интерес. Биологический материал может содержать или состоять из материала, содержащего или состоящего из биологического объекта, или может содержать или состоять из материала, изготовленного или полученного из биологического объекта, или любого другого подходящего материала, представляющего интерес, который содержит геномную нуклеиновую кислоту (т.е. геномную ДНК) из биологического объекта. В некоторых вариантах осуществления изобретения биологический материал может содержать или состоять из материала на основе растения, материала на основе гриба, материала животного происхождения, материала на основе вируса или материала на основе бактерии. Например, в некоторых вариантах осуществления изобретения биологический материал может включать или состоять из пищевого продукта или напитка, содержащего или состоящего из растения или другого биологического объекта, или изготовленного из него, где пищевой продукт или напиток содержит геномную ДНК биологического объекта. В некоторых вариантах осуществления изобретения биологический материал может содержать или состоять из салата, шпината или другой листовой зелени, или, например, представлять собой пищевой продукт, содержащий или состоящий из них, или изготовленный из них.

В некоторых вариантах осуществления способов, описанных в настоящем документе, может быть получен или предоставлен образец, содержащий геномную нуклеиновую кислоту (т.е. геномную ДНК, где биологический объект имеет геном на основе ДНК) из представляющего интерес биологического материала (например, из биологического материала, для которого требуется идентификация). Образец может быть получен или предоставлен в очищенной или частично очищенной форме, так что геномная ДНК может быть легко использована, или образец может быть получен или предоставлен по существу в том виде, как он

существует (т.е. в виде образца пищевого продукта) или в другой неочищенной форме, или в форме предшественника, которые могут быть подвергнуты одной или более стадиям обработки или очистки, чтобы содержащаяся в них геномная ДНК могла быть легко использована на последующих стадиях. В некоторых вариантах осуществления изобретения предполагается, что для подготовки образцов можно использовать любую подходящую стандартную методику очистки и/или выделения геномных нуклеиновых кислот.

В некоторых вариантах осуществления изобретения, а также в качестве примера, одна или более стадий выделения, очистки и/или экстракции нуклеиновой кислоты (например, генома) могут быть выполнены как часть подготовки образцов для последующих стадий. Выделение или экстракция ДНК может включать, например, одну или более стадий получения ДНК из образца. В некоторых вариантах осуществления изобретения выделение или экстракция ДНК может включать разрушение клеток (например, лизис), или, например, разрушение с помощью физических стадий обработки, обработки ультразвуком или химической обработки; удаление мембраны с помощью детергента; необязательное удаление белков протеазой; и осаждение ДНК с использованием спирта (такого как этанол (холодный) или изопропанол). Таким образом, осадок ДНК может быть получен центрифугированием. В некоторых вариантах осуществления изобретения активность ДНК-азных ферментов может быть подавлена с помощью хелатирующего агента, как известно специалистам в данной области техники. В некоторых вариантах осуществления изобретения клеточные и гистоновые белки могут быть удалены с помощью протеазы, или путем осаждения ацетатом натрия или аммония, или путем экстракции фенолом-хлороформом перед осаждением ДНК. Принимая во внимание изложенные здесь идеи, специалисту в данной области техники понятно, что для подготовки образцов и/или для выделения, очистки и/или экстракции геномной нуклеиновой кислоты доступен широкий спектр методов, используемых по мере необходимости.

В некоторых вариантах осуществления способов, описанных в настоящем документе, последовательность уникального идентификатора (упоминаемая здесь для удобства как ДНК последовательность уникального идентификатора, DUID; хотя при этом следует понимать, что в некоторых примерах, например, когда биологический объект имеет геном на основе РНК, последовательность уникального идентификатора может представлять собой РНК, а не ДНК), вставленная или интегрированная в геном биологического объекта/биологического материала, необязательно может быть амплифицирована.

В некоторых вариантах осуществления изобретения интеграция в геном может включать интеграцию в нативную хромосому. В некоторых вариантах осуществления изобретения интеграция в геном может включать стабильное введение в геном искусственной хромосомы, где искусственная хромосома имеет центромерную последовательность и наследуется вместе с нативным геномным материалом. Пример 2, представленный ниже, описывает пример использования искусственных хромосом, например, в дрожжах.

Такая амплификация может быть осуществлена, как правило, с использованием любой подходящей методики амплификации, известной специалисту в данной области техники, принимая во внимание изложенные здесь идеи, такой как полимеразная цепная реакция (ПЦР). В некоторых вариантах осуществления изобретения, как более подробно описано здесь, последовательность уникального идентификатора, подлежащая амплификации, может сопровождаться в геноме последовательностями праймеров отжига для амплификации и/или секвенирования. В некоторых вариантах осуществления изобретения последовательности праймеров отжига могут быть выбраны и расположены таким образом, чтобы обеспечить возможность амплификации с помощью вложенной ПЦР для снижения вероятности непреднамеренной или нецелевой амплификации, как более подробно описано здесь далее.

В некоторых вариантах осуществления изобретения для амплификации можно использовать подход, основанный на ПЦР. ПЦР-амплификация может предусматривать использование прямых и обратных праймеров, где праймеры могут быть комплементарными (или по существу комплементарными) участкам на 5'- и 3'-концах амплифицируемой последовательности нуклеиновой кислоты, представляющей интерес. Прямые и обратные праймеры к конкретным последовательностям праймеров отжига могут быть получены любым подходящим методом, известным специалисту в данной области. Примеры таких методов можно найти, например, в Dieffenbach CW, Dveksler GS. 1995. PCR primer: a laboratory manual, New York, NY: Cold Spring Harbor Laboratory Press; New England Biolabs Inc., 2007-08 Catalog & Technical Reference, которые включены в настоящее описание посредством ссылки. В некоторых вариантах осуществления изобретения для считывания биоинформации праймеры для ПЦР могут содержать множество наборов прямых и обратных праймеров, которые могут действовать независимо друг от друга. В некоторых вариантах осуществления изобретения может обеспечиваться или использоваться идентичность некоторых праймеров, в то время как доступ к другим праймерам может контролироваться, так что разные участвующие стороны могут иметь возможность легкого доступа к различным областям и/или информации о последовательности нуклеиновой кислоты по желанию.

В некоторых вариантах осуществления способов, описанных в настоящем документе, последовательность уникального идентификатора, такая как ДНК последовательность уникального идентификатора (DUID), может содержать любую подходящую последовательность нуклеиновой кислоты, экзогенно введенную в геном биологического объекта для целей идентификации. Как правило, последовательностью уникального идентификатора может быть либо ДНК, либо РНК, чтобы она соответствовала типу генома (ДНК или РНК) биологического объекта. Понятно, что геном многих биологических объектов, например растений, является двухцепочечным, поэтому последовательность уникального идентификатора обычно находится в геноме в двухцепочечной форме. Таким образом, понятно, что в некоторых вариантах осуществления изобретения, приведенные в настоящем документе ссылки на последовательность уникального идентификатора (например, при описании секвенирования последовательности идентификатора) могут пониматься как ссылки на любую цепь из двухцепочечной конструкции или на обе цепи, как желательное или уместное.

В некоторых вариантах осуществления изобретения последовательность уникального идентификатора может быть включена в кассету или другой подобный конструкт, содержащий один или более функциональных элементов в дополнение к последовательности уникального идентификатора. В некоторых вариантах осуществления кассета может содержать последовательность уникального идентификатора, фланкированную одной или более последовательностями праймеров отжига для ПЦР-амплификации ДНК последовательности уникального идентификатора, секвенирования ДНК последовательности уникального идентификатора или для того и другого. Понятно, что последовательность праймеров отжига может относиться к заранее определенной последовательности или к области нуклеиновой кислоты, имеющей известную нуклеотидную последовательность, так что один или более праймеров могут быть сконструированы или выбраны для отжига с такой последовательностью праймеров отжига, чтобы инициировать полимеризацию, например, полимеразой. Как правило, последовательности праймеров отжига выбирают таким образом, чтобы они были уникальными в пределах генома биологического объекта, представляющего интерес, чтобы уменьшить или исключить непреднамеренную или нецелевую амплификацию. В некоторых вариантах осуществления изобретения последовательность уникального идентификатора может быть известной и заранее определенной последовательностью, выбранной для

конкретного применения, или она может быть случайной последовательностью, полученной из рандомизированного пула последовательностей нуклеиновых кислот, которые впоследствии могут быть определены и записаны в базу данных, например, как описано подробно здесь. В некоторых вариантах осуществления изобретения последовательность уникального идентификатора или кассета, содержащая последовательность уникального идентификатора, может иметь размер до около 1500 нуклеотидов (нт) в длину; до около 1000 нт в длину; от около 200 до около 600 нт в длину; от около 200 нт до около 400 нт в длину; или от около 400 до около 600 нуклеотидов в длину; или иметь любой размер или поддиапазон размеров в пределах любых из этих двух показателей. Понятно, что более длинные последовательности уникальных идентификаторов могут обеспечить большее количество уникальных последовательностей в пуле, и они могут обеспечить снижение риска дублирования. Кроме того, в вариантах осуществления, где требуется кодирование или шифрование идентификационной информации в пределах последовательности уникального идентификатора, более длинные последовательности могут позволить хранить относительно больше информации и/или использовать, например, более сложные схемы шифрования или кодирования. Тем не менее, поддерживая разумную длину последовательности, такие как упомянутые в настоящем документе, можно проводить более надежную и/или быструю амплификацию и/или секвенирование, и/или затраты на идентификацию могут быть относительно снижены.

В некоторых вариантах осуществления изобретения последовательность уникального идентификатора может содержать последовательность размером до около 1500 нуклеотидов (нт) в длину; до около 1000 нт в длину; от около 200 до около 600 нт в длину; от около 200 нт до около 400 нт в длину; или от около 400 нт до около 600 нт в длину. В некоторых вариантах осуществления последовательность уникального идентификатора может быть относительно короткой, и длиной, например, около 20 пар оснований. Понятно, что в некоторых вариантах осуществления изобретения размер последовательности уникального идентификатора может быть выбран в соответствии с конкретным применением и требуемыми параметрами. В некоторых вариантах осуществления изобретения последовательность уникального идентификатора может иметь размер от около 20 нуклеотидов до около 1500 нуклеотидов в длину, или иметь любой размер или поддиапазон размеров в пределах любых этих двух значений.

В некоторых вариантах осуществления изобретения последовательность уникального идентификатора может быть получена случайным образом из пула, и она необязательно может подвергаться скринингу на приемлемость (например, скринингу на уникальность, скринингу на исключение нежелательных мотивов последовательности) или она может быть, например, рационально сконструирована (например, сконструирована с целью ее уникальности, или сконструирована таким образом, чтобы избежать нежелательных мотивов последовательности).

В некоторых вариантах осуществления изобретения ДНК последовательность уникального идентификатора может быть фланкирована одной или более последовательностями праймеров отжига для ПЦР-амплификации ДНК последовательности уникального идентификатора, секвенирования ДНК последовательности уникального идентификатора или для того и другого.

В некоторых вариантах осуществления изобретения предполагается, что последовательность уникального идентификатора может быть предоставлена в кассете, или она может быть иным образом введена или встроена в геномную нуклеиновую кислоту таким образом, чтобы она была фланкирована одной или более последовательностями праймеров отжига для ПЦР-амплификации ДНК последовательности уникального идентификатора, секвенирования ДНК последовательности уникального идентификатора или для того и другого. Примеры подходящих кассет и конфигураций описаны здесь более подробно. В некоторых

вариантах осуществления изобретения кассета может быть включена в плазмиду, вектор или другой подобный носитель, пригодный для использования при inserции/встраивании/интеграции кассеты в геном биологического объекта.

Понятно, что с учетом изложенных здесь идей любая подходящая методика генетической модификации, известная специалисту в данной области, может быть использована для введения/инсерции/встраивания/интеграции последовательности уникального идентификатора или кассеты/вектора, с последовательностью уникального идентификатора, в геном биологического объекта. Также понятно, что метод генетической модификации может быть выбран на основе последовательности уникального идентификатора или используемой кассеты/вектора, а также на основе конкретного модифицируемого биологического объекта. Методы модификации генома для широкого круга биологических объектов, включая растения, животных, грибы, бактерии и вирусы, хорошо известны, и они могут быть легко адаптированы для экзогенного введения последовательности уникального идентификатора, как описано в настоящем документе.

Например, специалисту в данной области техники, имеющего отношение к изложенным здесь идеям, известны векторы для включения ДНК в организм, которые можно сконструировать в соответствии с известными принципами молекулярной биологии. Такие векторы могут быть, например, предназначены для стабильного введения представляющей интерес последовательности ДНК в геном организма. В некоторых вариантах осуществления изобретения векторы могут иметь, например, вирусное происхождение или получены из них. Если организм представляет собой растение, предполагается, что, например, для введения в растение представляющей интерес ДНК можно использовать встраивание, опосредованное *Agrobacterium tumefaciens*. Квалифицированному специалисту, имеющему отношение к приведенным здесь идеям, известны, среди прочих, несколько других методов трансформации, таких как баллистические методы или методы генной пушки, которые могут быть адаптированы по желанию или в зависимости от конкретного интересующего применения. В некоторых вариантах осуществления изобретения можно использовать систему доставки генов, основанную на принципах генной инженерии, так чтобы представляющая интерес последовательность была введена или вставлена в геном организма-хозяина. Например, в одном из вариантов осуществления изобретения можно использовать систему транспозонов для встраивания в геном хозяина, которым может быть, например, микроорганизмом, клеткой животного или клеткой растения (Insect Molecular Biology (2007), 16(1), 37-47, Plant Physiology Preview. 2007, DOI: 10.1104/pp.107.111427, The American Society of Plant Biologists; research on production of lactoferrin from transformed silkworms and functionality thereof, the Ministry of Agriculture and Forestry, 2005). В некоторых вариантах осуществления изобретения может быть использован любой подходящий метод, известный в области молекулярной биологии и/или генной инженерии, который позволяет встраивать один или более фрагментов или компонентов ДНК, представляющих интерес, в геном хозяина (см., например, Transgenic Plants Methods and Protocols., Methods in Molecular Biology 2019, Editors: Kumar, Sandeep, Barone, Pierluigi, Smith, Michelle, ISBN 978-1-4939-8778-8, содержание которого полностью включено в настоящее описание посредством ссылки).

В некоторых вариантах осуществления изобретения, когда требуется идентификация биологического материала или биологического объекта, содержащего последовательность уникального идентификатора, последовательность уникального идентификатора может быть определена путем секвенирования. Понятно, что последовательность уникального идентификатора может быть секвенирована, как правило, с помощью любого подходящего метода секвенирования, известного специалисту в данной области, принимая во внимание идеи, изложенные в настоящем документе. В некоторых вариантах осуществления изобретения

секвенированию может способствовать включение или использование последовательности праймера отжига для секвенирования, связанной с последовательностью уникального идентификатора в геномной нуклеиновой кислоте. Примеры такой секвенирующей последовательности праймера отжига, которая может быть включена, например, в кассету, содержащую последовательность уникального идентификатора, подробно описаны в настоящем документе.

В некоторых вариантах осуществления изобретения секвенирование может быть выполнено с использованием любого подходящего метода секвенирования, известного специалисту в данной области, с учетом изложенных здесь идей, который может быть выбран на основе конкретного применения и/или используемой конфигурации. В некоторых вариантах осуществления изобретения секвенирование можно проводить любым подходящим методом секвенирования для определения порядка нуклеотидных оснований в молекуле ДНК (или РНК). Примеры методов секвенирования могут включать, например, секвенирование Максама-Гилберта, методы обрыва цепи, секвенирование с терминатором красителя, автоматическое секвенирование ДНК, амплификацию клонированием *in vitro*, параллельное секвенирование путем синтеза, секвенирование путем лигирования, секвенирование по Сэнгеру, такое как микрофлюидное секвенирование по Сэнгеру, и, например, секвенирование путем гибридизации.

В некоторых вариантах осуществления изобретения после определения последовательности уникального идентификатора биологического материала эта последовательность может быть использована для представления запроса на поиск в базе данных (также именуемой в настоящем документе реестром), содержащей набор последовательностей уникальных идентификаторов, ассоциированных или иным образом связанных с соответствующей информацией об идентификации и/или по отслеживанию. Если найдена совпадающая запись в базе данных, то эта запись в базе данных может быть извлечена, чтобы предоставить информацию об идентификации и/или по отслеживанию интересующего биологического материала. Таким образом, может быть определена релевантная информация об идентификации и/или по отслеживанию биологического материала, которая может быть использована, например, для информирования о событии, таком как, например, отзыв пищевых продуктов, или о другом действии.

В другом варианте осуществления изобретения в настоящем документе предложен способ обеспечения возможности отслеживания биологического материала, включающий:

определение последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК биологического объекта;

подтверждение идентификации биологического объекта посредством: проверки наличия ДНК последовательности уникального идентификатора в геномной ДНК; и сравнение ДНК последовательности уникального идентификатора с базой данных для подтверждения того, что ДНК последовательность уникального идентификатора еще не используется в базе данных;

предоставление указания о соответствии получения биологического материала из биологического объекта, где биологический материал содержит геномную ДНК из биологического объекта; и

ввод последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в запись базы данных, и связывание ДНК последовательности уникального идентификатора с информацией об идентификации и/или по отслеживанию биологического материала;

тем самым обеспечивая возможность отслеживания биологического материала путем считывания ДНК последовательности уникального идентификатора в биологическом материале и извлечения соответствующей записи из базы данных, предоставляющей информацию об идентификации и/или по отслеживанию биологического материала.

Блок-схема, изображающая вариант осуществления такого способа, показана на фиг. 8.

Понятно, что биологический объект может включать, как правило, любой подходящий биологический материал, представляющий интерес. Биологический объект может содержать или состоять из клетки (т.е. из клетки растения, клетки гриба, клетки животного или бактериальной клетки), или семени, или ткани, содержащей одну или более клеток, или вируса, или организма, такого как растение, животное или гриб, или содержать или состоять из любой их части. В некоторых вариантах осуществления изобретения биологический объект может включать клетку растения, клетку гриба, клетку животного, вирус или бактериальную клетку. Когда биологический объект должен быть генетически модифицирован для включения последовательности уникального идентификатора, то клетку или вирус, который содержит биологический объект, обычно можно размножить после генетической модификации для получения большего количества биологических объектов, каждый из которых содержит вставленную последовательность уникального идентификатора.

В некоторых вариантах осуществления изобретения стадия проверки может быть выполнена для проверки наличия последовательности уникального идентификатора в геномной ДНК биологического объекта, и/или для определения ее последовательности, и/или для определения того, не является ли последовательность уникального идентификатора уже используемой в базе данных (т.е. что она представляет собой новую последовательность, которая ранее не была связана ни с какой записью в базе данных). Если проверка прошла успешно (т.е. последовательность уникального идентификатора правильно вставлена и она уникальна для базы данных), то в некоторых вариантах осуществления изобретения в базе данных может быть создана запись для последовательности уникального идентификатора (которая может быть связана с соответствующей информацией об идентификации и/или по отслеживанию, и которая может обновляться на постоянной основе), то указание на приемлемость производства биологического материала из биологического объекта может быть предоставлено заинтересованной стороне, такому как производителю сельскохозяйственной продукции, фермеру или другому субъекту, связанному с сельским хозяйством, который может далее производить или выращивать биологический материал.

Таким образом, возможность отслеживания биологического материала может быть обеспечена путем считывания (т.е. секвенирования) последовательности уникального идентификатора биологического материала, которая может использоваться для извлечения соответствующей записи из базы данных для получения информации об идентификации и/или по отслеживанию.

В некоторых вариантах осуществления изобретения способы, описанные в настоящем документе, могут дополнительно предусматривать вставку по меньшей мере одной ДНК последовательности уникального идентификатора в геномную ДНК биологического объекта или модификацию уже существующей последовательности идентификатора в геномной ДНК биологического объекта путем редактирования гена для создания ДНК последовательность уникального идентификатора в геномной ДНК биологического объекта, тем самым обеспечивая его идентификацию.

В еще одном варианте осуществления изобретения способы, описанные в настоящем документе, могут дополнительно предусматривать предоставление по меньшей мере одной ДНК последовательности уникального идентификатора для вставки в геномную ДНК биологического объекта. В некоторых вариантах осуществления ДНК последовательность уникального идентификатора может быть представлена в виде рандомизированного пула последовательностей, как дополнительно описано в настоящем документе.

Понятно, что предполагается, что в определенных вариантах осуществления изобретения способы, описанные в настоящем документе, могут предусматривать использование одной последовательности

уникального идентификатора или они могут предусматривать использование двух или более последовательностей идентификаторов, включенных в геном, для обеспечения идентификации и/или возможности отслеживания.

В некоторых вариантах осуществления изобретения последовательность уникального идентификатора может быть из рандомизированного пула последовательностей уникального идентификатора. Идентичность вставленной последовательности уникального идентификатора не может быть определена до тех пор, пока не будет осуществлена вставка (т.е. трансформация или генетическая модификация). Таким образом, предполагается, что заинтересованным сторонам может быть предоставлен рандомизированный пул последовательностей уникальных идентификаторов, и они могут выполнить генетическую модификацию представляющего интерес биологического объекта, так что одна, две или более последовательностей уникальных идентификаторов будут вставлены в геном. После процесса генетической модификации вставленная последовательность (последовательности) уникального идентификатора может быть секвенирована для определения нуклеотидной последовательности вставленной последовательности (последовательностей) уникального идентификатора. Учитывая, что типичная длина последовательности уникального идентификатора обычно может быть выбрана достаточно большой, чтобы обеспечить большое количество различных последовательностей в рандомизированном пуле, то статистическая вероятность того, что два разных представителя заинтересованных сторон вставят одну и ту же последовательность уникального идентификатора будет чрезвычайно низкой. Соответственно, таким образом предполагается, что в некоторых вариантах осуществления изобретения многим различным представителям заинтересованных сторон, стремящимся извлечь выгоду из идентификации и/или возможности отслеживания продукта способами, описанными в настоящем документе, может быть представлен образец из одного и того же рандомизированного пула последовательностей, предназначенный для вставки в биологические объекты, представляющие интерес. Таким образом, предполагается, что в некоторых вариантах осуществления изобретения процессы могут быть оптимизированы и/или затраты на выполнение способа по изобретению могут быть снижены.

В еще одном варианте осуществления описанных здесь способов считывание ДНК последовательности уникального идентификатора в биологическом материале и поиск соответствующей записи в базе данных может предусматривать:

получение или предоставление образца, содержащего геномную ДНК, из биологического материала; амплификацию по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК из биологического материала, и секвенирование ДНК последовательности уникального идентификатора; и

сравнение ДНК последовательности уникального идентификатора с базой данных, и извлечение записи из базы данных о соответствующей ДНК последовательности уникального идентификатора, где запись в базе данных предоставляет информацию об идентификации и/или по отслеживанию биологического материала.

В некоторых вариантах осуществления изобретения предполагается, что последовательность (последовательности) уникального идентификатора может быть вставлена в геном биологического объекта в по существу безвредном месте (т.е. она не может существенно влиять на экспрессию гена или фенотип). Например, в некоторых вариантах осуществления изобретения предполагается, что последовательность(и) уникального идентификатора может быть вставлена в одну или более межгенных областей геномной ДНК.

В некоторых вариантах осуществления изобретения информация об идентификации и/или по отслеживанию, представленная в базе данных или реестре, может включать информацию о цепочке поставок биологического материала. В некоторых вариантах осуществления изобретения информация об идентификации и/или по отслеживанию в базе данных может содержать информацию об источнике происхождения биологического материала. В некоторых вариантах осуществления изобретения информация об идентификации и/или по отслеживанию в базе данных может включать информацию о производителе, регионе, партии, лоте, дате или о другой соответствующей информации о цепочке поставок или любую их комбинацию. Специалисту в данной области техники, имеющему отношение к изложенным здесь идеям, известно о различной информации об идентификации и/или по отслеживанию, которая может быть включена в базу данных, и которая может быть выбрана по желанию или в соответствии с конкретным применением. В некоторых вариантах осуществления изобретения в базу данных могут быть включены, например, существующие сведения для отслеживания цепочки поставок, такие как штрих-код или номер партии или лота. В некоторых вариантах осуществления изобретения такая информация, как географический регион, даты, покупатели, фермеры, партии, части партии, урожаи, лоты, другие продукты с DUID, организмы, договоры, сертификаты, сведения о связанных отраслях производства и предприятиях, данные от сенсорных датчиков, данные о погоде, или любые их комбинации могут быть включены/сохранены в базе данных.

В другом варианте осуществления изобретения в настоящем документе предложен способ идентификации биологического материала, включающий:

получение на вычислительном устройстве ДНК последовательности уникального идентификатора (DUID), извлеченной из известного биологического материала;

поиск в вычислительном устройстве базы данных DUID, в которой хранится множество DUID вместе с соответствующей информацией о биологическом материале, для совпадения с полученной DUID;

если поиск в базе данных DUID не дает совпадения с полученной DUID, то в базе данных DUID сохраняют полученную DUID вместе с информацией о биологическом материале, связанной с известным биологическим материалом;

после сохранения полученной DUID и информации, связанной с известным биологическим материалом в базе данных DUID, прием на вычислительном устройстве запроса DUID, извлеченной из неизвестного биологического материала;

поиск в вычислительном устройстве базы данных DUID для совпадения с полученным запросом DUID; и

если поиск DUID обеспечивает совпадение с полученным запросом DUID, возврат в ответ на полученный запрос DUID биологической информации, сохраненной в ассоциации с DUID, совпадающей с запросом DUID.

Блок-схема, изображающая вариант осуществления такого способа, показана на фиг. 7. На этой фигуре ДНК последовательность уникального идентификатора (DUID – DuID 4 в изображенном примере) извлекается (т.е. считывается, определяется или секвенируется) из известного биологического материала и подается на вычислительное устройство. Вычислительное устройство используется для поиска в базе данных DUID (т.е. в хранилище данных DuID), хранящей множество DUID в ассоциации с соответствующей информацией о биологическом материале, для совпадения с полученной DUID 4. Если поиск в базе данных DUID не дает соответствия полученной DUID, то полученная DUID (DuID 4) сохраняется в базе данных DUID вместе с информацией о биологическом материале (т.е. информацией о производителе 4), связанной с известным биологическим материалом, обеспечивая таким образом регистрацию DUID и биологического

материала в базе данных. Затем заинтересованной стороне может быть направлено уведомление об успешной регистрации и одобрению продолжения размножения биологического объекта/материала для производства биологического материала, такого как пищевой продукт. После сохранения полученной DUID и информации, связанной с известным биологическим материалом в базе данных DUID, запрашиваемая DUID, извлеченная (т.е. прочитанная, например, путем секвенирования) из неизвестного биологического материала (т.е. продукта, с подозрением на контаминацию), может быть получена вычислительным устройством, и может быть выполнен поиск в базе данных DUID для совпадения с полученным запросом DUID. Если поиск в базе данных DUID дает совпадение с полученным запросом DUID, то биологическая информация, хранящаяся в ассоциации с DUID, совпадающей с запросом DUID, может быть возвращена в ответ на полученный запрос DUID, предоставляя таким образом информацию об отслеживании и/или по идентификации биологического материала, которую можно использовать для принятия ответных мер, таких как, например, отзыв пищевого продукта.

В другом варианте осуществления изобретения поиск совпадения с полученной DUID в базе данных DUID может предусматривать:

поиск в базе данных DUID точного совпадения с полученной DUID; и

если точное совпадение не найдено, выполнение выравнивания/поиска идентичности для DUID, хранящихся в базе данных DUID, которые являются близким совпадением с полученной DUID.

В еще одном варианте осуществления изобретения поиск в базе данных DUID совпадения с запросом DUID может предусматривать:

поиск в базе данных DUID точного совпадения с запросом DUID; и

если точное совпадение не найдено, выполнение выравнивания/поиска идентичности для DUID, хранящихся в базе данных DUID, которые близко соответствуют запросу DUID.

Понятно, что поскольку используется последовательность нуклеиновой кислоты, то может иметь место возможность мутации последовательности уникального идентификатора во время размножения и/или амплификации, и/или могут возникать ошибки секвенирования. Соответственно, в некоторых вариантах осуществления изобретения такое выравнивание/поиск идентичности может выполняться для определения того, может ли существовать запись для близкого или очень похожего совпадения. Существует множество алгоритмов сравнения последовательностей для выполнения такой оценки выравнивания/идентичности/сходства (см., например, инструменты BLAST, доступные в NCBI), и специалист в данной области, принимая во внимание идеи, изложенные в настоящем документе, сможет выбрать или адаптировать соответствующий алгоритм, желательный для конкретного применения.

В еще одном варианте осуществления изобретения способы, описанные в настоящем документе, могут дополнительно предусматривать:

если поиск обеспечивает близкое совпадение с DUID запроса, сохранение DUID запроса в ассоциации с DUID, которая является близким соответствием DUID запроса.

Таким образом, база данных может быть обновлена, если, например, идентифицирована мутация последовательности.

В другом варианте осуществления изобретения в настоящем документе предложена вычислительная система для идентификации биологического материала, содержащая:

блок обработки, способный выполнять инструкции; и

блок памяти, хранящий инструкции, которые при выполнении их блоком обработки конфигурируют вычислительную систему для выполнения любого из описанного здесь способа или способов.

В другом варианте осуществления изобретения предлагается машиночитаемая память, содержащая хранящиеся в ней инструкции, которые при выполнении их блоком обработки вычислительной системы конфигурируют систему для выполнения любого описанного здесь способа или способов.

В другом варианте осуществления изобретения в настоящем документе предложен способ идентификации биологического материала, включающий:

получение или предоставление образца, содержащего геномную ДНК, из биологического материала; амплификацию по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК из биологического материала, и секвенирование ДНК последовательности уникального идентификатора; и

декодирование или расшифровку информации об идентификации и/или по отслеживанию биологического материала, хранящейся в ДНК последовательности уникального идентификатора.

Такие варианты осуществления способа могут быть аналогичны тем, которые описаны в настоящем документе, с использованием базы данных или реестра, за исключением того, что вместо сохранения в базе данных информации об идентификации и/или по отслеживанию, информация может быть закодирована (зашифрована или нет) в самой последовательности уникального идентификатора. Подходы к хранению информации в последовательности нуклеиновых кислот известны в данной области, и они обычно могут включать использование нуклеотидов А, Т, G, С, по аналогии с битами 0 и 1 при хранении цифровых данных. Примеры подходов к хранению/кодированию/шифрованию информации можно найти, например, в публикации Clelland, C., Risca, V. & Bancroft, C. Hiding messages in DNA microdots. *Nature* 399, 533–534 (1999) doi:10.1038/21092 (включена сюда посредством ссылки).

Блок-схема, изображающая вариант осуществления такого способа, показана на фиг. 11.

В некоторых вариантах осуществления изобретения предполагается, что последовательность уникальных идентификаторов может использоваться для кодирования ключа, и именно этот ключ хранится в базе данных вместе с информацией об идентификации и/или по отслеживанию. Таким образом, следует понимать, что ссылки в настоящем документе на хранение DUID в базе данных и поиск в базе данных DUID могут рассматриваться как охватывающие как прямые (т.е. хранение и поиск первичной последовательности нуклеиновой кислоты самой последовательности уникального идентификатора), так и косвенные (т.е. получение ключа из первичной последовательности нуклеиновой кислоты последовательности уникального идентификатора, и использование ключа для сохранения в базе данных и поиска) варианты. Специалисту в данной области техники, имеющему отношение к приведенным здесь идеям, известно о множестве комбинаций, которые могут быть использованы, и все они включены в настоящее описание.

В другом варианте осуществления изобретения в настоящем документе предложен способ обеспечения возможности отслеживания биологического материала, включающий:

определение последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК биологического объекта;

подтверждение идентификации биологического объекта посредством: проверки наличия ДНК последовательности уникального идентификатора в геномной ДНК; и декодирование или расшифровку информации об идентификации и/или по отслеживанию, хранящейся в ДНК последовательности уникального идентификатора, для проверки ДНК последовательности уникального идентификатора; и

предоставление указания о соответствии получения биологического материала из биологического объекта, где биологический материал содержит геномную ДНК из биологического объекта;

обеспечивая тем самым возможность отслеживания биологического материала путем считывания ДНК последовательности уникального идентификатора в биологическом материале и декодирования или расшифровки информации, хранящейся в ДНК последовательности уникального идентификатора, предоставляя информацию об идентификации и/или по отслеживанию биологического материала.

Блок-схема, изображающая вариант осуществления такого способа, показана на фиг. 10.

Такие варианты осуществления способа могут быть аналогичны тем, которые описаны в настоящем документе, с использованием базы данных или реестра, за исключением того, что вместо сохранения информации об идентификации и/или по отслеживанию в базе данных, информация может быть закодирована (зашифрована или нет) в самой последовательности уникального идентификатора. Подходы к хранению информации в последовательности нуклеиновых кислот известны в данной области и обычно могут включать использование нуклеотидов А, Т, G, С аналогично битам 0 и 1 при хранении цифровых данных. Примеры подходов к хранению/кодированию/шифрованию информации можно найти, например, в публикации Clelland, C., Risca, V. & Bancroft, C. Hiding messages in DNA microdots. *Nature* 399, 533–534 (1999) doi:10.1038/21092 (включена сюда посредством ссылки).

В еще одном варианте осуществления изобретения в настоящем документе предложен способ идентификации биологического материала, включающий:

получение на вычислительном устройстве ДНК последовательности уникального идентификатора (DUID), извлеченной из неизвестного биологического материала; и

декодирование или расшифровку информации об идентификации и/или по отслеживанию неизвестного биологического материала, хранящейся в ДНК последовательности уникального идентификатора.

Блок-схема, изображающая вариант осуществления такого способа, показана на фиг. 12.

В еще одном варианте осуществления изобретения в настоящем документе предложен способ обеспечения возможности отслеживания биологического материала, где указанный способ предусматривает:

встраивание по меньшей мере одной ДНК последовательности уникального идентификатора в геномную ДНК биологического объекта для использования при получении биологического материала.

В другом варианте осуществления вышеуказанного способа ДНК последовательность уникального идентификатора может быть встроена с помощью любой кассеты или кассет, как описано в настоящем документе.

В другом варианте осуществления любого из вышеуказанных способов, способ может дополнительно предусматривать стадию определения последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК биологического объекта.

В другом варианте осуществления любого из вышеуказанного способа или способов, способ может дополнительно предусматривать стадию проверки подтверждения подлинности идентификации биологического объекта путем: проверки наличия ДНК последовательности уникального идентификатора в геномной ДНК; и сравнения ДНК последовательности уникального идентификатора с базой данных для подтверждения того, что ДНК последовательность уникального идентификатора еще не используется в базе данных.

В еще одном варианте осуществления любого из вышеуказанного способа или способов, способ может дополнительно предусматривать стадию:

получение биологического материала из биологического объекта, где биологический материал содержит геномную ДНК из биологического объекта; и/или

предоставление указания о соответствии получения биологического материала из биологического объекта, где биологический материал содержит геномную ДНК из биологического объекта.

В еще одном варианте осуществления любого из вышеуказанных способа или способов, способ может дополнительно предусматривать стадию ввода последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в запись базы данных, и связывания ДНК последовательности уникального идентификатора с информацией об идентификации и/или с информацией по отслеживанию биологического объекта и/или биологического материала.

В еще одном варианте осуществления любого из вышеуказанного способа или способов, способ может дополнительно предусматривать стадию:

обеспечение возможности отслеживания биологического объекта и/или биологического материала путем считывания ДНК последовательности уникального идентификатора в биологическом объекте и/или в биологическом материале, и извлечения соответствующей записи из базы данных, обеспечивающей предоставление информации об идентификации и/или по отслеживанию биологического объекта и/или биологического материала.

Олигонуклеотидные конструкторы, кассеты, плазмиды, векторы, клетки и наборы

В другом варианте осуществления изобретения в настоящем документе предложена кассета, содержащая последовательность уникального идентификатора, где последовательность уникального идентификатора фланкирована по меньшей мере одной последовательностью 5'-праймера отжига и по меньшей мере одной последовательностью 3'-праймера отжига для амплификации ДНК последовательности уникального идентификатора, секвенирования ДНК последовательности уникального идентификатора или для того и другого.

Понятно, что в некоторых вариантах осуществления такие кассеты могут быть использованы в любом способе или способах, как описано в настоящем документе.

В некоторых вариантах кассет ДНК последовательность уникального идентификатора может быть фланкирована двумя последовательностями 5'-праймеров отжига и двумя последовательностями 3'-праймеров отжига, чтобы обеспечить амплификацию ДНК последовательности уникального идентификатора с помощью вложенной ПЦР. В некоторых вариантах осуществления может использоваться вложенный дизайн структуры, например, для улучшения точности воспроизведения. В еще одном варианте кассеты две последовательности 5'-праймеров отжига могут частично перекрываться; последовательности двух 3'-праймеров отжига могут частично перекрываться; или имеют место оба случая. В еще одном варианте выполнения кассеты она может дополнительно содержать последовательность праймера отжига для секвенирования, расположенную на 5'-конце ДНК последовательности уникального идентификатора, для секвенирования ДНК последовательности уникального идентификатора. В еще одном варианте выполнения кассеты последовательность праймера отжига для секвенирования может быть расположена между двумя последовательностями 5'-праймеров отжига. В другом варианте выполнения кассеты последовательность праймеров отжига для секвенирования может по меньшей мере частично перекрываться с одной или с обеими последовательностями 5'-праймеров отжига. В еще одном варианте выполнения кассеты две последовательности 5'-праймеров отжига могут частично перекрываться, и по меньшей мере часть последовательности секвенирующих праймеров отжига может располагаться в месте перекрытия. В еще одном варианте выполнения кассеты ее последовательность может иметь размер до около 1500 нуклеотидов

(нт) в длину; до около 1000 нт в длину; от около 200 до около 600 нт в длину; от около 200 нт до около 400 нт в длину; или от около 400 нт до около 600 нт в длину.

Вариант осуществления кассеты, описанной в настоящем документе, и пример способа ее получения показаны на фиг. 2, где кассета может быть получена с использованием рандомизированного пула последовательностей олигонуклеотидов. Рандомизированные пулы олигонуклеотидов могут быть коммерчески получены или синтезированы по желанию. Они могут быть собраны посредством ферментативной полимеризации или лигирования, или, например, химически синтезированы. Случайные фрагменты олигонуклеотидов могут быть очищены, например, разделением на колонке, для выделения фрагментов приблизительно одинакового или близкого размера (например, размером около 300-400 нуклеотидов, как показано в примере), и они могут быть вставлены в кассеты. Может быть получен пул кассет, содержащих большое количество различных последовательностей уникальных идентификаторов (т.е. около 10^7 в некоторых примерах). Кассета может содержать последовательности праймеров отжига (т.е. сайты связывания праймеров) и по меньшей мере одну последовательность праймера отжига для секвенирования (т.е. сайт связывания праймера для секвенирования) в подходящем расположении, позволяющем проводить амплификацию и/или секвенирование DUID, например, в конфигурации, показанной на фиг. 2. Сайты праймеров и секвенирования могут быть проверены на соответствие геному хозяина, чтобы убедиться в отсутствии нативной амплификации. При желании кассеты с разными праймерами можно использовать для разных организмов или для разных геномов. Кассета может содержать сайты массива рестрикционных ферментов, и она может быть предоставлена, например, в форме плазмиды или вектора-носителя кассеты для вставки. В некоторых вариантах осуществления кассета может иметь размер около 500 п.н. в длину, и она может быть представлена в плазмиде или векторе-носителе размером, например, около 1200 п.н.

Понятно, что последовательность праймеров отжига кассеты может относиться к заранее определенной последовательности или области нуклеиновой кислоты, имеющей известную нуклеотидную последовательность, так что один или более праймеров могут быть сконструированы или выбраны для отжига с такой последовательностью, чтобы например, обеспечить первичную полимеризацию полимеразой. Последовательность праймеров отжига может использоваться для амплификации последовательности уникального идентификатора, секвенирования последовательности уникального идентификатора или для того и другого.

На фиг. 13 показаны дополнительные примеры конструкций кассет, описанных в настоящем документе, включая последовательность уникального идентификатора (UID). На фиг. 13(a) показана конструкция с двойным праймером, на фиг. 13(b) показана конструкция с одним праймером, а на фиг. 13(c) показана конструкция без праймера (один UID). В конструкции кассеты для вставки двойного праймера, показанной на фиг. 13(a), вариант ее осуществления включает массив рестрикционных ферментов, 5'-область «Праймера А» и 5'-область «Праймера В» (где 5'-праймер для секвенирования может отжигаться в области, охватывающей области между областями «Праймера А» и «Праймера В»), за которым следует сайт лигирования с тупым концом. Затем следует UID-область (например, случайная ДНК с переменными п.о. или другая последовательность идентификатора), и, необязательно, может быть расположен сайт CAS 9 PAM, как показано. Далее следует сайт лигирования с тупым концом, а затем расположена 3'-область «Праймера В» и 3'-область «Праймера А», после которой следует массив рестрикционных ферментов. В конструкции кассеты для вставки с одним праймером, показанной на фиг. 13(b), вариант ее осуществления включает массив рестрикционных ферментов, 5'-область «Праймера А» (где 5'-праймер для секвенирования может отжигаться), за которым следует сайт лигирования с тупым концом. Затем следует UID-область (например,

случайная ДНК с переменными п.о. или другая последовательность идентификатора), и, необязательно, может быть расположен сайт CAS 9 PAM, как показано. Далее следует сайт лигирования с тупым концом, а затем расположена 3'-область «Праймера В», после которой следует массив рестрикционных ферментов. На фиг. 13(с) показан вариант конструкции автономной кассеты для вставки, которая включает массив рестрикционных ферментов, UID-область (например, случайную ДНК с переменными п.о. или другую последовательность идентификатора), необязательно сайт CAS 9 PAM, и массив рестрикционных ферментов, как показано.

Как показано на фиг.13, предлагается множество различных конструкций кассет. Кассеты могут различаться, например, с точки зрения присутствующих элементов, с точки зрения размера, а также с точки зрения эффективности амплификации. В зависимости от того, присутствуют ли пары праймеров (см. фиг. 13(A)-(C)), общий размер кассеты может изменяться. Например, общий размер кассеты может уменьшаться (например, в некоторых вариантах осуществления, около на 40 п.н.) по мере удаления отдельных пар праймеров. Понятно, что в некоторых вариантах осуществления изобретения эффективность амплификации UID может снижаться в результате элиминации пары праймеров. Например, для структуры с двумя праймерами, для амплификации можно использовать любую перестановку праймеров, что дает 4 возможных варианта, а не один, как было бы для структуры с одной парой праймеров. Также следует понимать, что в некоторых вариантах осуществления уменьшение размера кассеты может обеспечить, например, уменьшение возможности непреднамеренных эффектов. В некоторых вариантах осуществления можно использовать необязательный сайт CAS 9 PAM, чтобы сделать возможным эффективное редактирование последовательности UID на основе CRISPR, например, среди потомства трансформированного организма. В некоторых вариантах осуществления, где все праймеры исключены из конструкции кассеты, предполагается, что необязательно может быть представлен CAS 9 PAM, и в некоторых вариантах осуществления сайт CAS 9 PAM может позволять конструировать автономную кассету полностью из ДНК генома хозяина, например, при использовании метода расщепления/лигирования ДНК. В некоторых вариантах осуществления изобретения последовательность UID может быть переменной длины. Предполагается, что в некоторых вариантах осуществления изобретения можно безопасно использовать даже короткие последовательности UID, в частности, когда выполняется стадия проверки, которая включает, например, проверку любых коллизий между существующими в реестре UID и вновь вставленным UID.

В еще одном варианте выполнения кассет, описанных в настоящем документе, последовательности праймеров отжига в геноме целевого биологического объекта могут не встречаться в природе. Таким образом, можно уменьшить или избежать непреднамеренной и/или нецелевой амплификации и/или секвенирования.

В другом варианте осуществления изобретения в настоящем документе предложена композиция, содержащая множество из любой кассеты или кассет, описанных здесь, где каждая кассета содержит одни и те же последовательности праймеров отжига, и каждая кассета содержит рандомизированную ДНК последовательность уникального идентификатора. Такие композиции могут представлять собой пример рандомизированного пула последовательностей, как описано в настоящем документе.

В еще одном варианте осуществления изобретения в настоящем документе предложена композиция, содержащая множество из любой кассеты или кассет, описанных здесь, где каждая кассета содержит одни и те же последовательности праймеров отжига и одну и ту же последовательность секвенирующих праймеров отжига, и каждая кассета содержит рандомизированную ДНК последовательность уникального идентификатора. Такие композиции могут представлять собой пример рандомизированного пула последовательностей, как описано в настоящем документе.

В другом варианте осуществления изобретения в настоящем документе предложена плазмида, вектор экспрессии или другой одно- или двухцепочечный олигонуклеотидный конструкт, содержащий любой олигонуклеотид или олигонуклеотиды, описанные здесь, или любую кассету или кассеты, описанные в настоящем документе.

В другом варианте осуществления изобретения в настоящем документе предложена кассета, содержащая любой олигонуклеотид или олигонуклеотиды, описанные в настоящем документе.

В еще одном варианте осуществления изобретения в настоящем документе предложена клетка или вирус, которые содержат любой олигонуклеотид или олигонуклеотиды, описанные в настоящем документе, или любую кассету или кассеты, описанные в настоящем документе, включенные в геном клетки или вируса. В другом варианте осуществления изобретения в настоящем документе предложена клетка или вирус, которые содержат последовательность уникального идентификатора, встроенную в геном клетки или вируса. В другом варианте осуществления любой клетки или вируса, описанного в настоящем документе, последовательность уникального идентификатора может быть встроена в межгенную область геномной нуклеиновой кислоты клетки или вируса. В еще одном варианте осуществления любой клетки или вируса, клетка может быть клеткой растения, клеткой гриба, клеткой животного или клеткой бактерии.

В другом варианте осуществления в настоящем документе предложен набор, содержащий любой один или более из:

- ДНК последовательности уникального идентификатора;
- рандомизированного пула ДНК последовательностей уникальных идентификаторов;
- любого олигонуклеотида или олигонуклеотидов, описанных в настоящем документе;
- любой кассеты или кассет, описанных в настоящем документе;
- одного или более праймеров или пар праймеров для амплификации и/или секвенирования ДНК последовательности уникального идентификатора;
- буфера;
- полимеразы; и
- инструкции по выполнению любого из описанного здесь способа или способов;
- или любой их комбинации.

В еще одном варианте осуществления изобретения в настоящем документе предложен способ обеспечения возможности отслеживания интересующего продукта, причем указанный способ включает:

- получение или предоставление образца представляющего интерес продукта, где образец содержит геномную ДНК из части биологического материала, смешанного или связанного иным образом с представляющим интерес продукт;

- амплификацию по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК из биологического материала, и секвенирование ДНК последовательности уникального идентификатора; и

- поиск ДНК последовательности уникального идентификатора в базе данных и извлечение записи из базы данных, соответствующей ДНК последовательности уникального идентификатора, где запись в базе данных предоставляет информацию об идентификации и/или по отслеживанию представляющего интерес продукта.

В другом варианте осуществления вышеуказанного способа, способ может предусматривать введение или добавление любого биологического материала или биологических материалов, или биологического объекта или биологических объектов, как описано здесь, к представляющему интерес

продукту, биологическому материалу или объекту, содержащему по меньшей мере одну ДНК последовательность уникального идентификатора, как описано в настоящем документе, в виде части ее геномного материала.

В еще одном варианте осуществления любого из вышеуказанного способа или способов, информация об идентификации и/или по отслеживанию в записи в базе данных может содержать информацию о цепочке поставок интересующего продукта.

В еще одном варианте осуществления любого из вышеуказанного способа или способов представляющий интерес продукт может включать пищевой продукт, сельскохозяйственный продукт, фармацевтический препарат, розничный продукт, текстильный продукт, товар широкого потребления, химический продукт или другой элемент цепочки поставок.

ПРИМЕР 1 – Типовая система DUID для обеспечения возможности отслеживания пищевых продуктов

В этом примере описаны варианты выполнения типовой системы отслеживания пищевых продуктов, называемой здесь системой уникального ДНК идентификатора (DUID). В этом примере используется устойчивость и репликативная способность ДНК последовательностей для безопасного кодирования уникальных идентификаторов в ядерном геноме организма. Кодирование идентифицирующей информации в ДНК организма описанным здесь способом может обеспечить детализацию возможности отслеживания по всей цепочке поставок. В частности, система DUID может предоставить возможность следующего:

1. Безопасно обеспечить идентификацию популяции на уровне ДНК, не влияя на наследуемые признаки целевого организма;
2. Создать логические связи между DUID и справочной информацией;
3. Сократить время, затрачиваемое на отслеживание продукта до источника его происхождения, с месяцев до около одного дня;
4. Обеспечить быструю идентификацию как источника происхождения продукта, так и окончательной точки по цепочке поставок;
5. Предоставлять ценную информацию специалистам в области здравоохранения и отраслевым регулирующим органам;
6. Укреплять доверие потребителей и промышленности к стабильности, прозрачности и эффективности контроля цепочки поставок пищевых продуктов; и/или
7. Поддерживать механизмы обеспечения выполнения обязательств членами ассоциаций и поддержки интеллектуальной собственности на пищевые продукты.

Предполагается, что в некоторых вариантах осуществления изобретения система DUID может использоваться, например, для значительного расширения возможностей наблюдения за заинтересованными сторонами продовольственной системы. Описанные здесь системы DUID, помимо обеспечения возможности отслеживания, могут перевернуть традиционное представление о точке идентификации с ног на голову, и представить его как восходящее, а не нисходящее отслеживание. Такие подходы, описанные здесь, могут быть особенно желательны, учитывая, что усиление консолидации в цепочке поставок становится нормой. Системы DUID, как описано здесь, могут обеспечить практически гарантированную возможность отслеживания источника происхождения, как правило, из любой точки цепочки поставок, и при желании - в течение около одного дня. Системы могут извлекать выгоду из репликативных и стабильных клеточных свойств организма, и в результате этого предельные затраты могут приближаться к нулю по мере создания

потомства этого организма. Финансовые затраты и риск подделки и/или мошенничества для традиционных систем отслеживания довольно высоки, а юридические последствия злонамеренной деятельности могут быть значительными. Системы DUID, описанные здесь, могут быть отредактированы выгодными способами, например, таким образом, чтобы потомство популяции сохраняло части исходного идентификатора. DUID также может использоваться медицинскими работниками, если им необходимо протестировать человеческие экскременты, например, для определения недавно съеденной пищи.

Предполагается, что вышеуказанная идентификация на уровне популяции населения может дополнительно включать ссылку на юридические соглашения. Например, предполагается, что владельцы ИС на продукт могут целенаправленно связывать материал для размножения, например, с конкретным производителем и/или регионом. Генетическая идентификация на уровне популяции населения в сочетании с традиционными методами отслеживания всей цепочки может обеспечить значительный уровень контроля над движением продукта. Возьмем, например, сорт шпината, генетически модифицированный таким образом, чтобы он был устойчив к различным вредителям. Используя систему DUID, можно значительно снизить затраты на его обнаружение. В дополнение к снижению затрат система DUID может играть роль реестра, обеспечивая, например, централизованную точку контакта для отслеживания ИС.

Многие организмы подпадают под регулирующее законодательство. Например, определенный сорт растения может быть предшественником наркотика. Такие организмы в некоторых вариантах осуществления могут быть неразрывно связаны, например, с субъектами юридического права. Соответственно, предполагается, что такие случаи могут выиграть от использования описанных здесь стратегий.

В качестве примера можно рассмотреть регулирование в отношении каннабиса в Канаде. Производство и распространение растений каннабиса, также как и материала для его размножения регулируется законодательно. Предполагается, что лицензированные производители каннабиса в некоторых вариантах могут, например, включать DUID в свои продукты, что может использоваться для помощи в регулировании. В некоторых вариантах такой DUID может быть полезен для регулирования путем идентификации и/или отслеживания каннабиса, даже в сложных случаях, когда каннабис смешивается с чем-либо еще (например, в пищевых продуктах).

В качестве другого примера можно рассмотреть ассоциацию производителей шпината. В некоторых случаях для выращивания и продажи шпината может потребоваться членство в ассоциации. В некоторых вариантах предполагается, что такие материалы для размножения могут быть получены из растения с DUID. Затем могут быть проведены выборочные проверки на уровне розничной торговли, чтобы убедиться, например, что весь продаваемый шпинат аккредитован.

Система DUID:

Система DUID может охватывать, например, идентификацию продукта, проверку DUID, считывание DUID и последующее отслеживание совокупности продуктов. Она также может функционировать как центральный реестр для всех данных DUID.

В следующих примерах платформа DUID может содержать набор исполнителей, бизнес-сервис, задач, событий и систем. Исполнители могут выполнять или запускать бизнес-сервис и задачи. Системы и бизнес-сервис следует понимать с точки зрения событий, которые они производят. События могут быть напрямую связаны с состоянием следов пищевого продукта.

Исполнители: например, сотрудник службы безопасности потребителей (Исполнитель) из FDA (Исполнитель) может запросить, чтобы Платформа DUID (Исполнитель) попыталась прочитать (Бизнес-

сервис) DUID из поставляемого интересующего органического материала. Исполнители - это двигатели платформы DUID. Исполнителями могут быть системы, организации и/или отдельные лица. Они могут инициировать события и делать запросы в бизнес-сервисы. Исполнители также могут выполнять задачи. В следующем перечне приведены некоторые примеры исполнителей; однако это неполный список, который предназначен только для иллюстративных целей:

- Платформа DUID
 - Реестр DUID
 - DUID-API (интерфейс прикладной программы для DUID)
 - Химик-аналитик
 - Микробиолог
 - Секвенатор ДНК
- Производитель
 - Ботаник
 - CFO (финансовый директор)
 - Программное обеспечение для отслеживания
- Производитель сельскохозяйственных продуктов
 - Руководитель службы безопасности пищевых продуктов
 - Система планирования ресурсов предприятия
- Упаковщик/грузоотправитель
 - Водитель грузовика
 - Управляющий делами
 - СЕО (исполнительный директор)
- Розничный продавец
 - Руководитель службы безопасности пищевых продуктов
 - CFO (финансовый директор)
 - Генеральный представитель
- Государственный регулятор
 - Сотрудник службы безопасности потребителей
 - Сотрудник службы управления
- Страховщик
 - Представитель страховщика
 - Специалист по урегулированию претензий

Бизнес-сервис: Например, после аутентификации/авторизации сотрудника по безопасности потребителей (Исполнитель) и успешного завершения чтения (Бизнес-сервис) чтение (Событие) может быть зарегистрировано в реестре (Система). Бизнес-сервисы могут охватывать критически важные процессы и задачи, которые в конечном итоге могут привести к событию. Эти сервисы могут быть организованы так, чтобы они не сохраняли постоянного состояния, поскольку от них не требуется нахождения в каком-либо конкретном предшествующем состоянии для их запуска. Они могут поддерживать определенные события для успешного их завершения. В любом случае бизнес-сервис может использовать систему, но чаще всего предусматривает некоторое участие человека. Например, в некоторых вариантах действие должно быть запрошено или инициировано исполнителем. Бизнес-сервисы могут называться аналогично событию, которое они обеспечивают, например, «Проверка» (бизнес-сервис) → «Проверено» (событие).

Системы: например, после того, как чтение (Событие) было зарегистрировано в реестре (Система), потоковый процессор (Система) может прочитать вновь созданную запись чтения (Событие) из реестра, и может транслировать ее аутентифицированным/авторизованным адресатам (Система). Один из адресатов может обновить панель уведомлений, которую использует владелец бренда продукта (Исполнитель). С другой стороны, с системами могут взаимодействовать только другие системы или, иным образом, клиент, управляемый человеком. Другими словами, системы обычно могут быть цифровыми системами. Примером системы на платформе DUID может быть интерфейс прикладной программы (API). API может предоставлять интерфейс авторизованным исполнителям, которые действуют за пределами платформы. Другим примером системы может быть реестр DUID (т.е. база данных), который может функционировать как постоянное хранилище для всех данных DUID. Реестр не может напрямую подвергаться воздействию внешних исполнителей.

События: например, сотрудник службы безопасности потребителей (Исполнитель) может запросить чтение записи (Бизнес-сервис) из FDA (Исполнитель). После авторизации/аутентификации бизнес-сервис может обеспечить успешное чтение (Событие). События могут относиться к результатам действия бизнес-сервисов и систем. События обычно регистрируются в ассоциации с DUID. То есть организм может быть идентифицирован; проверен или прочитан бизнес-службой; и он отслеживается внутренними или внешними системами. В следующей таблице показано каждое событие и его связь с различными бизнес-службами, исполнителями, системами и задачами, представленными в этом примере.

Таблица 1. События и их связь с различными бизнес-сервисами, исполнителями, системами и задачами в этом примере.

<i>Идентифицирование</i>	<i>Идентифицированное</i> событие может относиться к процессу, посредством которого кассета собирается или редактируется, встраивается в геном организма и впоследствии проверяется на наличие ряда свойств.
<i>Подтверждение</i>	<i>Подтвержденное</i> событие может относиться к результату подтверждения бизнес-службы. <i>Подтверждение</i> может указывать на то, что производитель успешно трансформировал рассматриваемый организм, и можно начать регенерацию.
<i>Считывание</i>	Событие <i>считывания</i> может относиться к чтению DUID. Это может быть вызвано идентифицированным событием. Событие <i>считывания</i> может потребоваться для подтверждения <i>отслеживаемого</i> события.
<i>Отслеживание</i>	<i>Отслеживаемое</i> событие может относиться ко всем регистрационным действиям для идентифицированного организма. Оно может включать регистрацию распределения продукта получателю в цепочке поставок. <i>Отслеживаемые</i> события могут регистрироваться как подтвержденные или неподтвержденные. Подтвержденное <i>отслеживаемое</i> событие может потребовать аутентифицированного

	<p><i>считывания</i> DUID в органическом материале с использованием обычных методов секвенирования. Неподтвержденные <i>отслеживаемые</i> события могут быть зарегистрированы с использованием какой-либо метки или штрих-кода, внешнего по отношению к ДНК организма, например, идентификационная часть DUID может быть включена, например, в штрих-коды.</p>
--	--

Как описано, платформа DUID в этом примере может охватывать различных участников, бизнес-сервисы, события, системы и/или задачи. Все эти компоненты могут соответствовать определенному технологическому маршруту. В этом разделе будет подробно описан иллюстративный маршрут. На схемах, используемых для иллюстрации этих процессов, используется система обозначений согласно системе BPMN 2.0 (BPMN 2.0: <https://www.omg.org/spec/BPMN/2.0/PDF>; полностью включенной в настоящий документ посредством ссылки). Схемы показаны на фигурах, которые более подробно описаны ниже.

Обзор процесса:

На фиг. 3 показан общий вид приведенного в качестве примера процесса для экосистемы DUID этого примера.

Запуск процесса:

Перед началом этого иллюстративного процесса можно ожидать, что соответствующие соглашения были заключены в отношении условий обслуживания. Это может включать проверки «знай своего клиента» (KYC), такие как подтверждение права собственности, идентификация юридического лица и платежеспособность. В дополнение к требованиям KYC клиенты могут указать степень доступа пользователей и другие параметры системы/учетной записи через административную панель.

Создание сайта для праймеров и секвенирования:

Это может быть текущая/выполняемая задача, которая может выполняться независимо от процесса. Разработка праймеров для DUID может зависеть, например, от требований клиента к организму-хозяину, или от степени исследований и разработок, или от того и другого. Наличие пригодных для использования праймеров может быть использовано для бизнес-сервисов по идентификации.

Идентификация:

Бизнес-сервис по идентификации подробно представлен на фиг. 4. Физическим результатом этого бизнес-сервиса может быть кассета на основе ДНК последовательности, которую производитель может использовать во время трансформации организма. В рамках этой деятельности могут быть реализованы два сценария.

Во-первых, предполагается, что при наличии существующей кассеты можно использовать стандартную CRISPR и/или родственную технику для модификации частей существующего идентификатора. Например, если существующий идентификатор был сопоставлен с географическим регионом, то в конце последовательности может быть отредактировано несколько оснований. Это редактирование может быть сопоставлено с более конкретной информацией, например, с ожидаемым преобразованным состоянием после обработки. Идентифицированное событие может быть зарегистрировано после его завершения.

Если существующей кассеты нет, то ее можно создать: см. фиг. 2 для получения подробной информации о таком процессе. Как показано на фиг. 2, кассету можно получить с использованием последовательностей рандомизированного пула олигонуклеотидов. Рандомизированные пулы

олигонуклеотидов могут быть коммерчески получены или синтезированы по желанию. Их можно собрать, например, посредством ферментативной полимеризации или лигирования. Случайные фрагменты олигонуклеотидов могут быть очищены, например, разделением на колонке, с выделением фрагментов приблизительно одинакового или близкого размера (например, размером около 300-400 нуклеотидов в длину для изображенного примера), и они могут быть вставлены в кассеты. Может быть получен пул кассет, содержащих большое количество различных последовательностей уникальных идентификаторов (т.е. около 10^7 в некоторых примерах). Кассета может содержать последовательности праймеров отжига (т.е. сайты праймеров) и по меньшей мере одну последовательность праймеров отжига для секвенирования (т.е. сайт секвенирования) в подходящей конфигурации, позволяющей проводить амплификацию и/или секвенирование DUID, например, в конфигурации, показанной на фиг. 2. Сайты праймеров и секвенирования могут быть проверены на соответствие геному хозяина, чтобы убедиться в отсутствии нативной амплификации. При желании кассеты с разными праймерами можно использовать для разных организмов или для разных геномов. Кассета может содержать сайты массива рестрикционных ферментов, и она может быть предоставлена, например, в форме плазмиды-носителя кассеты вставки. В некоторых вариантах кассета может иметь размер около 500 п.н. в длину, и она может быть представлена в плазмиде или векторе-носителе размером, например, около 1200 п.н.

Как только кассета получена, можно инициировать событие идентифицирования, и кассета может быть отправлена клиенту. Клиент, как правило, является производителем, например, производителем сельскохозяйственной продукции. Производитель может использовать подходящие методы трансформации и регенерации для регенерации интересующего организма, который теперь содержит кассету, встроенную в геном. Затем они могут создать проверочный пакет, содержащий по крайней мере образец геномной ДНК из трансформированного биологического объекта, который затем может быть отправлен обратно.

Подтверждение:

Процесс подтверждения может начаться после получения проверочного пакета, аутентификации запрашивающей стороны и проверки авторизации. На фиг. 5 показан пример процесса подтверждения. Подтверждение DUID может быть выполнено для:

- Стабильной интеграции в ядерный геном хозяина.
 DUID можно легко амплифицировать из экстракта цельной ДНК.
 Последовательность DUID может быть восстановлена из кассеты DUID и в пределах предсказуемых спецификаций.
- Подтверждения уникального значения.
 Если значение уже присутствует в реестре, событие трансформации может быть отброшено.
- Определения количества копий интеграции.
 События трансформации, где имеется более одной копии DUID, могут быть отброшены (хотя также предполагается, что в некоторых примерах может использоваться более одной DUID).
- Места интеграции
 DUID может быть нацелена на некодирующие/межгенные области, чтобы уменьшить возможность вставки, влияющей на нативные кодирующие области.

Расположение DUID также может быть сопоставлено с определенной хромосомой и хромосомным плечом.

- Установления отсутствия экспрессии.

Если имеет место какая-либо экспрессия РНК для DUID, событие трансформации может не учитываться.

DUID можно амплифицировать независимо с обоими наборами праймеров (когда используется более одного набора, как показано, например, на фиг. 2), а случайную ID можно секвенировать. В некоторых вариантах осуществления этот процесс может быть повторен три раза для уменьшения ошибок секвенирования. Бизнес-сервис проверки подтверждения подлинности может использовать пошаговый процесс оценки успешного или неудачного выполнения для каждой стадии проверки подлинности кассеты. Это может снизить стоимость проверки в некоторых вариантах осуществления. Если происходит сбой, то такой результат может быть зарегистрирован. Если проверка подтверждения подлинности каждой последовательности проходит успешно, то результаты могут быть зарегистрированы, и могут начаться тесты отзыва.

В некоторых вариантах осуществления такая организация отзыва может включать введение представляющего интерес органического материала в различные среды окружения. Эти среды могут привести к различным органическим материалам, которые впоследствии могут быть переданы бизнес-сервису для считывания. В этом примере может иметь место любой из следующих четырех параллельных тестов в соответствующих средах окружения:

- Вся среда окружения свежая
- Вся среда окружения сухая
- Моделированная кислотная среда ЖКТ

Это может имитировать переваривание материала.

Это может имитировать потенциальный отзыв из анализа материала фекалий.

- Среда окружения с ионизирующим ультрафиолетовым излучением

Это может имитировать воздействие на органический материал солнечного света или других методов стерилизации пищевых продуктов, таких как гамма-облучение или стерилизация облучением электронами.

После завершения этих тестов полученный органический материал может быть независимо передан, и запущен бизнес-сервис для считывания. Все результаты после считывания бизнес-сервисом могут быть зарегистрированы. Для успешного завершения проверки подлинности не все результаты для органических материалов, полученных в результате этих испытаний по зависимости от среды окружения, должны быть успешно считаны, и это может быть сделано, в частности, для каждого конкретного случая.

Подтверждение подлинности после проведения проверки:

В зависимости от результатов проверки может быть несколько потенциальных выходов. Если проверка подлинности была несанкционированной, то сервис в отношении DUID может быть прекращен, и заинтересованные стороны могут быть уведомлены. Если один из тестов проверки подлинности последовательности не пройден, то может быть проведено подтверждение подлинности после проведения проверки. Подтверждение подлинности после проведения проверки может позволить определить причину сбоя. В зависимости от этой причины может быть два результата процесса (ошибка кассеты или ошибка трансформации): можно либо инициировать повторную попытку бизнес-службы по идентификации, либо запросить повторную попытку трансформации у производителя.

Если результатом проверки бизнес-сервисом подлинности является подтвержденное событие, то реестр DUID (т.е. база данных) может быть обновлен соответствующей информацией. Это событие также может инициировать сообщение или уведомление об утверждении размножения, которое может быть получено производителем. Затем можно перейти к производству материала для размножения производителем, который, в свою очередь, может продолжать свою деятельность в обычном режиме.

Предварительное считывание цепочки поставок:

Как описано в настоящем документе, остальная часть цепочки поставок может продолжать работу в обычном режиме. Тем не менее, участники цепочки поставок могут иметь возможность интегрировать DUID в свои существующие процессы. Если они решат этого не делать, то наличие DUID может обеспечить по меньшей мере отслеживание источника происхождения. В некоторых вариантах осуществления предполагается, что DUID может быть интегрирована в существующие штрих-коды. Следует обратить внимание, что в некоторых вариантах осуществления часть уникального идентификатора (UID) DUID может быть по существу строкой символов, охарактеризованной нуклеотидами (A, T, G, C). В некоторых вариантах осуществления, если явное считывание не требуется, они могут независимо отслеживать готовый к DUID организм, используя свои собственные технологии сбора данных (например, штрих-кодирование). Это может привести к неподтвержденному отслеживаемому событию.

Если требуется событие считывания, заинтересованная сторона может отправить запрос в бизнес-службу считывания. В этом примере может быть два типа запросов. Один тип может быть обязательным, а другой может быть добровольным. Содержимое пакета считывания может зависеть от типа запроса. Например, если запрос на считывание является обязательным, то могут быть определенные требования, которые необходимо выполнить для удовлетворения требований заинтересованных сторон, например, предоставления образцов органических материалов, собранных в определенные даты.

Считывание:

Бизнес-сервис считывания подробно показан на фиг. 6. Как и в случае с другими бизнес-сервисами, авторизация может быть немедленно проверена. Часто пакет считывания может содержать различные типы органического материала. В зависимости от этого материала может быть выполнена очистка и/или амплификация. Если праймеры обнаружены, то можно начинать секвенирование (а в некоторых случаях и стадии декодирования UID). Если праймер не обнаружен, то регистрируют результаты и завершают процесс как неудачный.

После секвенирования и/или декодирования UID может быть предпринята попытка найти все соответствующие данные в реестре DUID. Вполне возможно, что в некоторых случаях DUID может быть не найдена в реестре, и в этом случае может быть проведено подтверждение подлинности после проведения проверки. Этот обзор данных может позволить найти причину ошибки. С другой стороны, если DUID найдена, то результаты могут быть зарегистрированы и может быть создано событие считывания.

Также возможно, что в некоторых вариантах осуществления одобренный ассоциированный партнер, например FDA, может сделать запрос в бизнес-сервис считывания. В некоторых юрисдикциях могут действовать правила, требующие, например, обмена данными о возможности отслеживания.

Пакет данных после считывания

После завершения считывания бизнес-сервисов может быть сгенерирован пакет данных считывания и возвращен запрашивающей заинтересованной стороне. Пакет считывания может содержать все предыдущие отслеженные события, результаты проверки и данные для начинающих. Он также может содержать

договорные обязательства, которые в первую очередь требуют использования DUID. Это может включать информацию KYC для каждой вовлеченной стороны.

Вспомогательные системы:

В этом примере на диаграмме глобального представления DUID могут быть отмечены две вспомогательные системы. Ни одна из них не может играть неотъемлемую роль в общем процессе, но вместо этого они могут функционировать как интерфейсы и процессоры для реестра DUID.

API: API может функционировать как интерфейс к реестру DUID. Это может позволить одобренным ассоциированным сторонам получить доступ к утвержденным данным. В некоторых случаях они могут изменять эти данные - см. степень доступа пользователей, как описано выше.

Потоковый процессор: потоковый процессор может считывать данные из реестра в режиме реального времени и в результате этого запускать функции. Например, если неавторизованный субъект запросил бизнес-услугу считывания, то, например, владелец DUID может быть автоматически уведомлен.

Таким образом в этом примере подробно описаны варианты системы DUID, способы и композиции, которые можно использовать в соответствии с представленными здесь идеями. Понятно, что этот пример представлен для иллюстративных целей, и он предназначен для специалистов в данной области, а не для ограничения.

ПРИМЕР 2 – Стабильная интеграция DUID в виды дрожжей

Стабильная интеграция DUID в виды дрожжей:

В этом примере описана стабильная интеграция DUID в виды дрожжей.

Методы и материалы для стабильной интеграции DUID в виды дрожжей.

Обзор:

В этом примере описаны подходы к конструированию, интеграции и проверке уникальных идентификаторов на основе последовательностей ДНК (DUID) в модельном организме, в частности, в дрожжах. Эти методы включают использование как лабораторных штаммов дрожжей, так и промышленных штаммов дрожжей. Описанные здесь методы подтверждают полезность и эффективность интеграции DUID в геном для обеспечения активности по отслеживанию. Эти лабораторные методы молекулярной биологии включают:

1. Конструирование *in silico* DUID, DUID векторов и DUID праймеров.
2. Метод стабильной интеграции в геном с помощью центромерных дрожжевых плазмид (YCr).
3. Метод стабильной интеграции в геном путем встраивания в нативные хромосомы дрожжей.
4. Метод подтверждения интеграции DUID.
5. Метод обнаружения сигнала от DUID и пределы обнаружения сигнала.

Предполагается, что эти методы применимы к широкому кругу лабораторных и промышленных штаммов дрожжей, включая прототрофные штаммы. Подход YCr позволяет интегрировать конструкции DUID в геном путем клеточного и ядерного управления в виде независимых хромосом посредством ассоциации веретена центромерных последовательностей, встроенных в каркас вектора. Для встраивания в нативные хромосомы дрожжей выбраны четыре геномных сайта для минимального вмешательства в нормальную кодирующую способность и экспрессию генов в геноме. Эти сайты включают субтеломерные области, которые обычно считаются гетерохроматическими, где гены обычно молчат, и эухроматиновую область с низкой кодирующей способностью, которая действует как положительный контроль. Подход путем вставки в нативные хромосомы дрожжей фокусируется на: 1) совместной трансформации плазмиды, несущей

устойчивость к антибиотикам, для отбора трансформантов вместе с линейным фрагментом, содержащим DUID, фланкированным гомологичными областями, фланкирующими выбранные сайты-мишени; и 2) методы на основе CRISPR, нацеленные на сайт интеграции с использованием специфических гидовых РНК (гРНК) и специфических матриц восстановления гомологии (HRT), которые служат в качестве матриц для сайтов-мишеней PAM, расщепленных Cas9.

Конструирование и создание конструкторов и векторов:

Конструирование конструкторов DUID

На фиг. 14 показаны карты двух конструкторов DUID размером 370 п.о. А) Конструкция конструктора DUID для амплификации ПЦР и количественной ПЦР. Конструктор имеет 370 п.о. Этот конструктор DUID включает 2 прямых праймера и два обратных праймера. В нем имеется два идентификатора (ID1 и ID2). ID1 идеально подходит для ПЦР-амплификации. ID2 идеально подходит для амплификации методом количественной ПЦР (qPCR). В) Конструкция конструктора DUID для петлевой изотермической амплификации (LAMP) и ПЦР. Эта карта включает праймеры как для ПЦР, так и для LAMP. Помимо традиционных решений по конструированию структур для амплификации, следует обратить внимание на структуры, выделенные розовым цветом, которые являются дополнительными сайтами CAS PAM, и они позволяют редактировать и обнаруживать последовательности конструкторов DUID с использованием систем на основе CRISPR. Такой сайт PAM может позволить редактировать интегрированную конструкцию DUID.

На фиг. 17 показан пример сопоставления идентификатора (ID) с реестром, как описано здесь. Следует обратить внимание, что на этой фигуре показан упрощенный пример, и предполагается, что все последовательности DUID обычно не будут такими короткими, как показано в таблице.

В данном примере нет более одного выравнивания последовательности ID в базе данных. В этом примере последовательности ID всегда уникальны для одного конструктора DUID, но в одном конструкторе DUID может быть несколько последовательностей ID. Однако внутри последовательности ID может быть один или более участков, гомологичных другим последовательностям DUID. Предполагается, что в конструкторе DUID могут быть последовательности, которые можно использовать везде в конструкторах DUID; однако сами идентификаторы должны быть уникальными, и, соответственно, DUID также будут уникальными. Это решение в отношении конструирования, а именно, иметь гомологичные участки в последовательностях ID для любого количества DUID в конструкторах, может позволить создавать варианты DUID несколькими способами.

Пример секции гомологичного идентификатора - одна гомологичная секция для трех DUID:

- Есть несколько причин, по которым желательно иметь гомологичные последовательности для нескольких идентификаторов. В некоторых случаях идентификатор может иметь гомологичную последовательность с целью предоставления варианта, связанного с этим идентификатором. Возможность использования вариантов идентификаторов может позволить пользователям ссылаться на связанный протокол, который будет информировать, как они могут взаимодействовать с DUID. Например, конкретный вариант идентификатора DUID может содержать внутри открытый ключ в контексте криптографии, который может каким-то осмысленным образом информировать о последующих взаимодействиях с DUID. В других случаях гомологичная последовательность может ссылаться, например, на систему или объект, который первоначально создал или имеет идентификатор. В следующей таблице показаны три DUID, имеющие такой гомологичный участок: в этих иллюстративных примерах DUID участок 1:10 является гомологичным, и участок 11:50 является уникальным.

Денатурация	98	10 с	30
Отжиг	60	20 с	30
Удлинение	72	15 с	30
Окончательное удлинение	72	10 мин	1
Выдержка	4	Время выдержки	Цикл выдержки

Проверяли праймеры, и температуры отжига оптимизировали для реакционных объемов 20 мкл. Для создания фрагментов линейной интеграции HR проводили реакции в объеме 2×50 мкл, и продукты очищали с использованием набора Qiagen PCR Purification Kit (<https://www.qiagen.com/ie/shop/pcr/qiaquick-pcr-purification-kit/>). Очищенные фрагменты ДНК элюировали в 50 мкл элюирующего буфера (10 mM трис-Cl, pH 8,5). Продукты проверяли, пропуская 5 мкл через 1% агарозный гель.

Вектор CRISPR и генерация матриц восстановления гомологии (HRT):

Эксперименты CRISPR проводили с использованием плазмиды pCC-036, которая содержит CAS9, экспрессируемый *TDH3p*, *SNR52p* для управления экспрессией гидовой РНК (гРНК) и *hygR* для селекции на гигромицине, как описано в Kroeger et al., 2019. Три гРНК были сконструированы для каждого из целевых сайтов интеграции с помощью программного обеспечения Benchling (<https://www.benchling.com/>). Праймеры, содержащие последовательности гРНК (Приложение В), использовали в реакциях ПЦР, где в качестве матрицы использовали pCC-036. Реакционные компоненты и условия реакции приведены ниже (Таблица 3 и Таблица 4). Эти реакции ПЦР трансформировали в *E. coli*. Плазмиды выделяли из трансформантов и их подвергали скринингу путем секвенирования для подтверждения кооректности клонов (фиг. 14 и Приложение В). Конструировали один клон гРНК для Chr6 (Chr6_2) и два для Euch (Euch_1; Euch_2). Праймеры были сконструированы так, чтобы они частично перекрывались (~8-10 пар оснований, не перекрывающихся с каждой стороны) с мутацией в середине обоих праймеров, и ПЦР проводили в соответствии с протоколом, приведенном в Zheng, et al., 2004.

Таблица 4: Реакционная смесь ПЦР для вставки гРНК и мутаций сайта РАМ – полимеразы высокой точности Phusion

Компонент	Объем (20 мкл)	Объем (50 мкл)	Окончательная концентрация
5-X буфер Phusion	4 мкл	10 мкл	1X
10 mM dNTP	0,4 мкл	1 мкл	200 мкМ, каждый
Прямой праймер	1 мкл	2,5 мкл	0,5 мкМ
Обратный праймер	1 мкл	2,5 мкл	0,5 мкМ
Матрица	1 мкл	1 мкл	25 нг
Полимераза Phusion	0,2 мкл	0,5 мкл	0,02 ЕД/мкл
Вода	12,4 мкл	32,5 мкл	

Таблица 5: Условия амплификации

Стадия	Температура	Время	Циклы
--------	-------------	-------	-------

Начальная денатурация	98	30 с	1
Денатурация	98	10 с	16
Отжиг	55	20 с	16
Удлинение	68	24 мин	16
Окончательное удлинение	68	1 час	1
Выдержка	4	Время выдержки	Цикл выдержки

Проверяли праймеры, и температуру отжига оптимизировали для используемых реакционных объемов 20 мкл; для интеграции проводили реакции в объеме 5×50 мкл с последующим расщеплением вектора с HindIII и BamHI (NEB). ДНК очищали с использованием фенола/хлороформа/изоамилового спирта с последующим осаждением этанолом в присутствии 0,1 М ацетата аммония и гликогена. ДНК ресуспендировали в 30 мкл воды, свободной от нуклеаз. Амплификацию проверяли, пропуская 5 мкл через 1% агарозный гель.

Таблица 6: Реакционная смесь для вставки последовательностей гРНК

Компонент	Объем на реакцию
Буфер HF	10 мкл
Матричная ДНК	1 мкл
F-праймер	2,5 мкл
R-праймер	2,5 мкл
dNTP	1 мкл
Полимераза Phusion	1 мкл
Вода	32 мкл

Таблица 7: Условия амплификации ПЦР

Стадия	Температура	Время
Горячий старт	94	Время выдержки
Начальная денатурация	94	3 мин
*Денатурация	94	1 мин
*Отжиг	52	1 мин
*Удлинение	68	24 мин
Окончательное удлинение	68	1 час
Конец	24	Время выдержки

* 16 циклов денатурация/отжиг/удлинение

После проведения ПЦР 10 мкл помещали в 1% агарозный гель (выход ДНК в SDM с Phusion низкий). Расщепление 10 мкл ПЦР-ампликона с DpnI (NEB) в реакционном объеме 30 мкл проводили в течение ночи при 37°C для линеаризации метилированной матричной ДНК. Еще 10 мкл разделяли на геле, а затем 5 мкл трансформировали в *E. coli*. Минипрепараты получали на 12 колониях и последовательностях.

Трансформация дрожжей:

Трансформация векторов YCr-DUID

Стандартный протокол трансформации дрожжей на основе ацетата лития, как описано в Mertenes et al. 2017, использовали для трансформации как плазмиды CRISPR, так и для матрицы восстановления в целевые штаммы, и выполняли, как описано ниже.

1. Дрожжи выращивали в течение ночи в 100 мл 2% питательной среды YPD при 30°C до OD ~ 0,7-0,8.
2. Затем культуру дрожжевых клеток центрифугировали (3 минуты при 3000 об/мин), однократно промывали стерильной водой и клетки ресуспендировали в 200 мкл 0,1 М раствора ацетата лития.
3. После 10-минутной инкубации при комнатной температуре 50 мкл клеточной культуры смешивали с 500 нг плазмиды, 300 мкл PLI (142 М полиэтиленгликоля, 0,12 М ацетата лития, 0,01 М Трис (pH 7,5) и 0,001 М EDTA) и 5 мкл ДНК спермы лосося (1 мг/мл). Параллельно проводили трансформацию отрицательного контроля, не содержащего ДНК (стерильная вода).
4. Суспензию дрожжей инкубировали 30 минут при 42°C.
5. Клетки центрифугировали (3 минуты при 3000 об/мин) и ресуспендировали в свежей 2% среде YPD, после чего клетки выдерживали для инкубации в течение одной ночи при 30°C.
6. 200 мкл суспензии дрожжей высевали на YPD + G418 300 мкг/мл с последующей 2-дневной инкубацией при 30°C. 200 мкл также высевали на YPD, не содержащий антибиотика, для подтверждения жизнеспособности клеток после этих обработок.

Котрансформация

Этот метод включал получение компетентных клеток с помощью ацетата лития с последующей трансформацией ДНК с помощью электропорации, как описано в Bernardi et al., 2019.

Таблица 8: Обобщение результатов трансформации

Штамм	Вектор	Линейный конструктор DUID
S288C (BY4743; 2n)	pRS41K	S288C_Euch_Chr2
		S288C_Chr6
Vermont(пивные дрожжи; 4n)	pRS41K	Verm_Chr2
		Verm_Chr6
French Saison(дрожжи для эля; 4n)	pRS41K	FrenSais_Chr2
		FrenSais_Chr6

Стадии котрансформации:

1. Клетки выращивали в 100 мл среды YPD при встряхивании до желаемой фазы роста (на основе кривых роста или OD).
2. Клетки собирали в середине логарифмического роста ($OD_{600} = 0,7-0,8$). Культуру центрифугировали, а супернатант отбрасывали.
3. Осадок однократно промывали стерильной водой. Культуру центрифугировали, супернатант отбрасывали и ресуспендировали в 25 мл раствора 0,1 М ацетата лития/10 мМ DTT/10 мМ TE (трис-HCl:EDTA = 10:1). Культуру инкубировали в течение 1 часа при комнатной температуре. Культуру центрифугировали, и супернатант отбрасывали.
4. Примечание. При работе с флокулянтами необходимо обязательно переворачивать пробирки несколько раз каждые 10 минут, чтобы клетки не оседали на дно пробирки.

5. Осадок промывали 25 мл ледяной дистиллированной стерильной воды, культуру центрифугировали при 4°C и удаляли надосадочную жидкость. Стадию повторяли (всего две промывки).
6. Осадок промывали 10 мл охлажденного льдом сорбита, центрифугировали при 4°C и удаляли надосадочную жидкость. Ресуспендировали осадок в 100 мкл ледяного сорбита.
7. Для трансформации использовали 100 мкл клеточной суспензии.
8. Смешивали 15 мкл (1 мкг pRS41K [плазида YCr] + 1 мкг линейного фрагмента DUID; молярное соотношение 1:10; и молярное соотношение 1:20) трансформирующей ДНК с клеточной суспензией и инкубировали на льду в течение 5 минут.
9. Суспензии клеток подвергали электропорации при 1,8 кВ в кюветах диаметром 0,1 см.
10. В кювету для электропорации добавляли 1 мл холодного сорбита и смешивали с клеточной суспензией. Суспензию переносили в пробирку с 300 мкл YPD.
11. Примечание. При использовании маркера-антибиотика необходимо инкубировать суспензию в течение 3 часов при 30°C, чтобы обеспечить экспрессию антибиотика. * - в эту культуру не добавляют антибиотик, поскольку это приведет к гибели всех клеток, так как они еще не экспрессируют плазмиду, обеспечивающую устойчивость к антибиотикам.*
12. 100 мкл трансформированной культуры высевали в чашки на селективные среды (YPD + 300 мг/л G418), и инкубировали при 30°C в течение 5 дней до появления колоний.
13. Трансформированные культуры также высевали на чашки с YPD без каких-либо маркеров/антибиотиков, чтобы убедиться, что клетки являются живыми.

Трансформация векторами CRISPR и HRT:

Для трансформации как плазмиды CRISPR, так и матрицы восстановления в целевые штаммы использовали стандартный протокол трансформации дрожжей на основе ацетата лития, как описано в Mertens, et al., 2019. Этот протокол, описанный ниже, основан на стандартных процедурах трансформации, при которых клетки становятся компетентными путем обработки раствором LiOAc, после чего клетки инкубировали с молекулами ДНК (плазмидой и матрицей восстановления) и ДНК-носителем (ДНК спермы лосося) перед тепловым шоком для встраивания ДНК. После восстановления клетки высевали на гигромицин для отбора от всех нетрансформированных клеток. Посев на YPD без гигромицина показал рост клеток после процедуры трансформации; в частности, сама процедура не приводила к гибели клеток. Трансформация плазмиды CRISPR без HRT должна привести к гибели клеток, поскольку DSB не восстанавливается; это подтверждает успешное функционирование плазмиды CRISPR, что означает, что Cas9 экспрессируется, а гРНК нацеливают Cas9 на геном. Трансформация плазмиды CRISPR вместе с HRT должна восстанавливать DSB и поддерживать рост клеток.

Плазмиды pCC-036_Chrr6_2/Chrr6_HRT и pCC-036_Euch_1/Euch-HRT представляют собой соответствующие комбинации молекул ДНК, трансформированных в штаммы дрожжей S288c, Vermont и French Saison. Был использован следующий протокол.

1. Дрожжи выращивали в течение ночи в 5 мл среды YPD при 30°C, 200 об/мин, после чего 1 мл прекультуры переносили в 50 мл YPD и инкубировали еще 4 часа (30°C, 200 об/мин).
2. Затем культуру дрожжевых клеток центрифугировали (3 минуты при 3000 об/мин) и клетки ресуспендировали в 200 мкл 0,1 М раствора ацетата лития.
3. После 10-минутной инкубации при комнатной температуре 50 мкл клеточной культуры смешивали с 500 нг плазмиды, в которой клонировали соответствующую одиночную гидовую РНК с 5-25 мкг

(скорректированный протокол) ДНК HRT и без нее, 300 мкл PLI (142 М полиэтиленгликоля, 0,12 М ацетата лития, 0,01 М трис (рН 7,5) и 0,001 М EDTA) и 5 мкл ДНК спермы лосося (1 мг/мл).

4. Инкубировали 30 минут при 42°C.
5. Клетки центрифугировали (3 минуты при 3000 об/мин) и ресуспендировали в свежей среде YPD, после чего клетки восстанавливали в течение одной ночи при 30°C.
6. 200 мкл суспензии дрожжей высевали на YPD, содержащей гигромицин в концентрации 300 мг/л, с последующей инкубацией в течение 3-5 дней при 30°C.

Скрининг трансформантов:

Протокол выделения геномной ДНК

1. Реплицировали чашки для совместной трансформации на YPD+G418 (300 мг/л).
2. Разделяли мастер-чашки на 4-8 колоний на срез. Соскабливали колонии в стерильную пробирку с 3 мл YPD и выращивали в течение ночи при встряхивании при 30°C.
3. Получали осадок 2 мл культуры в 2 мл пробирке с завинчивающейся крышкой.
4. Промывали один раз 1 мл воды MQ.
5. Ресуспендировали в 200 мкл буфера для разрушения (2% TX-100, 1% SDS, 100 mM NaCl, 100 mM Трис, рН 7,5).
6. Добавляли 200 мкл стеклянных шариков и 200 мкл смеси фенол/хлороформ/изоамиловый спирт.
7. Подвергали вихревому перемешиванию на высокой скорости в течение 3 мин.
8. Центрифугировали при максимальных оборотах в течение 5 мин.
9. Переносили верхний водный слой в чистую микроцентрифужную пробирку.
10. Добавляли 1 мл 100% EtOH и перемешивали путем переворачивания.
11. Центрифугировали при максимальных оборотах в течение 3 мин.
12. Декантировали этанолом, высушивали осадок и ресуспендировали в 400 мкл 1X TE.
13. Добавляли 30 мкл 1 мг/мл РНКазы А.
14. Инкубировали 5 мин при 37°C.
15. Добавляли 10 мкл 4М ацетата аммония и 1 мл 100% EtOH. Смешивали с инверсией.
16. Центрифугировали 3 мин при максимальных оборотах. Промывали осадок в 1 мл 70% EtOH, и давали высохнуть.
17. Ресуспендировали в 100 мкл воды.

Идентификация интегрантов и ПЦР-скрининг трансформантов:

гДНК, выделенная, как описано выше, служила матрицей для реакций ПЦР с использованием праймеров, которые связывают геномную ДНК в определенных областях, расположенных выше и ниже гомологичных областей HRT, которые фланкируют сайт мишени интеграции (подробности в отношении праймеров см. в Приложении А; состав компонентов и условия реакции указаны ниже в Таблицах 9 и 10). Для эухроматического сайта мишени интеграции на Chr2 использовали праймеры Euch_Seq F/R, а для субтеломерного гетерохроматинового сайта мишени интеграции Chr6 использовали праймеры Chr6_Seq F/R. Эти праймеры давали фрагмент ДНК ~600 п.н. из гДНК без какой-либо вставки в сайте интеграции. При интеграции размер этого фрагмента увеличивался до ~970 п.н. Контроли включали реакции без какой-либо матрицы гДНК и гДНК, выделенной из нетрансформированных штаммов (например, S288c/BY4743). Реакции

ПЦР контролировали с помощью гель-электрофореза с использованием маркера молекулярной массы GeneRuler 100bp Plus для подтверждения размера получаемых фрагментов ДНК.

Как только было подтверждено получение интегрантов, правильную интеграцию подтверждали с помощью ПЦР-праймеров, которые связываются с геномом вне интегрирующегося фрагмента, и другого ПЦР-праймера, который связывается внутри трансформированного фрагмента. Эти праймеры дают фрагмент ДНК, если интеграция произошла в правильном сайте-мишени, и не дают фрагмента ДНК, если интеграция не произошла.

Полученные фрагменты ДНК, подтвержденные с помощью тестов подтверждения интеграции, а также с помощью проверочных анализов, представляют собой последовательности для подтверждения интеграции.

Таблица 9: ПЦР-реакция для скрининга DUID

Компонент	Количество для реакции 20 мкл
Стандартный буфер	4 мкл
Матричная ДНК	~ 100-400 нг (1 мкл)
F-праймер	1 мкл
R-праймер	1 мкл
dNTP	0,4 мкл
Полимераза OneTaq HotStart	0,1 мкл
Вода	12,6 мкл

Таблица 10: Программа ПЦР-скрининга DUID

Стадия	Температура	Время
Начальная денатурация	98	3 мин
* Денатурация	98	30 с
*Отжиг	60	30 с
*Удлинение	68	1 мин
Окончательное удлинение	68	10 мин
Конец	12	Время выдержки

* выполнено 30 циклов денатурация/отжиг/удлинение

После ПЦР 10 мкл реакционной смеси разделяли на 1% агарозном геле (1ХТАЕ, с красителем для нуклеиновой кислоты SYBR Safe).

Подтверждение количества копий и места вставки:

Выполняли полногеномное секвенирование (WGS) для родительских штаммов и интегрантов, чтобы подтвердить количество копии вставки и определить наличие любых нецелевых событий интеграции. Для этого объединяли данные секвенирования с коротким чтением (Illumina) и длинным чтением (PacBio), чтобы собрать полные геномы как родительских, так и трансформированных штаммов. Комбинация этих двух подходов позволила получить общую структуру генома интегранта и, следовательно, определить, произошли ли множественные вставки или могли ли быть нецелевые интеграционные события. Полные геномы интегрантов и родительских штаммов секвенировали с помощью Genome Québec (Монреаль, Канада), как описано ранее (Preiss et al., 2018). Короче говоря, выделяли ДНК и использовали ее в качестве матриц для создания библиотек для методов Illumina и PacBio. Результаты чтения секвенирования подвергали

качественному анализу с помощью программы FastQC (версия 0.11.5) (Andrews, 2010), а также обрезали и фильтровали с помощью программы Trimmomatic (версия 0.36) (Bolger, Lohse, & Usadel, 2014). Результаты чтения сопоставляли с эталонным геномом *S. cerevisiae* S288c (R64-2-1) с использованием SpeedSeq (0.1.0) (Chiang et al., 2015). Качество выравнивания оценивали с помощью QualiMap (2.2.1) (Garcia-Alcalde et al., 2012). Анализ вариантов выполняли на выровненных результатах чтения с использованием FreeBayes (1.1.0-46-g8d2b3a01) (Garrison & Marth, 2012). Варианты для всех штаммов вызывали одновременно (мультиэмпл). Перед анализом вариантов выравнивания выполняли фильтрацию до минимального MAPQ 50 с помощью SAMtools (1.2) (Li et al., 2009). Аннотирование и предсказание эффекта вариантов выполняли с использованием SnpEff (1.2) (Cingolani et al., 2012). Вариации количества копий хромосом и генов оценивали на основе охвата Control-FREEC (11.0) (Boeva et al., 2012). Статистически значимые вариации количества копий идентифицировали с использованием критерия суммы рангов Уилкоксона ($p < 0,05$). Среднее значение размера охвата вставки и количество гетерозиготных SNP в окнах размером 10000 п.н. рассчитывали с помощью BEDTools (2.26.0) (Quinlan & Hall, 2010) и визуализировали в R.

Определение экспрессии DUID в интегрантах с помощью капельной цифровой ПЦР:

Использовали капельную цифровую ПЦР (ddPCR), которая позволяет количественно определить абсолютное количество молекул в образце. Это, в частности, позволяет проводить количественную оценку количества копий или генов с низкой экспрессией. Процедура включала выделение из дрожжей РНК, свободной от гДНК, с последующим синтезом кДНК и, в конечном итоге, получение S288c с плазмидой pRS41K-Euch и без нее, при этом штамм интегрантов выращивали в YPD в трех повторностях. РНК экстрагировали с помощью обычно используемого фенольного метода с горячей кислотой (COLLART AND OLIVIERO, 2001) и ее количественно определяли с помощью спектрофотометра NanoDrop 2000C (NanoDrop Technologies Inc.). Образцы РНК обрабатывали с использованием набора для удаления ДНК RapidOut (Thermo Fisher), проверяли ДНК на загрязнение, и оценивали ее качество с использованием биоанализатора Agilent 2100. РНК (1000 нг на образец) использовали для создания кДНК с использованием набора для обратной транскрипции кДНК высокой емкости (Applied BioSystems).

Эти образцы вместе с разбавленной pRS41K-Euch представляли для анализа методом ddPCR в Центре геномики Университета Гвельфа. Эти образцы вместе с «контролем без матрицы» использовали в качестве матриц в реакциях ddPCR с использованием ddPCR EvaGreen Supermix (с эмульгированием) во всех реакциях вместе с праймерами для qPCR для DUID, GAT3 (низкий уровень экспрессии, контроль) и ACT1 (высокий уровень экспрессии, контроль). Капли наноразмера получали на приборе AutoDGTM (Bio-Rad), затем проводили ПЦР-амплификацию с использованием термоциклера C1000 Touch (Bio-Rad). После циклов ПЦР планшет для ddPCR считывали в устройстве для считывания капель QX200 от Bio-Rad, и данные анализировали с использованием программного обеспечения QuantaSoft Analysis Pro версии 1.0.596 (Bio-Rad Laboratories).

Протокол анализа LOD/LOQ:

гДНК, приготовленную с использованием протокола выделения гДНК, использовали для скрининга вставки. Векторы готовили с использованием набора QiaQuick Miniprep из культур DH5 α K12, выращенных в присутствии ампициллина,.

Стандартный протокол ПЦР

Праймеры: F- и R-S288C DUID

Серия разбавлений: 100 нг, 10 нг, 1 нг, 100 мкг, 10 мкг, 1 мкг, 100 фг, 10 фг, 1 фг, 100 аг.

Таблица 11: Реакции ПЦР, проведенные с использованием полимеразы GoTaq в реакциях объемом 20 мкл

Компонент	Количество для реакции 20 мкл
2X зеленый ММ	10 мкл
Матричная ДНК	На серию разбавления
F-праймер	1 мкл
R-праймер	1 мкл
Вода	7 мкл

Таблица 12: Условия реакции ПЦР

Стадия	Температура	Время
Начальная денатурация	98	3 мин
*Денатурация	98	30 с
*Отжиг	60	30 с
*Удлинение	72	30 с
Окончательное удлинение	72	10 мин
Конец	12	Время выдержки

*30 циклов денатурация/отжиг/удлинение

После проведения ПЦР 10 мкл реакционной смеси разделяли на 1% агарозном геле (1XTAE, с красителем для нуклеиновой кислоты SYBR Safe).

Протокол количественной ПЦР (qPCR):

Реакции количественной ПЦР проводили в Центре геномики Университета Гвельфа с использованием SensiFAST Hi-ROX SYBR Master Mix в системе для ПЦР в реальном времени StepOnePlus. Условия циклов количественной ПЦР показаны в таблице 12. Анализ проводили с использованием программного обеспечения Applied Biosystems StepOnePlus. гДНК получали с использованием протокола выделения гДНК, описанного выше. Контрольный вектор DUID готовили с использованием набора QiaQuick Miniprep из культур DH5α K12, выращенных в присутствии ампициллина.

Аmplификацию проводили как на образцах плазмиды, так и на образцах гДНК дрожжей YCr с использованием следующего праймера и в серии разбавлений.

Праймеры: F- и R- S288C DUID для qPCR

Серия разбавлений: 100 нг, 10 нг, 1 нг, 100 мкг, 10 мкг, 1 мкг, 100 фг, 10 фг, 1 фг, 100 аг.

Результаты и обсуждение

Проверка трансформации:

DUID стабильно трансформировали в геном штамма дрожжей (BY4743) с помощью вектора YCr. Трансформированные дрожжи культивировали и экстрагировали геномную ДНК, как описано выше. Стабильную интеграцию (Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* fly strain (w1118; iso-2; iso-3, Austin. 2012 Apr-Jun; 6(2):80-92. doi:

10.4161/fly.19695. PMID: 22728672; PMCID: PMC3679285.on) подтверждали с помощью ПЦР с конечной точкой (фиг. 15, В1-В3) и количественной ПЦР (фиг. 16). Анализ амплификации методом ПЦР с конечной точкой, с использованием праймеров отзыва DUID, которые фланкируют конструкт DUID, дал положительную амплификацию как для вектора YCr-DUID (фиг. 15А), так и для экстрактов геномной ДНК из клеток дрожжей, трансформированных векторами YCr-DUID (фиг. 15С). Полоса 370 п.н., указывающая на амплификацию DUID, четко видна, когда в качестве матрицы использовали количества 100 мкг-100 нг вектора YCr-DUID (фиг. 15А, дорожки 1-4), при этом отсутствовала заметная амплификация с нетрансформированной геномной ДНК BY4743 при любом количестве введенной ДНК (фиг. 15В, дорожки 1-8). Аналогичные анализы с использованием общей ДНК, выделенной из клеток, трансформированных вектором YCr-DUID, привели к положительной амплификации в диапазоне 1-100 нг вводимой ДНК с очень слабым сигналом от 100 мкг вводимой ДНК (фиг. 15С, дорожки 1-4), что указывает на то, что DUID присутствует в количестве 1-2 копий на клетку, и при этом количество копий, отражающее количество копий хромосомных признаков, может быть легко обнаружено в изолятах гДНК дрожжей с использованием стандартных процедур ПЦР с конечной точкой.

На фиг. 15 показано обнаружение YCr-DUID в геномной ДНК дрожжей с помощью ПЦР с конечной точкой. ПЦР-амплификацию проводили с использованием (А) вектора YCr-DUID, (В) гДНК, выделенной из BY4743 и (С) штамма дрожжей BY4743, трансформированного вектором YCr-DUID, в качестве матриц с праймерами отзыва DUID. Реакции проводили с использованием последовательно разбавленной ДНК-матрицы с входными количествами (1) 100 нг, (2) 10 нг, (3) 1 нг, (4) 100 мкг, (5) 10 мкг, (6) 1 мкг, (7) 100 мкг и (8) 10 фг, и разделение выполняли на 1% агарозном геле с GeneRuler™ 100bp Plus Ready-to-use Ladder в качестве стандарта.

Анализ по оценке LOD/LOQ (предел обнаружения/предел количественного определения):

Количественную ПЦР в реальном времени проводили с использованием серийных 10-кратных разбавлений очищенного вектора YCr-DUID (фиг. 16); в этих анализах амплификацию DUID обнаружили при всех измеренных концентрациях, что указывает на то, что DUID можно надежно идентифицировать при таких низких концентрациях, как 500 мкг. Стандартную кривую строили путем нанесения средних значений C_q в зависимости от известных входных концентраций ДНК с использованием MS Excel. На основании этой стандартной кривой вычислили значение R^2 как 0,9993, при эффективности праймера 105,5% (рассчитано с использованием калькулятора Agilent QPCR Standard Curve to Slope Efficiency, https://www.chem.agilent.com/store/biocalculators/calcSlopeEfficiency.jsp?_requestid=1116919), что указывает на то, что эффективность реакции находится в пределах принятых стандартов для высококачественного анализа количественной ПЦР (<https://www.gene-quantification.de/roche-rel-quant.pdf>). Аналогичный анализ количественной ПЦР, проведенный с использованием ДНК, выделенной из штамма BY4743, трансформированного с помощью YCr-DUID, показал, что DUID может быть обнаружена в пределах 50 нг общей дрожжевой ДНК со средним значением C_q 29,02. Это значение C_q наносили на график относительно стандартной кривой, описанной выше (фиг. 3; оранжевая полоса). Эти результаты подтверждают, что методы отзыва DUID могут амплифицировать DUID из матрицы культуры клеток дрожжей.

Эти результаты продемонстрировали:

1. DUID можно успешно создавать и стабильно трансформировать в дрожжах.
2. С целью возможности отслеживания DUID они могут быть извлечены из биологической матрицы с помощью как стандартных методов ПЦР с конечной точкой, так и методов количественной ПЦР.

На фиг. 16 показано обнаружение DUID общей ДНК дрожжей в экстрактах. Количественную ПЦР в реальном времени проводили на серийных 10-кратных разбавлениях вектора YCr в диапазоне от 50 нг до 500 аг, и строили стандартную кривую (синяя линия) с использованием MS Excel. Результаты аналогичного эксперимента с количественной ПЦР с использованием ДНК, полученной из штамма BY4743, трансформированного вектором YCr-DUID, наносили на график (оранжевая полоса) и сопоставляли со значениями стандартной кривой для определения количественного обнаружения DUID в биомассе дрожжей.

Приложение А: Праймеры, используемые для создания фрагментов линейной трансформации или отзыва DUID

Праймер	Последовательность (5'-3')	Применение
Chr6_DUID F	CATTCCGCCTGACCTGGAG (SEQ ID NO: 5)	Синтез линейных фрагментов для котрансформации из фрагментов Twist
Euch_DUID F	CATTCCGCCTGACCCCTTAAT (SEQ ID NO: 6)	Синтез линейных фрагментов для котрансформации из фрагментов Twist
DUID_synth R	CACTGAGCCTCCACCTAGC (SEQ ID NO: 7)	Синтез линейных фрагментов для котрансформации из фрагментов Twist
Chr6_Seq F	AAGCGTAATCCGAAAGGCA (SEQ ID NO: 8)	Фланкирующий праймер Chr6: связывает 5' генома сайта интеграции
Chr6_Seq R	TGCATACGCTTCTCTCGACT (SEQ ID NO: 9)	Фланкирующий праймер Chr6: связывает геном 3' сайта интеграции
Euch_Seq F	CAGAAATGGACAAGGAGATATGTGA (SEQ ID NO: 10)	Фланкирующий праймер Euch: связывает 5' генома сайта интеграции.
Euch_Seq R	TTGAGTACCTGGCCAATGGAG (гшв1)	Фланкирующий праймер Euch: связывает 3' генома сайта интеграции

S288C_recall F	GCTGATGGTTT TAGGCGTACA (SEQ ID NO: 12)	Отзыв S288c DUID
S288C_recall R	CCCTGGAAATGCACTTGGTC (SEQ ID NO: 13)	
S288C_qPCR F	TGGTCGTTTGGCTGTAGAGA (SEQ ID NO: 14)	Количественная ПЦР S288c DUID
S288C_qPCR R	CGTATAGAGCGGGTCATCGA (SEQ ID NO: 15)	
Verm_recall F	ACTCTCCATTAGTCGGCAG (SEQ ID NO: 16)	Отзыв Vermont DUID
Verm_recall R	AAGACCGCTTGTCCGACA (SEQ ID NO: 17)	
Verm_qPCR F	GGCCSTATCAGTACAGCAGT (SEQ ID NO: 18)	Количественная ПЦР Vermont DUID
Verm_qPCR R	AGTGCTGGCGAGAGAATGAA (SEQ ID NO: 19)	
FrenSais_recall F	GCGTACAATGCCCTGAAGAA (SEQ ID NO: 20)	Отзыв French Saison DUID
FrenSais_recall R	CTCCTGGAAATGCACTTGG (SEQ ID NO: 21)	
FrenSais_qPCR F	AGCGGGTCATCGAAAGGTTA (SEQ ID NO: 22)	Количественная ПЦР French Saison DUID
FrenSais_qPCR R	CACTGGTTCGTTTGGCTGTA (SEQ ID NO: 23)	

Приложение В: Праймеры CRISPR, используемые для клонирования гРНК или мутации сайтов PAM в HRT

Праймер	Последовательность (5'-3')	Применение
Chr6_1 F	AGTTGCAAAAAACAAGGGAAGTTTTAGAGCTAG AAATAGCAAGTAAAATAAGG (SEQ ID NO: 24)	Вставка гРНК в pCC-036: Chr 6_1
Chr6_1 R	TCCCTGTTTTTGCAACTGATCATTTATCTTTCACT GCGGAG (SEQ ID NO: 25)	
Chr6_2 F	GAGATCTTGTTTTATCATGAGTTTTAGAGCTAGAA ATAGCAAGTAAAATAAGG (SEQ ID NO: 26)	Вставка гРНК в pCC-036: Chr 6_2
Chr6_2 R	CATGATAAAACAAGATCTCGATCATTTATCTTTCA CTGCGGAG (SEQ ID NO: 27)	

Chr6_3 F	AGATCTTGTTTTATCATGAGGTTTTAGAGCTAGAA ATAGCAAGTTAAAATAAGG (SEQ ID NO: 28)	Вставка гРНК в pCC-036: Chr 6_3
Chr6_3 R	TCATGATAAAACAAGATCTGATCATTATCTTTCA CTGCGGAG (SEQ ID NO: 29)	
Euch_1 F	ATACTAAGTCAACATCAAGGGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGG (SEQ ID NO: 30)	Вставка гРНК в pCC-036: Euch_1
Euch_1 R	CCTTGATGTTGACTTAGTATGATCATTATCTTTTC ACTGCGGAG (SEQ ID NO: 31)	
Euch_2 F	GAAATACTAAGTCAACATCAGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGG (SEQ ID NO: 32)	Вставка гРНК в pCC-036: Euch_2
Euch_2 R	TGATGTTGACTTAGTATTTTCGATCATTATCTTTTC ACTGCGGAG (SEQ ID NO: 33)	
Euch_3 F	TCTTGGCTTTTACAACCGAGGTTTTAGAGCTAGA AATAGCAAGTTAAAATAAGG (SEQ ID NO: 34)	Вставка гРНК в pCC-036: Euch_3
Euch_3 R	CTCGGTTGTAAAAGCCAAGAGATCATTATCTTTTC ACTGCGGAG (SEQ ID NO: 35)	
Chr6_1 mut F	CAAGGGAAACGAAAATCAATCAAATTAG (SEQ ID NO: 36)	Мутирование сайта РАМ в HRT для Chr6_1
Chr6_1 mut R	GATTGATTTTCGTTTCCCTTGTTT (SEQ ID NO: 37)	
Chr6_2_3 mut R	CCTCCCACTAGCCTCCGCTCATGATAAA (SEQ ID NO: 38)	Мутирование сайта РАМ в HRT для Chr6_2 и Chr6_3
Chr6_Vector R	GCTGCAGCCCCTCATGATAA (SEQ ID NO: 39)	
Euch_1 mut F	GAGGAAATACTAAGTCAACATCAAGGTCGCA (SEQ ID NO: 40)	Мутирование сайта РАМ в HRT для Euch_1
Euch_1 mut R	TGCGACCTTGATGTTGACTTAGTATTTCTCTCGG (SEQ ID NO: 41)	
Euch_2 mut F	СТАAGTCACATCAACGTGGCA (SEQ ID NO: 42)	Мутирование сайта РАМ в HRT для Euch_2

Euch_2 mut R	TGCCACGTTGATGTTGACTTAGTATTCC (SEQ ID NO: 43)	
Euch_3 mut F	TACAACCCGAGACGAAATACTAAGTCAACATC (SEQ ID NO: 44)	Мутирование сайта РАМ в HRT для Euch_3
Euch_3 mut R	CTTAGTATTTTCGTCTCGGTTGTAAAAGCCA (SEQ ID NO: 45)	
Euch_Vector R	GGGCTGCAGTCAGCAGAT (SEQ ID NO: 46)	

Приложение С: Гомологичные рекомбинационные конструкты

Полные последовательности, включая гомологичные плечи, для использования в гомологичной рекомбинации

Сайты хромосомной интеграции дрожжей

ChrII:809650..809799

Левая гомология:

ACACAAACTGGCGTAGAAGGGGAAACGGAAATAGGGTCTGACGAGGAAGATAGCATAGAGG
ACGAGGGAAGCAGC (SEQ ID NO: 47)

Правая гомология:

AGTGGAGGAAATAGTACGACAGAAAGACTAGTACCACACCAGCTGAGGGAACAAGCAGCCA
GACATATAGGAAAA (SEQ ID NO: 48)

ChrVI:261123..261272

Левая гомология:

TGGAGTTGCAAAAAACAAGGGAAAGGAAAATCAATCAAATTAGAATTAAGGTTTTTTTTGGA
CAGTGCAGCGTCA (SEQ ID NO: 49)

Правая гомология:

ATGCGCACGTAATGGCTTCGAAGAAAAAAGAAGGCAAATACAATGAAGCTGAGATCCTTGTT
TTATCATGAGGGG (SEQ ID NO: 50)

ChrXIV:764201..764350

Левая гомология:

CAAATAAATTAGGCTCATAACCGTAATTTTTCGAGACATTTTTGGTTACTTCAAATTTGTTAT
TATATAAA (SEQ ID NO: 51)

Правая гомология:

GATCATATAAAGTTCTTGGACAAGATTGGATACATTTAGTTTTATTTTTGAAAATCAAAAAGAT
GAAACAAAATA (SEQ ID NO: 52)

Приложение D: Компоненты конструктов и длины

Общий размер	520 о.
Левое плечо гомологии	75 о.
Прямой праймер 1	20 о.

ID1	210 п.о. (включая рам, фланкирующий на 3')
Прямой праймер 2	20 о.
ID1	80 п.о. (включая рам, фланкирующий на 3')
Обратный праймер 2	20 о.
Обратный праймер 1	20 о.
Правое плечо гомологии	75 о.

S288c

ID1 (праймеры для ПЦР)

ОЛИГО	начало	длина	Тпл.	GC %	любой	3'-й	шпилька	последовательность
ЛЕВЫЙ-ПРАЙМЕР	66	20	57,98	50	0,00	0,00	0,00	GCTGATGGT TTAGGCGTCA (SEQ ID NO: 53)
ПРАВЫЙ-ПРАЙМЕР	437	20	59,11	55	0,00	0,00	0,00	CCCTGGAAA TGCACTTGGTC (SEQ ID NO: 54)

ID2 (праймеры для количественной ПЦР)

ОЛИГО	начало	длина	Тпл.	GC %	любой	3'-й	шпилька	последовательность
ЛЕВЫЙ-ПРАЙМЕР	301	20	58,85	55	10,15	0,00	0,00	CGTATAGAG CGGGTCATC GA (SEQ ID NO: 55)
ПРАВЫЙ-ПРАЙМЕР	422	20	58,67	50	0,00	0,00	0,00	TGGTCGTTT GGCTGTAGA GA (SEQ ID NO: 56)

s288c - Хромосома 2

ACACAAACTGGCGTAGAAGGGGAAACGGAAATAGGGTCTGACGAGGAAGATAGCATAGAGG
ACGAGGGAAGCAGCGCTGATGGTTTAGGCGTACACGAGATCCTGGTTCAACGCGCTGCAAACCTACCC
TGCTCCAAACTGCTGTTCAACGCCACTCTAACTGGCAGGCAAATTATTAGTTTCTAAGTTCCCCAGGTG
CTGAAGAGCAGTCATTC AACGCCCTCAGATCATCCCGCAAGTTGGCTGGCGCGTTTGTCCGGAGGAT
CGTGTCGTACAACAACCATCTGACTATCAACCCTCCaggCGTATAGAGCGGGTCATCGATGCGCTCAGG

GAACAACAACGATAGGCCTGCGGCTGGTCACCATCGGGAAGTTTTGCTGGAGATCTGCTGCTGTAGGaggTCTCTACAGCCAAACGACCAGACCAAGTGCATTTCCAGGGAGTGGAGGAAATAGTACGACAGAAAGACTAGTACCACACCAGCTGAGGGAACAAGCAGCCAGACATATAGGAAAA (SEQ ID NO: 57)

s288c - Хромосома 6

TGGAGTTGCAAAAAACAAGGGAAGGAAAATCAATCAAATTAGAATTAAGGTTTTTTTTTGGACAGTGCAGCGTCAGCTGATGGTTTAGGCGTACACGAGATCCTGGTTCAACGCGCTGCAAACCTACCCTGCTCCAAACTGCTGTTCAACGCCACTCTAACTGGCAGGCAAATTATTAGTTTCTAAGTTCCCCAGGTGCTGAAGAGCAGTCATTCAACGCCCTCAGATCATCCCGGCAAGTTGGCTGGCGCGTTTGTCCGGAGGATCGTGTCGTACAACAACCATCTGACTATCAACCCTCCaggCGTATAGAGCGGGTCATCGATGCGCTCAGGGAACAACAACGATAGGCCTGCGGCTGGTCACCATCGGGAAGTTTTGCTGGAGATCTGCTGCTGTAGGaggTCTCTACAGCCAAACGACCAGACCAAGTGCATTTCCAGGGATGCGCACGTAATGGCTTCGAAGAAAAAAAGAAGGCAAATACAATGAAGCTGAGATCTTGTTTTATCATGAGGGG (SEQ ID NO: 58)

s288c - Хромосома 14

CAAATAAATTAGGCTCATAACCGTAATTTTATTCGAGACATTTTTGGTTACTTCAAATATTGTATTATATAAAGCTGATGGTTTAGGCGTACACGAGATCCTGGTTCAACGCGCTGCAAACCTACCCTGTCCAAACTGCTGTTCAACGCCACTCTAACTGGCAGGCAAATTATTAGTTTCTAAGTTCCCCAGGTGCTGAAGAGCAGTCATTCAACGCCCTCAGATCATCCCGGCAAGTTGGCTGGCGCGTTTGTCCGGAGGATCGTGTCGTACAACAACCATCTGACTATCAACCCTCCaggCGTATAGAGCGGGTCATCGATGCGCTCAGGGAAACAACAACGATAGGCCTGCGGCTGGTCACCATCGGGAAGTTTTGCTGGAGATCTGCTGCTGTAGGaggTCTCTACAGCCAAACGACCAGACCAAGTGCATTTCCAGGGGATCATATAAAGTTCTTGGACAAGATTGGATACATTTAGTTTTATTTTTGAAAATCACAAGATGAAACAAAATA (SEQ ID NO: 59)

Vermont

ID1 (ПЦР)

ОЛИГО	начало	длина	Тпл.	GC%	любой	3'-й	шпилька	последовательность
ЛЕВЫЙ-ПРАЙМЕР	173	20	58,88	55,00	0,00	0,00	0,00	ACTCTCCC ATTAGTTCG GCAG (SEQ ID NO: 60)
ПРАВЫЙ-ПРАЙМЕР	541	20	60,18	50,00	0,00	0,00	0,00	AAGACCGC TTTGTTC GACA (SEQ ID NO: 61)

ID2 (праймеры для количественной ПЦР)

ОЛИГО	начало	длина	Тпл.	GC%	любой	3'-й	шпилька	последовательность
-------	--------	-------	------	-----	-------	------	---------	--------------------

ЛЕВЫЙ- ПРАЙМЕР	353	20	58,88	55,00	0,00	0,00	0,00	GGCCCTAT CAGTACAG CAGT (SEQ ID NO: 62)
ПРАВЫЙ- ПРАЙМЕР	470	20	59,39	50,00	0,00	0,00	0,00	AGTGCTGG CGAGAGAA TGAA (SEQ ID NO: 63)

Vermont – Хромосома 2

ACACAAACTGGCGTAGAAGGGGAAACGGAAATAGGGTCTGACGAGGAAGATAGCATAGAGG
ACGAGGGAAGCAGCACTCTCCATTAGTCGGCAGCACGTTCCGAGTAATACCGGAGACAGAAAAAT
CTCGGAACAGTTTATCCGCAATTCTGAGGAAATCGTCGTCCGCAAGCTCCGTGCACAGCTAGTAGTAG
TCTCCGGTGCGGGGGGGGGGCGGAGTGGTCTCCACGATACGACGTTGTCTAGATACGTACCCACCTCG
CTGTGTGCTCTCTGGCTATCTGAACGTCCACTCCAGAaggGGCCCTATCAGTACAGCAGTCATAGCCGC
ACACAAGTCCAACGTCCCCAAACCTCCTGACCACGCAGTCGCCACCGGCGCAGACACTATTTCTCGTg
ggTTCATTCTCTCGCCAGCACTTGTTCGGAACAAAGCGGTCTTAGTGGAGGAAATAGTACGACAGAAAG
ACTAGTACCACACCAGCTGAGGGAACAAGCAGCCAGACATATAGGAAAA (SEQ ID NO: 64)

Vermont – хромосома 6

TGGAGTTGCAAAAACAAGGGAAAGGAAAATCAATCAAATTAGAATTAAGGTTTTTTTTTGGGA
CAGTGCAGCGTCAACTCTCCATTAGTCGGCAGCACGTTCCGAGTAATTACCGGAGACAGAAAAATC
TCGGAACAGTTTATCCGCAATTCTGAGGAAATCGTCGTCCGCAAGCTCCGTGCACAGCTAGTAGTAGT
CTCCGGTGCGGGGGGGGGGCGGAGTGGTCTCCACGATACGACGTTGTCTAGATACGTACCCACCTCGC
TGTGTGCTCTCTGGCTATCTGAACGTCCACTCCAGAaggGGCCCTATCAGTACAGCAGTCATAGCCGCA
CACAAGTCCAACGTCCCCAAACCTCCTGACCACGCAGTCGCCACCGGCGCAGACACTATTTCTCGTag
gTTCATTCTCTCGCCAGCACTTGTTCGGAACAAAGCGGTCTTATGCGCACGTAATGGCTTCGAAGAAAA
AAAGAAGGCAAATACAATGAAGCTGAGATCTTGTTTTATCATGAGGGG (SEQ ID NO: 65)

Vermont – хромосома 14

CAAATAAATTAGGCTCATAACCGTAATTTTATTCGAGACATTTTTGGTTACTTCAAAATATTGT
TATPATATAAACTCTCCATTAGTCGGCAGCACGTTCCGAGTAATTACCGGAGACAGAAAAATCTC
GGAACAGTTTATCCGCAATTCTGAGGAAATCGTCGTCCGCAAGCTCCGTGCACAGCTAGTAGTAGTCT
CCGGTGCGGGGGGGGGGCGGAGTGGTCTCCACGATACGACGTTGTCTAGATACGTACCCACCTCGCTG
TGTGCTCTCTGGCTATCTGAACGTCCACTCCAGAaggGGCCCTATCAGTACAGCAGTCATAGCCGCACA
CAAGTCCAACGTCCCCAAACCTCCTGACCACGCAGTCGCCACCGGCGCAGACACTATTTCTCGTaggT
TCATTCTCTCGCCAGCACTTGTTCGGAACAAAGCGGTCTTGATCATATAAAGTTCTTGGACAAGATTGG
ATACATTTAGTTTTATTTTTGAAAATCACAAAGATGAAACAAAATA (SEQ ID NO: 66)

French Saison

ID1 (праймеры для ПЦР)

ОЛИГО	начало	длина	Тпл.	GC%	любой	3'-й	шпилька	последовательность
ЛЕВЫЙ-ПРАЙМЕР	79	20	58,55	50,00	0,00	0,00	0,00	GCGTACAA TGCCCCTG AAGAA (SEQ ID NO: 67)
ПРАВЫЙ-ПРАЙМЕР	439	20	58,82	55,00	0,00	0,00	0,00	CTCCTGGA AATGCACT TGG (SEQ ID NO: 68)

ID2 (праймеры для количественной ПЦР)

ОЛИГО	начало	длина	Тпл.	GC%	любой	3'-й	шпилька	последовательность
ЛЕВЫЙ-ПРАЙМЕР	308	20	59,10	50,00	0,00	0,00	0,00	AGCGGGTC ATCGAAAG GTТА (SEQ ID NO: 69)
ПРАВЫЙ-ПРАЙМЕР	426	20	58,41	50,00	0,00	0,00	0,00	CACTTGGT CGTTTGGC TGТА (SEQ ID NO: 70)

French – Хромосома 2

ACACAAACTGGCGTAGAAGGGGAAACGGAAATAGGGTCTGACGAGGAAGATAGCATAGAGG
ACGAGGGAAGCAGCGCTACAATGCCCTGAAGAATTACTTCCGTAAGCGGATAGCACCAGAC
TGTAAGCTAACGAACGCCTGTTTGAGGCTCAGTCTGCTAAATTGGAACCGCGTCGCTCCTAGGCATAT
TTTGGTGAAGCACTCTGCCAAAAGCCTGTAGAATTCCGGACCGACGCTCTCTCACTCGAAGATTC
CGGGTAAGAAGTTTCAGCCAGGGCTGTCTCCATTAGAAaggAGCGGGTCATCGAAAGGTTACGTTGGTT
GTATCTGATTAGACGGTAGACATCCAGCTCATCTCTGATTAATAAGTTCTCCGCCGCTCCATCGGGCG
aggTACAGCCAAACGACCAAGTGCCAAGTGCATTTCCAGGGAGAGTGGAGGAAATAGTACGACAGAAA
GACTAGTACCACACCAGCTGAGGGAACAAGCAGCCAGACATATAGGAAAA (SEQ ID NO: 71)

French – Хромосома 6

TGGAGTTGCAAAAAACAAGGGAAAGGAAAATCAATCAAATTAGAATTAAGGTTTTTTTTTGG
CAGTGCAGCGTCAGCGTACAATGCCCTGAAGAATTACTTCCGTAAGCGGATAGCACCAGACTG
TAAGCTAACGAACGCCTGTTTGAGGCTCAGTCTGCTAAATTGGAACCGCGTCGCTCCTAGGCATATTTT

GGTGAAAGCACTCTGCCAAAAGCCTGTAGAATTCCGGACCGACGCTCTCTTCACTCGAAGATTCCGG
 GTAAGAAGTTTCAGCCAGGGCTGTCTCCATTAGAAaggAGCGGGTCATCGAAAGGTTACGTTGGTTGTA
 TCTGATTAGACGGTAGACATCCAGCTCATCTCTGATTACTAAAGTTCTCCGCCGCTCCATCGGGCGaggT
 ACAGCCAAACGACCAAGTGCCAAGTGCATTTCCAGGGAGATGCGCACGTAATGGCTTCGAAGAAAA
 AAGAAGGCAAATACAATGAAGCTGAGATCTTGTTTTATCATGAGGGG (SEQ ID NO: 72)

French – Хромосома 14

CAAATAAATTAGGCTCATAACCGTAATTTTATTCGAGACATTTTTGGTTACTTCAAAATATTGT
 TATTATATAAAGCGTACAATGCCCTGAAGAATTACTTCCGTAAGCGGATAGCACCAGACTGTA
 AGCTAACGAACGCCTGTTTGAGGCTCAGTCTGCTAAATTGGAACCGCGTCGCTCCTAGGCATATTTTG
 GTGAAAGCACTCTGCCAAAAGCCTGTAGAATTCCGGACCGACGCTCTCTTCACTCGAAGATTCCGGG
 TAAGAAGTTTCAGCCAGGGCTGTCTCCATTAGAAaggAGCGGGTCATCGAAAGGTTACGTTGGTTGTAT
 CTGATTAGACGGTAGACATCCAGCTCATCTCTGATTACTAAAGTTCTCCGCCGCTCCATCGGGCGaggTA
 CAGCCAAACGACCAAGTGCCAAGTGCATTTCCAGGGAGGATCATATAAAGTTCTTGGACAAGATTGG
 ATACATTTAGTTTTATTTTTGAAAATCACAAAGATGAAACAAAATA (SEQ ID NO: 73)

Здесь в качестве примера описан один или более иллюстративных вариантов осуществления изобретения. Специалистам в данной области техники понятно, что ряд изменений и модификаций может быть выполнен без отклонения от объема и сущности изобретения, определенного формулой изобретения.

Цитированные документы

FDA. (2019). Investigation Summary: Factors Potentially Contributing to the Contamination of Romaine Lettuce Implicated in the Fall 2018 Multi-State Outbreak of *E. coli* O157:H7. Получено из: <https://www.fda.gov/media/120722/download>

Food and Drug Regulations C.R.C. c. 870 (2019).

GS1 US. (2013). Integrated Traceability in Fresh Foods: Ripe Opportunity for Real Results. Получено с https://www.gs1us.org/DesktopModules/Bring2mind/DMX/Download.aspx?Command=Core_Download&EntryId=598.

Introduction of Organisms and Products Altered or Produced Through Genetic Engineering Which Are Plant Pests or Which There Is Reason to Believe Are Plant Pests C.F.R. §340.1 (2019).

WHO, Foodborne Disease Burden Epidemiology Reference Group. 2015. WHO Estimates of the Global Burden of Foodborne Diseases. Получено с <https://academicanswers.waldenu.edu/faq/73164>.

System of centromeric, episomal, and integrative vectors based on drug resistance markers for *Saccharomyces cerevisiae*. Christof Taxis and Michael Knop EMBL, Heidelberg, Germany. *BioTechniques* 40:73-78 (January 2006) doi 10.2144/000112040

Krogerus, K., Magalhães, F., Kuivanen, J. et al. A deletion in the STA1 promoter determines maltotriose and starch utilization in STA1+ *Saccharomyces cerevisiae* strains. *Appl Microbiol Biotechnol* 103, 7597–7615 (2019). <https://doi.org/10.1007/s00253-019-10021-y>

Zheng L, Baumann U, Reymond JL. An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* 2004;32(14):e115. Опубликовано 2004 Aug 10, doi:10.1093/nar/gnh110

Mertens S, Steensels J, G. B, V Kevin J. Rapid Screening Method for Phenolic Off-Flavor (POF) Production in Yeast. *J Am Soc Brew Chem.* 2017; 75(4):318–23

Mertens S, Gallone B, Steensels J, et al. Reducing phenolic off-flavors through CRISPR-based gene editing of the FDC1 gene in *Saccharomyces cerevisiae* x *Saccharomyces eubayanus* hybrid lager beer yeasts [исправленная публикация в: *PLoS One.* 2019 Oct 24;14(10):e0224525]. *PLoS One.* 2019;14(1):e0209124. Опубликовано: 2019 Jan 9. doi:10.1371/journal.pone.0209124

Mertens S, Gallone B, Steensels J, et al. Correction: Reducing phenolic off-flavors through CRISPR-based gene editing of the FDC1 gene in *Saccharomyces cerevisiae* x *Saccharomyces eubayanus* hybrid lager beer yeasts. *PLoS One.* 2019;14(10):e0224525. Опубликовано 2019 Oct 24. doi:10.1371/journal.pone.0224525

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014 Aug 1;30(15):2114-20. doi: 10.1093/bioinformatics/btu170. Epub 2014 Apr 1. PMID: 24695404; PMCID: PMC4103590.

Chiang C, Layer RM, Faust GG, Lindberg MR, Rose DB, Garrison EP, Marth GT, Quinlan AR, Hall IM. SpeedSeq: ultra-fast personal genome analysis and interpretation. *Nat Methods.* 2015 Oct;12(10):966-8. doi: 10.1038/nmeth.3505. Epub 2015 Aug 10. PMID: 26258291; PMCID: PMC4589466.

García-Alcalde F, Okonechnikov K, Carbonell J, Cruz LM, Götz S, Tarazona S, Dopazo J, Meyer TF, Conesa A. Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics.* 2012 Oct 15;28(20):2678-9. doi: 10.1093/bioinformatics/bts503. Epub 2012 Aug 22. PMID: 22914218.

Erik Garrison and Gabor Marth 2012. Haplotype-based variant detection from short-read sequencing

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009 Aug 15;25(16):2078-9. doi: 10.1093/bioinformatics/btp352. Epub 2009 Jun 8. PMID: 19505943; PMCID: PMC2723002

Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)*. 2012 Apr-Jun;6(2):80-92. doi: 10.4161/fly.19695. PMID: 22728672; PMCID: PMC3679285.

Boeva V, Popova T, Bleakley K, et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics*. 2012;28(3):423-425. doi:10.1093/bioinformatics/btr670

Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010 Mar 15;26(6):841-2. doi: 10.1093/bioinformatics/btq033. Epub 2010 Jan 28. PMID: 20110278; PMCID: PMC2832824

Все цитированные документы, представленные здесь и где-либо в описании, полностью включены в настоящее описание посредством ссылки.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ идентификации биологического материала, включающий:
получение или предоставление образца, содержащего геномную ДНК, из биологического материала;
амплификацию по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК из биологического материала, и секвенирование ДНК последовательности уникального идентификатора; и
поиск ДНК последовательности уникального идентификатора в базе данных, и извлечение записи из базы данных, соответствующей ДНК последовательности уникального идентификатора, где запись в базе данных предоставляет информацию об идентификации и/или по отслеживанию биологического материала.
2. Способ по п. 1, в котором биологический материал содержит материал растительного происхождения, материал на основе гриба, материал животного происхождения, материал на основе вируса или материал на основе бактерии.
3. Способ обеспечения возможности отслеживания биологического материала, включающий:
определение последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК биологического объекта;
подтверждение идентификации биологического объекта посредством: проверки наличия ДНК последовательности уникального идентификатора в геномной ДНК; и сравнение ДНК последовательности уникального идентификатора с базой данных для подтверждения того, что ДНК последовательность уникального идентификатора еще не используется в базе данных;
предоставление указания о соответствии получения биологического материала из биологического объекта, где биологический материал содержит геномную ДНК из биологического объекта; и
ввод последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в запись базы данных, и связывание ДНК последовательности уникального идентификатора с информацией об идентификации и/или по отслеживанию биологического материала;
обеспечивая тем самым возможность отслеживания биологического материала путем считывания ДНК последовательности уникального идентификатора в биологическом материале, и извлечения соответствующей записи из базы данных, предоставляющей информацию об идентификации и/или по отслеживанию биологического материала.
4. Способ по п. 3, дополнительно включающий вставку по меньшей мере одной ДНК последовательности уникального идентификатора в геномную ДНК биологического объекта или модификацию уже существующей последовательности идентификатора в геномной ДНК биологического объекта путем редактирования гена для создания ДНК последовательности уникального идентификатора в геномной ДНК биологического объекта, обеспечивая тем самым его идентификацию.
5. Способ по п. 4, дополнительно включающий предоставление по меньшей мере одной ДНК последовательности уникального идентификатора для вставки в геномную ДНК биологического объекта.
6. Способ по любому из пп. 3-5, в котором биологический материал содержит материал растительного происхождения, материал на основе гриба, материал животного происхождения, материал на основе вируса или материал на основе бактерии.
7. Способ по любому из пп. 3-6, в котором биологический объект содержит клетку растения, клетку гриба, клетку животного, вирус или бактериальную клетку.

8. Способ по любому из пп. 3-7, в котором получение биологического материала из биологического объекта включает размножение биологического объекта.

9. Способ по любому из пп. 3-8, в котором ДНК последовательность уникального идентификатора происходит из рандомизированного пула ДНК последовательностей уникального идентификатора.

10. Способ по любому из пп. 3-9, в котором считывание ДНК последовательности уникального идентификатора в биологическом материале и извлечение соответствующей записи из базы данных предусматривает:

получение или предоставление образца, содержащего геномную ДНК, из биологического материала; амплификацию по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК из биологического материала, и секвенирование ДНК последовательности уникального идентификатора; и

сравнение ДНК последовательности уникального идентификатора с базой данных, и извлечение записи из базы данных, соответствующей ДНК последовательности уникального идентификатора, где запись в базе данных предоставляет информацию об идентификации и/или по отслеживанию биологического материала.

11. Способ по любому из пп. 1-10, в котором ДНК последовательность уникального идентификатора содержит уникальную нуклеотидную последовательность, встроенную в межгенный участок геномной ДНК.

12. Способ по любому из пп. 1-11, в котором ДНК последовательность уникального идентификатора содержит последовательность размером до около 1500 нуклеотидов (нт) в длину; до около 1000 нт в длину; от около 200 до около 600 нт в длину; от около 200 нт до около 400 нт в длину; или от около 400 нт до около 600 нт в длину.

13. Способ по любому из пп. 1-12, в котором ДНК последовательность уникального идентификатора фланкирована одной или более последовательностями праймеров отжига для ПЦР-амплификации ДНК последовательности уникального идентификатора, секвенирования ДНК последовательности уникального идентификатора или для того и другого.

14. Способ по любому из пп. 1-13, в котором биологический материал включает пищевой продукт.

15. Способ по любому из пп. 1-14, в котором информация об идентификации и/или по отслеживанию в записи в базе данных содержит информацию о цепочке поставок биологического материала.

16. Способ по любому из пп. 1-15, в котором информация об идентификации и/или по отслеживанию в записи в базе данных содержит информацию об источнике происхождения биологического материала.

17. Способ по любому из пп. 1-16, в котором информация об идентификации и/или по отслеживанию в записи в базе данных включает информацию о производителе, регионе, партии, лоте, дате или другую соответствующую информацию о цепочке поставок или любые их комбинации.

18. Способ по любому из пп. 1-17, в котором кассета включена в геномную ДНК, при этом кассета содержит ДНК последовательность уникального идентификатора, фланкированную одной или более последовательностями праймеров отжига для ПЦР-амплификации ДНК последовательности уникального идентификатора, секвенирования ДНК последовательности уникального идентификатора или для того и другого.

19. Способ по любому из пп. 1-18, в котором ДНК последовательность уникального идентификатора представляет собой случайную последовательность, полученную из рандомизированного пула последовательностей нуклеиновых кислот размером до около 1500 нуклеотидов (нт) в длину; до около 1000

нт в длину; от около 200 до около 600 нт в длину; от около 200 нт до около 400 нт в длину; или от около 400 нт до около 600 нт в длину.

20. Олигонуклеотид, содержащий ДНК последовательность уникального идентификатора, фланкированную одной или более последовательностями праймеров отжига для ПЦР-амплификации ДНК последовательности уникального идентификатора, секвенирования ДНК последовательности уникального идентификатора или для того и другого.

21. Олигонуклеотид по п. 20, в котором ДНК последовательность уникального идентификатора содержит случайную последовательность размером до около 1500 нуклеотидов (нт) в длину; до около 1000 нт в длину; от около 200 до около 600 нт в длину; от около 200 нт до около 400 нт в длину; или от около 400 нт до около 600 нт в длину.

22. Кассета, содержащая олигонуклеотид по п. 20 или 21.

23. Клетка или вирус, содержащие олигонуклеотид по п. 20 или 21, или кассету по п. 22, встроенную в геном клетки или вируса.

24. Клетка или вирус, содержащие ДНК последовательность уникального идентификатора, встроенную в геном клетки или вируса.

25. Клетка или вирус по п. 23 или 24, где ДНК последовательность уникального идентификатора встроена в межгенную область геномной ДНК клетки или вируса.

26. Клетка или вирус по любому из пп. 23-25, где клетка представляет собой клетку растения, клетку гриба, клетку животного или бактериальную клетку.

27. Набор, содержащий любой один или более из следующего:

ДНК последовательность уникального идентификатора;

рандомизированный пул ДНК последовательностей уникальных идентификаторов;

олигонуклеотид по п. 20 или 21;

кассета по п.22;

одну или более пар праймеров для амплификации и/или секвенирования ДНК последовательности уникального идентификатора;

буфер;

полимеразу; или

инструкции по выполнению способа по любому из пп.1-19.

28. Способ идентификации биологического материала, включающий:

получение на вычислительном устройстве ДНК последовательности уникального идентификатора (DUID), извлеченной из известного биологического материала;

поиск в вычислительном устройстве базы данных DUID, в которой хранится множество DUID вместе с соответствующей информацией о биологическом материале, для совпадения с полученной DUID;

если поиск в базе данных DUID не даст совпадения с полученной DUID, сохранение в базе данных DUID полученной DUID вместе с информацией о биологическом материале, связанной с известным биологическим материалом;

после сохранения полученной DUID и информации, связанной с известным биологическим материалом в базе данных DUID, прием на вычислительном устройстве запроса DUID, извлеченной из неизвестного биологического материала;

поиск в вычислительном устройстве базы данных DUID для совпадения с полученным запросом DUID; и

если поиск DUID обеспечивает совпадение с полученным запросом DUID, возврат в ответ на полученный запрос DUID биологической информации, сохраненной в ассоциации с DUID, совпадающей с запросом DUID.

29. Способ по п. 28, в котором поиск в базе данных DUID совпадения с полученной DUID предусматривает:

поиск в базе данных DUID точного совпадения с полученной DUID; и

если точное совпадение не найдено, выполнение выравнивания/поиска идентичности для DUID, хранящихся в базе данных DUID, которые являются близким соответствием с полученной DUID.

30. Способ по п. 28 или 29, в котором поиск в базе данных DUID совпадения с запросом DUID предусматривает:

поиск в базе данных DUID точного совпадения с запросом DUID; и

если точное совпадение не найдено, выполнение выравнивания/поиска идентичности для DUID, хранящихся в базе данных DUID, которые близко соответствуют запросу DUID.

31. Способ по п. 30, дополнительно включающий:

если поиск обеспечивает близкое соответствие с DUID запроса, сохранение DUID запроса в ассоциации с DUID, который является близким соответствием DUID запроса.

32. Вычислительная система для идентификации биологического материала, содержащая:

блок обработки, способный выполнять инструкции; и

блок памяти, хранящий инструкции, которые при выполнении их блоком обработки конфигурируют вычислительную систему для выполнения способа по любому из пп. 28-31.

33. Машиночитаемая память с хранящимися в ней инструкциями, которые при выполнении их блоком обработки вычислительной системы конфигурируют систему для выполнения способа по любому из пп. 28-31.

34. Способ идентификации биологического материала, включающий:

получение или предоставление образца, содержащего геномную ДНК, из биологического материала; амплификацию по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК из биологического материала, и секвенирование ДНК последовательности уникального идентификатора; и

декодирование или расшифровку информации об идентификации и/или по отслеживанию биологического материала, хранящегося в ДНК последовательности уникального идентификатора.

35. Способ обеспечения возможности отслеживания биологического материала, включающий:

определение последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК биологического объекта;

подтверждение идентификации биологического объекта посредством: проверки наличия ДНК последовательности уникального идентификатора в геномной ДНК; и декодирование или расшифровку информации об идентификации и/или по отслеживанию, хранящейся в ДНК последовательности уникального идентификатора, для проверки ДНК последовательности уникального идентификатора; и

предоставление указания о соответствии получения биологического материала из биологического объекта, где биологический материал содержит геномную ДНК из биологического объекта;

обеспечивая тем самым возможность отслеживания биологического материала путем считывания ДНК последовательности уникального идентификатора в биологическом материале и декодирования или

расшифровки информации, хранящейся в ДНК последовательности уникального идентификатора, предоставляя информацию об идентификации и/или по отслеживанию биологического материала.

36. Способ идентификации биологического материала, включающий:

получение на вычислительном устройстве ДНК последовательности уникального идентификатора (DUID), извлеченной из неизвестного биологического материала; и

декодирование или расшифровку информации об идентификации и/или по отслеживанию неизвестного биологического материала, хранящейся в ДНК последовательности уникального идентификатора.

37. Кассета, содержащая ДНК последовательность уникального идентификатора, где ДНК последовательность уникального идентификатора фланкирована по меньшей мере одной последовательностью 5'-праймеров отжига и по меньшей мере одной последовательностью 3'-праймеров отжига для амплификации ДНК последовательности уникального идентификатора, секвенирование ДНК последовательности уникального идентификатора или для того и другого.

38. Кассета по п. 37, в которой ДНК последовательность уникального идентификатора фланкирована двумя последовательностями 5'-праймеров отжига и двумя последовательностями 3'-праймеров отжига, чтобы обеспечить амплификацию ДНК последовательности уникального идентификатора с помощью вложенной ПЦР.

39. Кассета по п. 38, в которой две последовательности 5'-праймеров отжига частично перекрываются; или две последовательности 3'-праймера отжига частично перекрываются; или имеют место оба случая.

40. Кассета по любому из пп. 37-39, в которой кассета дополнительно содержит последовательность праймера отжига для секвенирования, расположенную в положении 5' относительно ДНК последовательности уникального идентификатора, для секвенирования ДНК последовательности уникального идентификатора.

41. Кассета по п. 40, в которой последовательность праймеров отжига для секвенирования расположена между двумя последовательностями 5'-праймеров отжига.

42. Кассета по п. 41, в которой последовательность праймеров отжига для секвенирования по меньшей мере частично перекрывается с одной или обеими последовательностями 5'-праймеров отжига.

43. Кассета по п. 41, в которой две последовательности 5'-праймера отжига частично перекрываются, и при этом по меньшей мере часть последовательности праймера отжига для секвенирования расположена в месте перекрытия.

44. Кассета по любому из пп. 37-43, в которой последовательность кассеты имеет размер до около 1500 нуклеотидов (нт) в длину; до около 1000 нт в длину; от около 200 до около 600 нт в длину; от около 200 нт до около 400 нт в длину; или от около 400 нт до около 600 нт в длину.

45. Кассета по любому из пп. 37-44, в которой последовательности праймеров отжига не встречаются в природе в геноме целевого биологического объекта.

46. Композиция, содержащая множество кассет по любому из пп. 37-45, где каждая кассета содержит вышеуказанные последовательности праймеров отжига, и каждая кассета содержит рандомизированную ДНК последовательность уникального идентификатора.

47. Композиция, содержащая множество кассет, как определено в любом из пп. 40-43, где каждая кассета содержит вышеуказанные последовательности праймеров отжига и вышеуказанную

последовательность праймера отжига для секвенирования, и каждая кассета содержит рандомизированную ДНК последовательность уникального идентификатора.

48. Способ обеспечения возможности отслеживания биологического материала, включающий:

встраивание по меньшей мере одной ДНК последовательности уникального идентификатора в геномную ДНК биологического объекта для использования при получении биологического материала.

49. Способ по п. 48, в котором ДНК последовательность уникального идентификатора встроена в виде кассеты по любому из пп. 37-45.

50. Способ по п. 48 или 49, дополнительно включающий стадию определения последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК биологического объекта.

51. Способ по любому из пп. 48-50, дополнительно включающий стадию проверки подтверждения подлинности идентификации биологического объекта путем: проверки наличия ДНК последовательности уникального идентификатора в геномной ДНК; и сравнения ДНК последовательности уникального идентификатора с базой данных для подтверждения того, что ДНК последовательность уникального идентификатора еще не используется в базе данных.

52. Способ по любому из пп. 48-51, дополнительно включающий стадию:

получение биологического материала из биологического объекта, где биологический материал содержит геномную ДНК из биологического объекта; и/или

предоставление указания о соответствии получения биологического материала из биологического объекта, где биологический материал содержит геномную ДНК из биологического объекта.

53. Способ по любому из пп. 48-52, дополнительно включающий стадию ввода последовательности по меньшей мере одной ДНК последовательности уникального идентификатора в запись базы данных, и связывания ДНК последовательности уникального идентификатора с информацией об идентификации и/или по отслеживанию биологического объекта и/или биологического материала.

54. Способ по п. 53, дополнительно включающий стадию:

обеспечение возможности отслеживания биологического объекта и/или биологического материала путем считывания ДНК последовательности уникального идентификатора в биологическом объекте и/или биологическом материале, и извлечения соответствующей записи из базы данных, предоставляющей информацию об идентификации и/или по отслеживанию биологического объекта и/или биологического материала.

55. Плазмида или вектор экспрессии, содержащий олигонуклеотид по любому из пп. 20-21 или кассету по любому из пп. 22, 37-44 или 45.

56. Способ обеспечения возможности отслеживания представляющего интерес продукта, включающий:

получение или предоставление образца представляющего интерес продукта, причем образец содержит геномную ДНК из части биологического материала, смешанного или иным образом связанного с представляющим интерес продуктом;

амплификацию по меньшей мере одной ДНК последовательности уникального идентификатора в геномной ДНК из биологического материала, и секвенирование ДНК последовательности уникального идентификатора; и

поиск ДНК последовательности уникального идентификатора в базе данных и извлечение записи из базы данных, соответствующей ДНК последовательности уникального идентификатора, где запись в базе

данных предоставляет информацию об идентификации и/или по отслеживанию представляющего интерес продукта.

57. Способ по п. 56, причем указанный способ предусматривает введение или добавление биологического материала к представляющему интерес продукту, где биологический материал содержит по меньшей мере одну ДНК последовательность уникального идентификатора как часть его геномного материала.

58. Способ по п. 56 или 57, в котором информация об идентификации и/или по отслеживанию в записи в базе данных содержит информацию о цепочке поставок представляющего интерес продукта.

59. Способ по любому из пп. 56-58, в котором представляющий интерес продукт включает пищевой продукт, сельскохозяйственный продукт, фармацевтический препарат, розничный продукт, текстильный продукт, товар широкого потребления, химический продукт или другой элемент цепочки поставок.

Фиг. 1



Случайные нуклеотиды коммерчески доступны. Их можно собирать ферментативным лигированием. Случайные фрагменты ДНК разделяют на колонке для выделения фрагментов размером примерно 300 п.н., и встраивают в кассету.

Случайный пул нуклеотидов

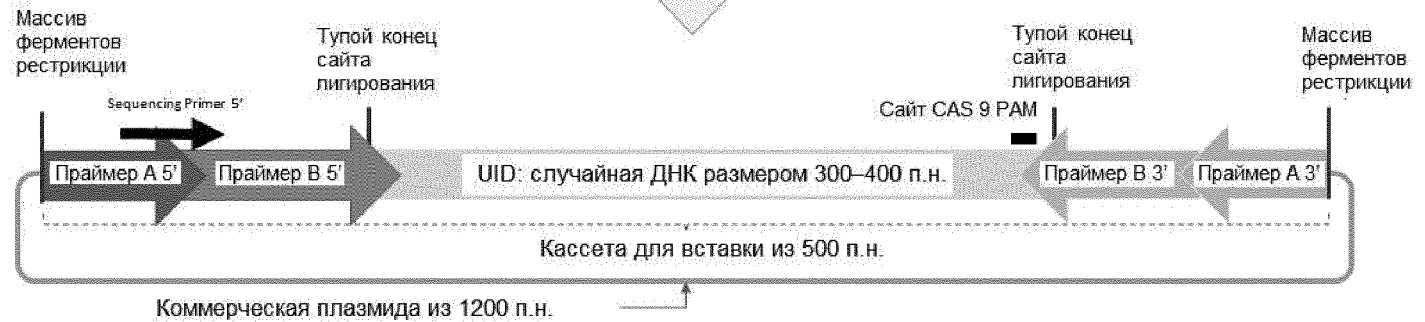


```

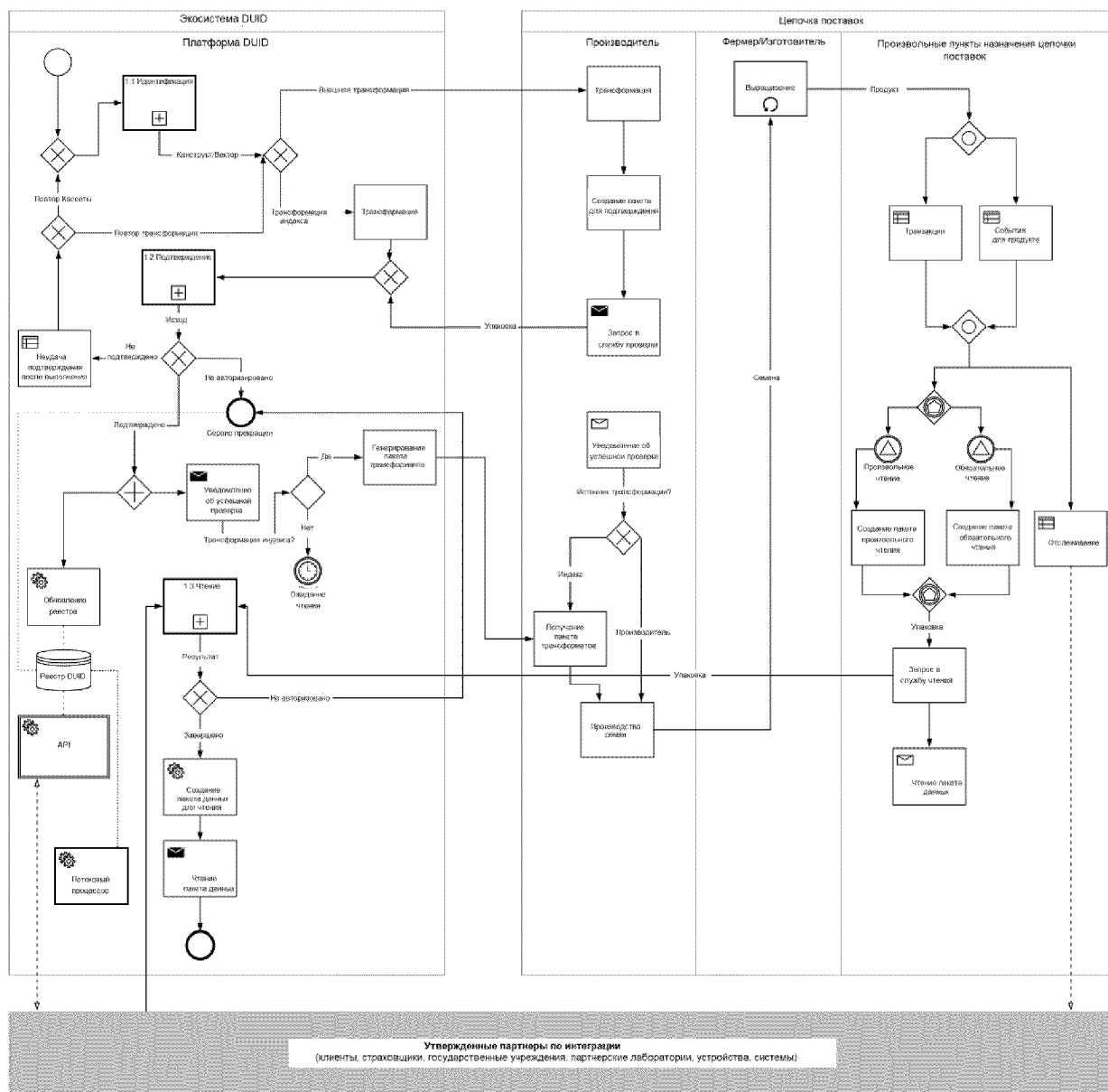
CGCCTAAGCTCCTGTGACGGGAGAGAGAACGGATTTCGGTGGTAGAACTCGCCCGTGCAGACATAAT
TGCGGATTCTCGGGGGCGTAGATGCCTAAGAATACCTCAACGCTTCTGCATGATGGTACTCACTCT
AGTTGTTTAAATTTACCACGCCAGTAGACTATGGGCAAGTCAGCGCGCAAGAGACACTCACCGTTGA
ATTAACAGCACCACCCCTTTACTCCACGCGAAAGGCGGATTATCGTATCTTTTGTGGCGAATTT
ATGTATCTCTCCTTGAGGAAGCAGAGCAGTCC
    
```

Это пример случайно созданной ДНК последовательности UID. Эта последовательность каждый раз будет отличаться, когда ее получают. В реакции образуется 10^7+ фрагментов.

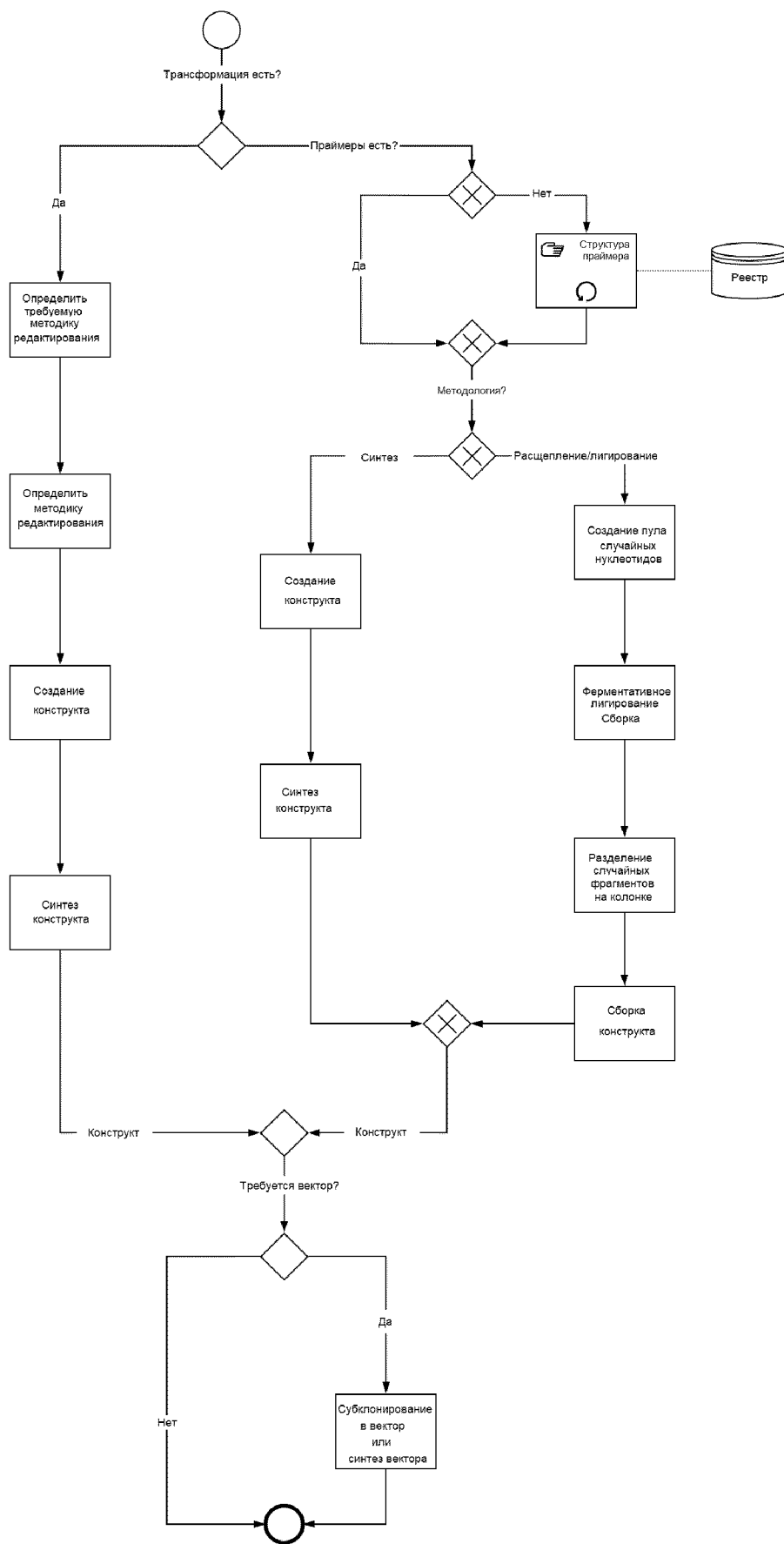
Создают праймеры и сайты секвенирования, которые проверяют на соответствие геному хозяина, чтобы гарантировать отсутствие нативной амплификации. Кассеты с разными праймерами можно использовать для разных геномов.



Фиг. 2

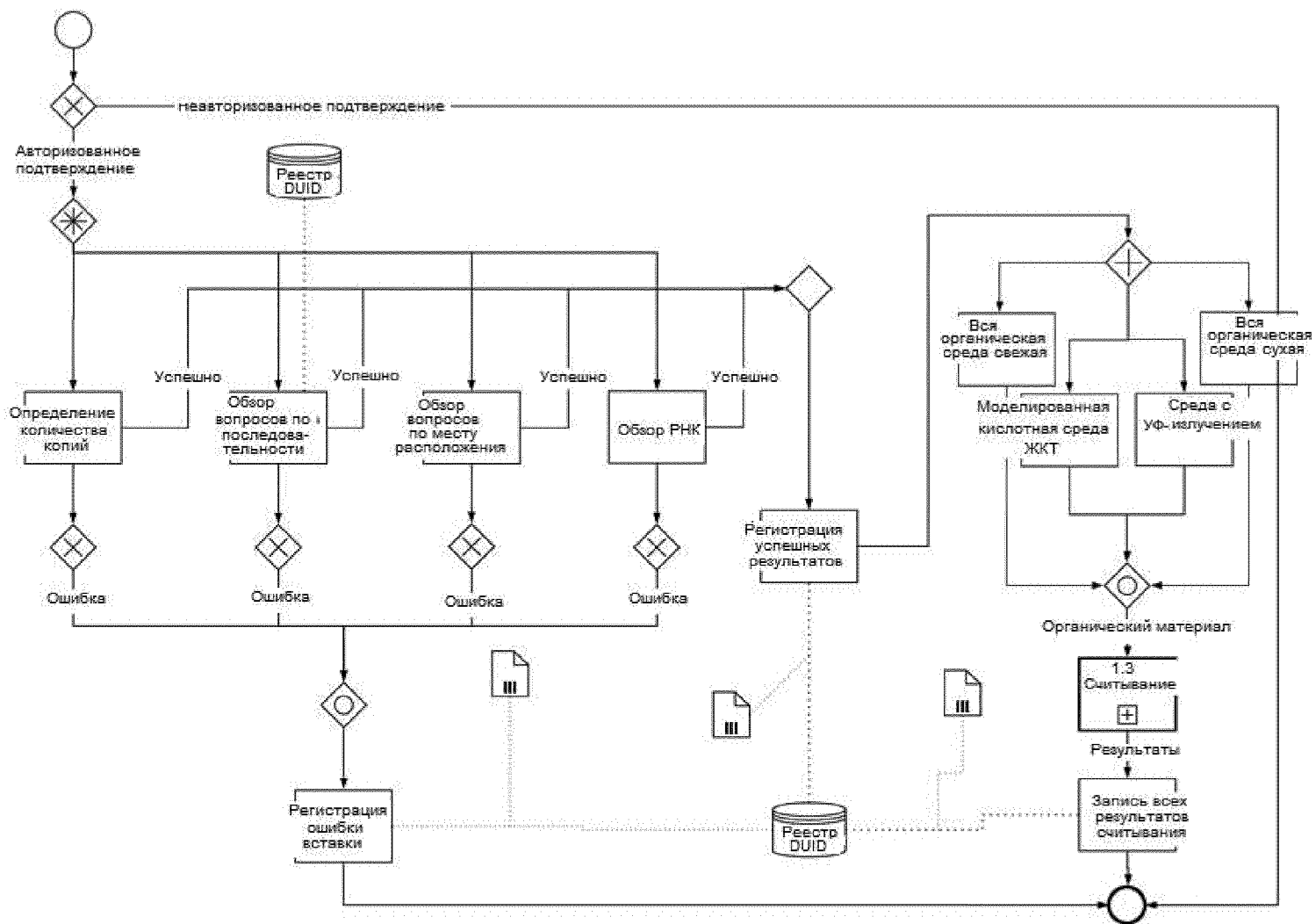


Фиг. 3

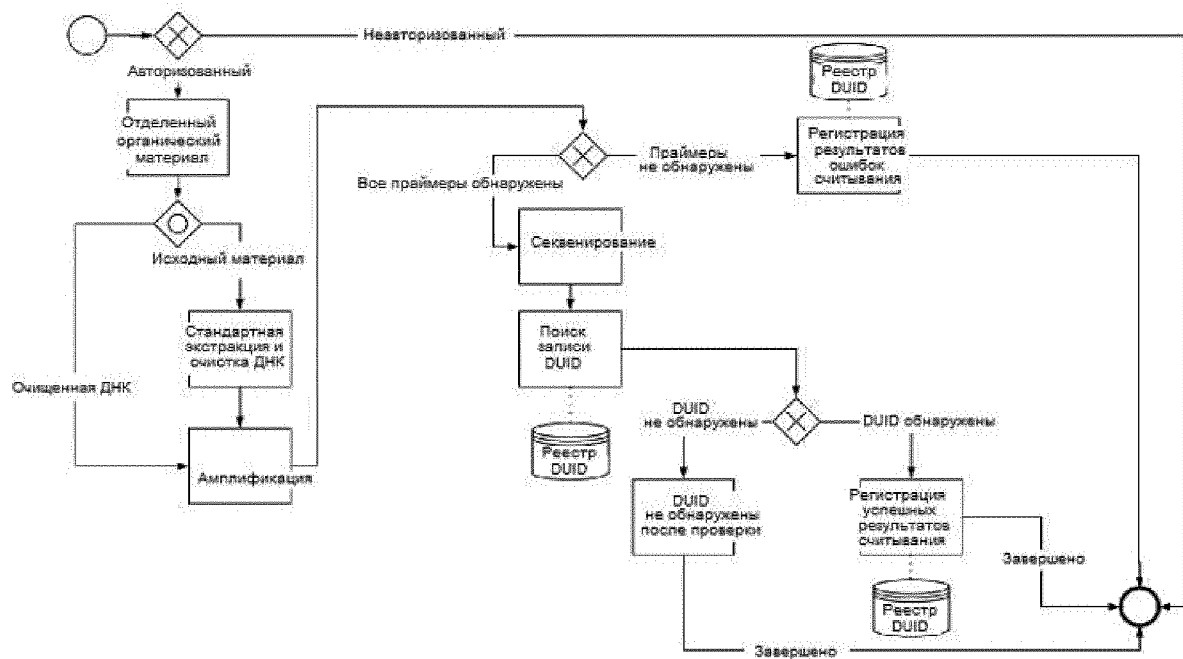


Фиг. 4

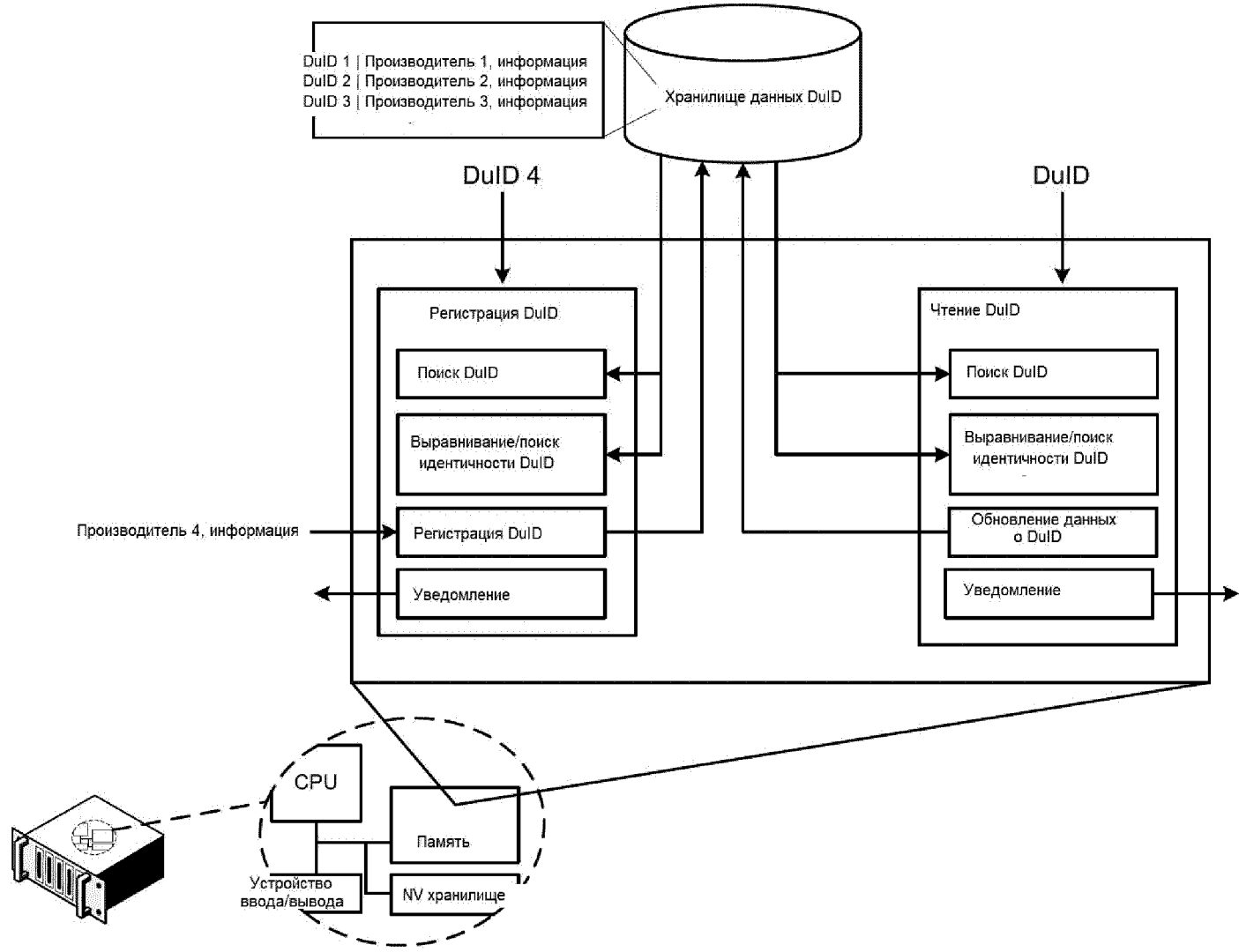
Фиг. 5

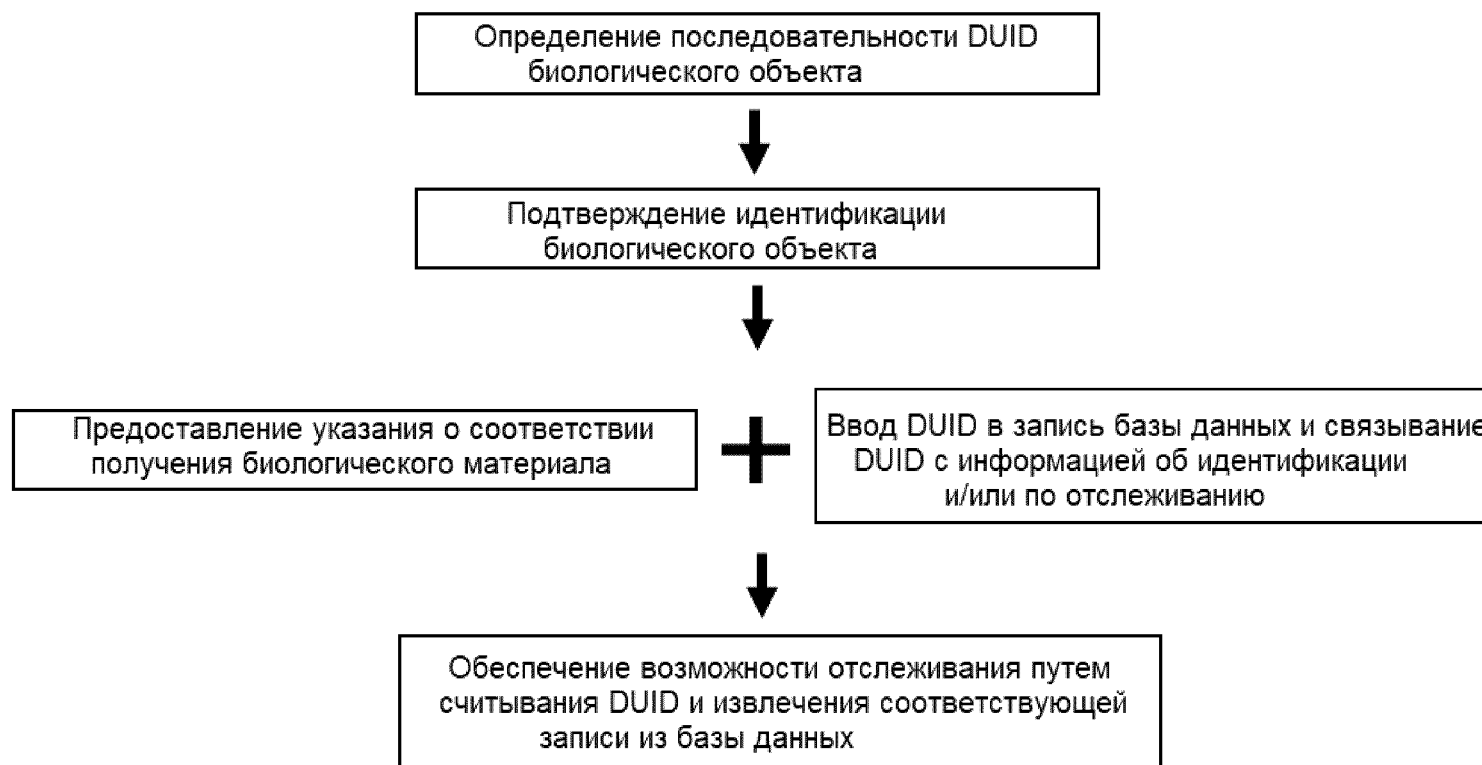


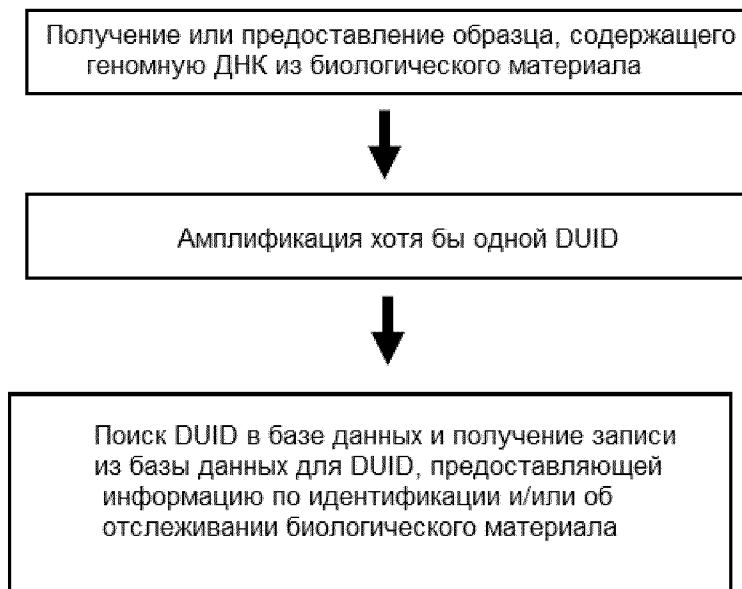
Фиг. 6



Фиг. 7



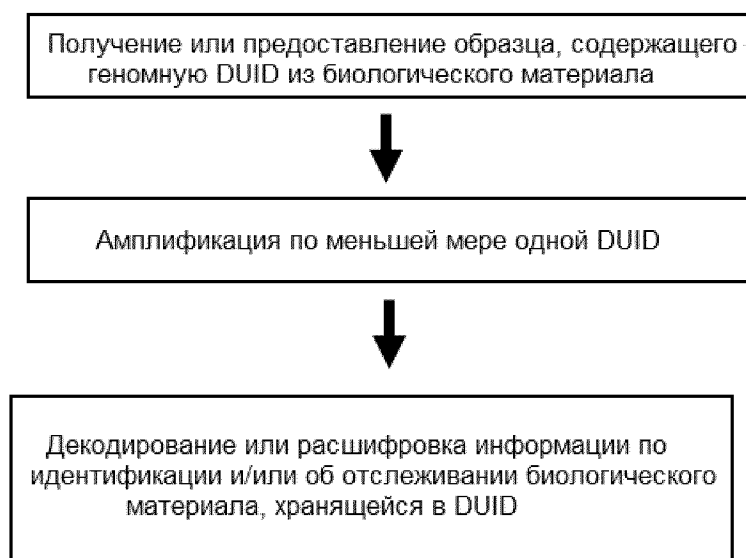




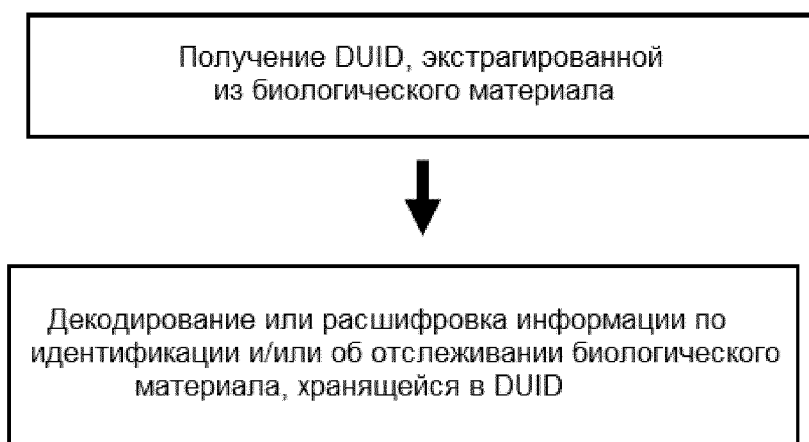
Фиг. 9



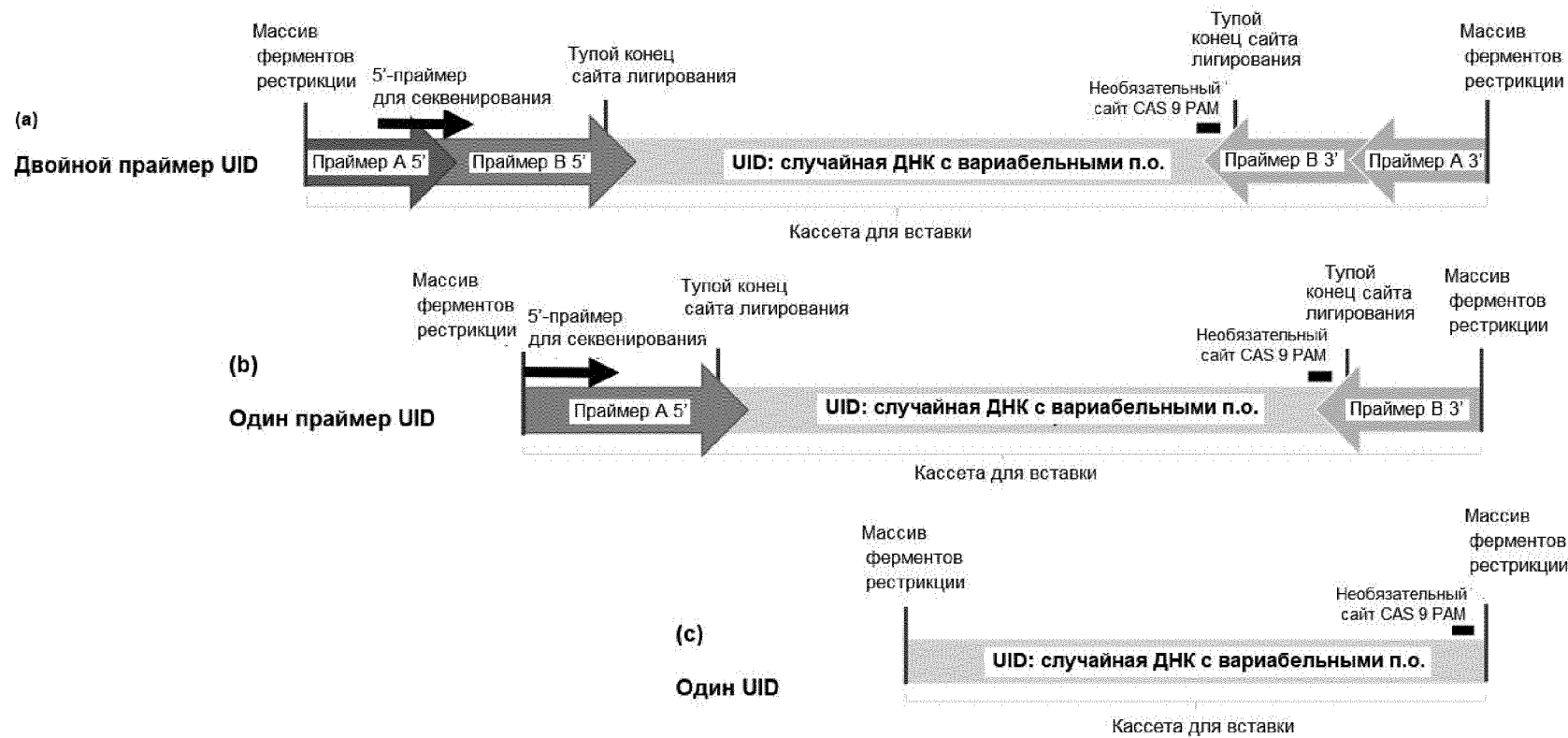
Фиг. 10



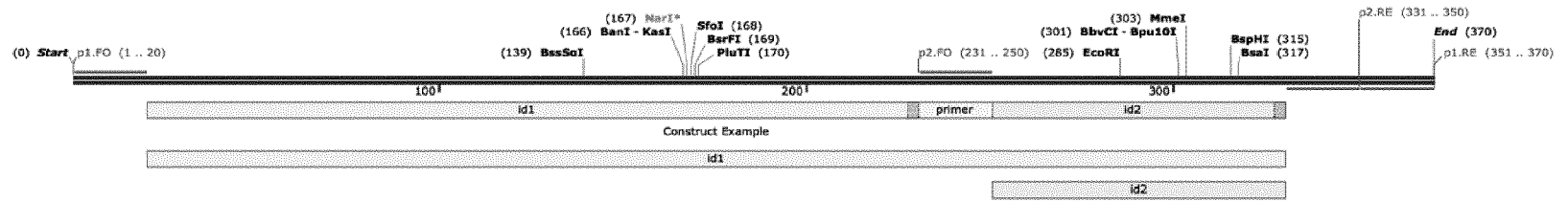
Фиг. 11



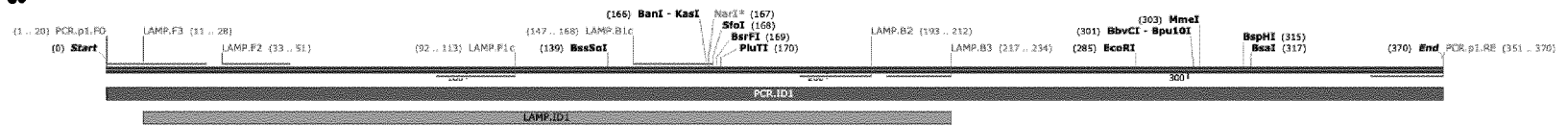
Фиг. 12

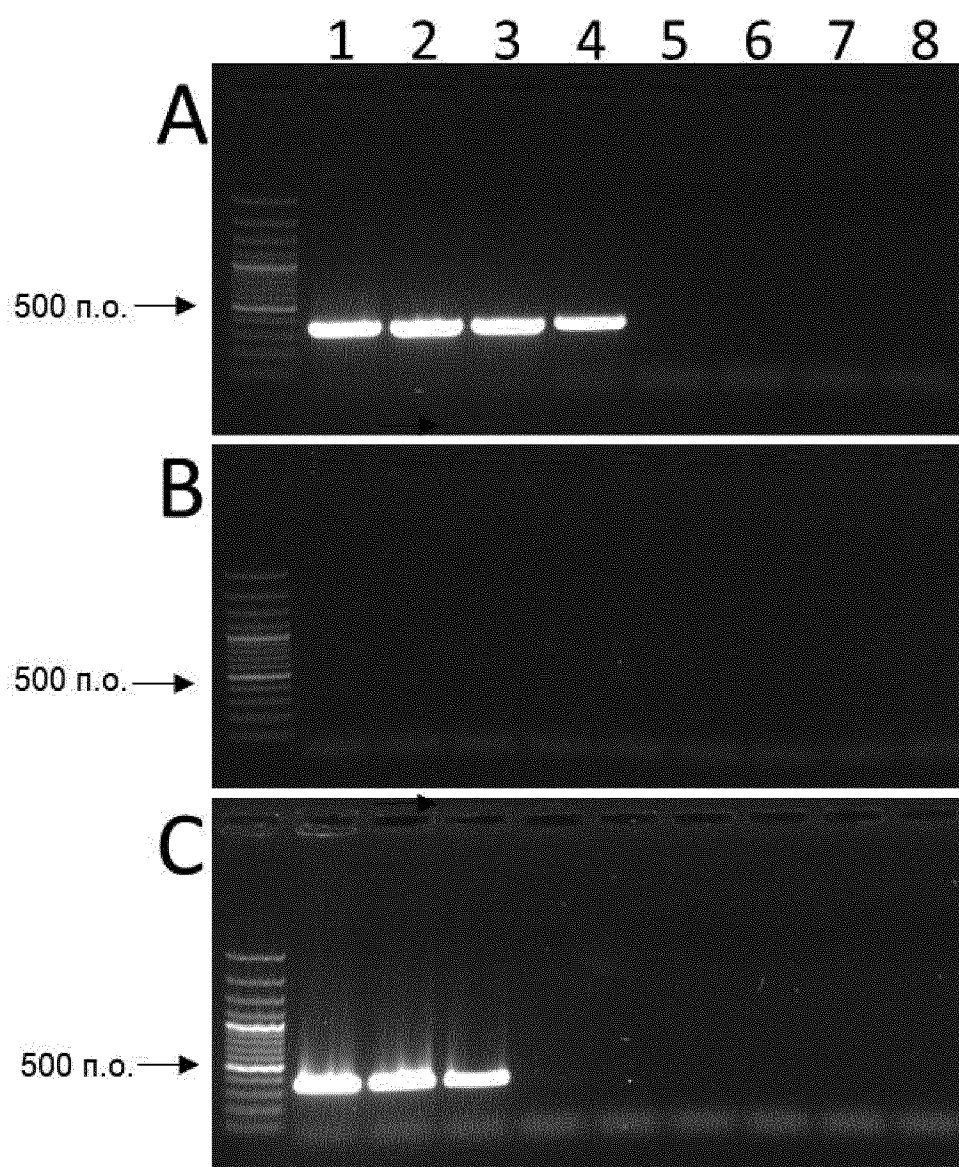


A

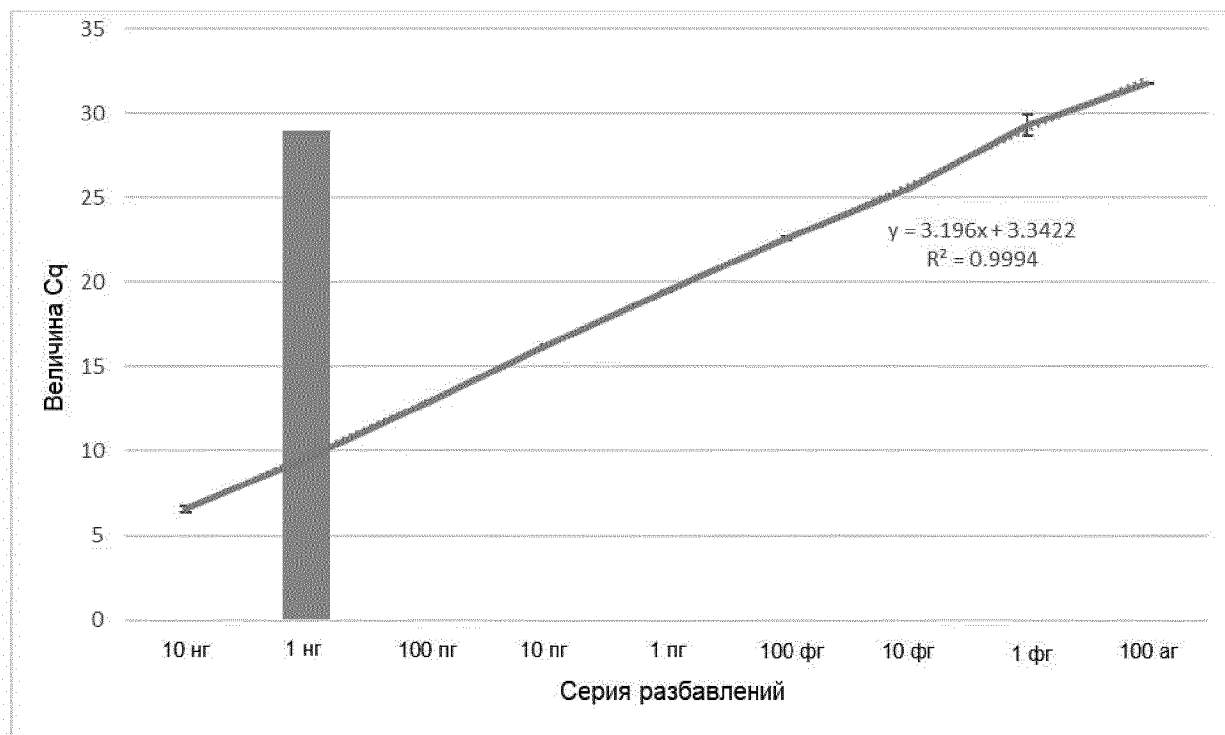


B





Фиг. 15



Фиг. 16

Фиг. 17

