

(19)



**Евразийское
патентное
ведомство**

(21) **202292161** (13) **A1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2022.10.06

(51) Int. Cl. **C12N 15/70** (2006.01)
C12P 19/00 (2006.01)
C07H 3/06 (2006.01)

(22) Дата подачи заявки
2021.01.22

(54) НОВЫЙ БЕЛОК СУПЕРСЕМЕЙСТВА ОСНОВНЫХ ПЕРЕНОСЧИКОВ (MFS) (Fred) В ПРОИЗВОДСТВЕ ОГМ

(31) **РА202000087**

(32) **2020.01.23**

(33) **DK**

(86) **PCT/EP2021/051479**

(87) **WO 2021/148620 2021.07.29**

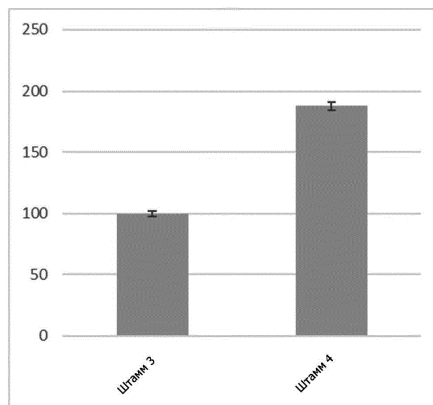
(71) Заявитель:
ГЛЮКОМ А/С (DK)

(72) Изобретатель:

**Педерсен Маргит, Д'Арриго Изотта,
Кампманн Катрине Бых, Пападакис
Манос (DK)**

(74) Представитель:
Фелицына С.Б. (RU)

(57) Изобретение относится к области рекомбинантного получения биологических молекул в генетически модифицированных клетках. Более конкретно, изобретение относится к способу рекомбинантного получения олигосахаридов грудного молока (ОГМ) с применением генетически модифицированной клетки, экспрессирующей белок суперсемейства основных посредников (MFS), экспрессируемый белок представляет собой Fred.



A1

202292161

202292161

A1

НОВЫЙ БЕЛОК СУПЕРСЕМЕЙСТВА ОСНОВНЫХ ПЕРЕНОСЧИКОВ (MFS) (FRED) В ПРОИЗВОДСТВЕ ОГМ

Область техники, к которой относится изобретение

Настоящее изобретение относится к области рекомбинантного получения биологических молекул в генетически модифицированных клетках. Более конкретно, оно относится к способу рекомбинантного получения олигосахаридов грудного молока (ОГМ) с применением генетически модифицированных клеток, экспрессирующих белок суперсемейства основных посредников (MFS), экспрессируемый белок представляет собой Fred.

Уровень техники

Грудное молоко представляет собой сложную смесь углеводов, жиров, белков, витаминов, минералов и микроэлементов. Безусловно, наиболее преобладающая фракция представлена углеводами, которые можно далее разделить на лактозу и более сложные олигосахариды (олигосахариды грудного молока, ОГМ). В то время как лактозу применяют в качестве источника энергии, сложные олигосахариды не метаболизируются младенцем. Фракция сложных олигосахаридов составляет до 1/10 от общей углеводной фракции и состоит, вероятно, из более чем 150 различных олигосахаридов. Наличие и концентрация этих сложных олигосахаридов специфичны для человека и поэтому не могут быть обнаружены в больших количествах в молоке других млекопитающих, таких как, например, домашние молочные животные.

На сегодняшний день определена структура не менее 115 ОГМ (см. Urashima *et al.*: Milk Oligosaccharides, Nova Biomedical Books, New York, 2011, ISBN: 978-1-61122-831-1), и, вероятно, значительно больше присутствует в грудном молоке (Kunz C. *et al.*, (2014) Food Oligo-saccharides: Production, Analysis and Bioactivity, 1st Edition, p 5-20, Eds. Moreno J. and Luz Sanz M., John Wiley & Sons, Ltd).

ОГМ вызвали большой интерес в последнее десятилетие в связи с открытием их важной функциональности в развитии человека. Помимо своих пребиотических свойств, ОГМ связаны с дополнительными положительными эффектами, что расширяет область их применения. (Kunz C. *et al.*, (2014) Food Oligosaccharides: Production, Analysis and Bioactivity, 1st Edition, p 5-20, Eds. Moreno J. and Luz Sanz M., John Wiley & Sons, Ltd). Польза ОГМ для здоровья позволила разрешить их применение в пищевых продуктах, таких как детские смеси и пищевые продукты, а также в потребительских товарах для здоровья.

Из-за ограниченной доступности ОГМ крайне желательно эффективное

коммерческое, т.е. крупномасштабное производство. Разработаны химические маршруты к некоторым ОГМ. Однако производство ОГМ путем химического синтеза, ферментативного синтеза или ферментации оказалось сложной задачей. Производство в больших количествах, а также качества, необходимые для пищевых и медицинских целей, путем химического синтеза еще предстоит обеспечить. Кроме того, химические синтетические пути к ОГМ включают несколько вредных химических веществ, которые создают риск загрязнения конечного продукта.

Чтобы обойти недостатки, связанные с химическим синтезом ОГМ, были разработаны несколько энзимологических способов и ферментативных подходов. Процессы на основе ферментации были разработаны для нескольких ОГМ, таких как 2'-фукозиллактоза, 3-фукозиллактоза, ласто-N-тетраоза, лакто-N-неотетраоза, 3'-сиалиллактоза и 6'-сиалиллактоза. В процессах, основанных на ферментации, обычно применяют генетически модифицированные бактериальные штаммы, такие как рекомбинантная кишечная палочка (*E. coli*).

Биотехнологическое производство ОГМ, такое как ферментативный способ, представляет собой ценный экономичный и масштабный подход к производству ОГМ. Оно основано на генетически модифицированных бактериях, сконструированных таким образом, чтобы экспрессировать гликозилтрансферазы, необходимые для синтеза желаемых олигосахаридов, и в нем применяют врожденный пул нуклеотидных сахаров бактерий в качестве предшественников ОГМ.

Недавние разработки в биотехнологическом производстве ОГМ позволили преодолеть некоторые присущие бактериальным экспрессионным системам ограничения. Например, бактериальные клетки, продуцирующие ОГМ, могут быть генетически модифицированы для увеличения ограниченного внутриклеточного пула нуклеотидных сахаров в бактериях (WO2012112777), для повышения активности ферментов, участвующих в продукции ОГМ (WO2016040531), или для облегчения секреции синтезированных ОГМ во внеклеточную среду (WO2010142305, WO2017042382). Кроме того, экспрессию представляющих интерес генов в рекомбинантных клетках можно регулировать с помощью конкретных промоторов или других регуляторов экспрессии генов, таких как, например, те, что недавно были описаны в WO 2019123324.

Подход, описанный в WO 2010142305 и WO 2017042382, имеет то преимущество, что он позволяет снизить метаболическую нагрузку, воздействующую на продуцирующую клетку, за счет высоких уровней экспрессии рекомбинантных генов, т.е. с применением способов WO2012112777, WO2016040531 или WO2019123324. Этот подход привлекает все большее внимание к конструированию рекомбинантных клеток,

продуцирующих ОГМ, например, недавно были описаны также несколько новых генов переносчиков сахара, кодирующих белки, и процедуры ферментации, которые могут способствовать выведению рекомбинантно продуцируемой 2'-фукозиллактозы (2'-FL), самого распространенного ОГМ грудного молока (WO2018077892, US201900323053, US201900323052).

Однако в настоящее время не существует алгоритма, который мог бы точно определить правильный белок-транспортер для оттока различных структур ОГМ, полученных рекомбинантно, среди многочисленных бактериальных белков с предсказанной функцией транспортера в нескольких базах данных белков, например, в UniProt, поскольку отношение структуры-функции, определяющие субстратную специфичность переносчиков сахаров, до сих пор недостаточно изучены и остаются крайне непредсказуемыми.

Раскрытие изобретения

Настоящее раскрытие показывает, что сверхэкспрессия в штаммах, продуцирующих ОГМ, гетерологичного гена *fred*, который кодирует белок Fred из суперсемейства основных переносчиков (MFS), увеличивает количество ОГМ, продуцируемого клетками. Выявление новых эффективных белков-транспортеров оттока сахаров, имеющих эффективность для различных рекомбинантно полученных ОГМ и разработка рекомбинантных клеток, экспрессирующих указанные белки, выгодны для крупномасштабного промышленного производства ОГМ.

Таким образом, в самом широком аспекте настоящее изобретение относится к генетически модифицированной клетке, способной продуцировать один или несколько олигосахаридов грудного молока (ОГМ), причем указанная генетически модифицированная клетка включает гетерологичную последовательность нуклеиновой кислоты, кодирующую полипептид главного облегчающего суперсемейства (MFS), показанный на SEQ ID NO: 1 или ее функциональный гомолог, аминокислотная последовательность которого более чем на 95,4% идентична SEQ ID NO: 1, например, по меньшей мере, на 95,5, например, по меньшей мере, на 96% идентична SEQ ID NO: 1.

Генетически модифицированная клетка по настоящему изобретению может дополнительно включать последовательность нуклеиновой кислоты, включающую регуляторный элемент для регуляции экспрессии гетерологичной последовательности нуклеиновой кислоты. Указанный регуляторный элемент в одном аспекте регулирует экспрессию полипептида MFS, показанного в SEQ ID NO: 1, или его функционального гомолога, аминокислотная последовательность которого более чем на 95,4% идентична SEQ ID NO: 1.

Аминокислотная последовательность, обозначенная в настоящем документе как SEQ ID NO: 1, представляет собой аминокислотную последовательность, которая на 100% идентична аминокислотной последовательности, имеющей номер доступа GenBank ID WP_087817556.1. Белок-транспортер MFS, имеющий аминокислотную последовательность SEQ ID NO: 1, идентифицируется в настоящем документе как «белок Fred» или «переносчик Fred» или «Fred», взаимозаменяемо; последовательность нуклеиновой кислоты, кодирующая белок Fred, обозначена в настоящем документе как SEQ ID NO: 2 «Fred, кодирующая нуклеиновую кислоту/ДНК», или «ген *fred*», или «*fred*».

Настоящее изобретение показывает, что применение рекомбинантных клеток, продуцирующих ОГМ, которые экспрессируют белок Fred, приводит к весьма отчетливым улучшениям процесса производства ОГМ как с точки зрения ферментации, так и очистки ОГМ. Описанные в настоящем документе рекомбинантные клетки и способы получения ОГМ обеспечивают как более высокие выходы всех произведенных ОГМ, сниженное образование побочных продуктов или отношение побочного продукта к продукту и облегченное извлечение ОГМ во время последующей обработки ферментационный бульон.

Неожиданно было обнаружено, что экспрессия последовательности ДНК, кодирующей Fred, в различных ОГМ-продуцирующих клетках связана с накоплением некоторых ОГМ во внеклеточной среде и других ОГМ внутри продуцирующих клеток, и с увеличением общей продукции ОГМ. Неожиданно оказалось, что увеличение выхода образующихся ОГМ характерно для ОГМ, состоящих либо из трех, либо из четырех звеньев моносахаридов, т.е. 2'-фукозиллактозы (2'-FL), 3-фукозиллактозы (3-FL), 3-сиалиллактозы (3'SL), 6'-сиалиллактозы (6'SL), лакто-N-триозы 2, (LNT-2), лакто-N-неотетраозы (LNnT) и лакто-N-тетраозы (LNT), особенно для 2'-фукозиллактозы (2'-FL), 3-фукозиллактозы (3-FL) и лакто-N-тетраозы (LNT), как видно из приведенных в настоящем документе Примеров, но не для более крупных олигосахаридных структур, таких как пентасахариды и гексасахариды, которые накапливаются внутри продуцирующих клеток.

Неожиданно также обнаружено, что общая продукция основного ОГМ, например, 2'-фукозиллактозы (2'-FL), 3-фукозиллактозы (3-FL), и лакто-N-тетраозы (LNT) в соответствующих ОГМ-продуцирующих клетках, экспрессирующих ген *fred*, также увеличивается, в то время как образование побочных продуктов, например, например дифукозиллактозы (DFL), лакто-N-фукопентаозы V (LNFP V) или паралактонеогексаозы-I (pLNH-I) в этих клетках, соответственно, часто снижены, и указанные побочные олигосахариды обычно накапливаются внутри продуцирующих клеток.

Неожиданно было также обнаружено, что общая продукция основных ОГМ, т.е. 2'-фукозиллактозы (2'-FL), 3-фукозиллактозы (3-FL) и лакто-N-тетраозы (LNT) в соответствующих ОГМ-продуцирующих клетках, экспрессирующих ген *fred*, также увеличивается, в то время как продукция побочных продуктов, т.е. дифукозиллактозы (DFL), лакто-N-фукопентаозы V (LNFP V) или паралактонеогексаозы-I (pLNH-I) в этих клетках, соответственно, часто снижается, и указанные побочные продукты олигосахаридов обычно накапливаются внутри продуцирующих клеток.

Генетически модифицированная клетка по настоящему изобретению обычно дополнительно включает последовательность нуклеиновой кислоты, включающую регуляторный элемент. Регуляторный элемент может регулировать экспрессию нуклеиновой кислоты, кодирующей полипептид суперсемейства основных переносчиков (MFS), и может быть выбран из группы, состоящей из *PglpF*, *PglpF_SD4* и *PglpF_SD7*.

Изобретение также относится к конструкции нуклеиновой кислоты, включающей последовательность нуклеиновой кислоты, кодирующую полипептид суперсемейства основных переносчиков (MFS), где последовательность нуклеиновых кислот, кодирующая полипептид суперсемейства основных переносчиков (MFS), имеет, по меньшей мере, 70%, более предпочтительно, по меньшей мере, 80%, более предпочтительно, по меньшей мере, 90%, более предпочтительно, по меньшей мере, 95% и еще более предпочтительно, по меньшей мере, 99% идентичности с последовательностью SEQ ID NO: 2, также изобретение относится к генетически модифицированной клетке, включающей конструкцию нуклеиновой кислоты, которая представляет собой *Escherichia coli*.

В одном аспекте конструкция нуклеиновой кислоты, включающая последовательность(последовательности) нуклеиновой кислоты, кодирующую(кодирующие) полипептид, относящийся к суперсемейству основных переносчиков (MFS), в которой последовательность нуклеиновой кислоты, по меньшей мере, на 70% идентична SEQ ID NO: 2.

Изобретение также обеспечивает способ биосинтетического производства одного или нескольких олигосахаридов грудного молока (ОГМ), включающий следующие стадии:

(i) обеспечение генетически модифицированной клетки по настоящему изобретению;

(ii) культивирование генетически модифицированной клетки в соответствии с (i) в подходящей среде для культивирования клеток для экспрессии указанного полипептида, способного к транспортировке выходящего сахара и продуцированию одного или нескольких олигосахаридов грудного молока (ОГМ), и;

(iii) сбор одного или нескольких ОГМ, полученных на стадии (ii).

Изобретение также относится к применению генетически модифицированной клетки или конструкции нуклеиновой кислоты, включающей гетерологичную последовательность нуклеиновой кислоты, кодирующую полипептид суперсемейства основных посредников (MFS), причем указанная последовательность нуклеиновой кислоты имеет, по меньшей мере, 70% идентичности последовательности с SEQ ID NO: 2, для биосинтетического производства одного или нескольких олигосахаридов грудного молока (ОГМ).

Как упоминалось выше, при культивировании генетически модифицированных клеток, способных продуцировать один или несколько ОГМ и включающих последовательность нуклеиновой кислоты, кодирующую белок-транспортер Fred, неожиданно было обнаружено, что соответствующие один или несколько ОГМ продуцируются с высокими выходами, тогда как образование побочных продуктов и биомассы снижается. Это облегчает извлечение этих ОГМ в последующих процессах, например, общая процедура извлечения и очистки может включать меньше стадий, и общее время очистки сокращается.

Эти эффекты повышенного выхода продукта и облегчения извлечения продукта делают настоящее изобретение превосходящим раскрытия предшествующего уровня техники.

Другие аспекты и полезные признаки настоящего изобретения подробно описаны и проиллюстрированы приведенными ниже неограничивающими рабочими примерами.

Краткое описание чертежей

На фиг. 1 показана относительная продукция 2'-FL модифицированным штаммом *E. coli* (штамм 1) с интеграцией и без интеграции гена *fred* в соответствии с SEQ ID NO: 2, кодирующего экспрессию белка-транспортера Fred MFS в соответствии с SEQ ID NO: 1, штамм 2 и штамм 1, соответственно. Данные получены анализом культуры с подпиткой способом с глубокими лунками.

На фиг. 2 показано относительное распределение 2'-FL внутри и снаружи клеток модифицированного штамма *E. coli* (штамм 1) с интеграцией и без интеграции гена *fred* в соответствии с SEQ ID NO: 2, кодирующего экспрессию белка-транспортера Fred MFS в соответствии с SEQ ID NO: 1, штамм 2 и штамм 1, соответственно. Данные получены анализом культуры с подпиткой способом с глубокими лунками.

На фиг. 3 показана относительная продукция 3-FL модифицированным штаммом *E. coli* (штамм 3) с интеграцией и без интеграции гена *fred* в соответствии с SEQ ID NO: 2, кодирующего экспрессию белка-транспортера Fred MFS в соответствии с SEQ ID NO: 1,

штамм 4 и штамм 3, соответственно. Данные получены анализом культуры с подпиткой способом с глубокими лунками.

На фиг. 4 показано относительное распределение 3-FL внутри и снаружи клеток модифицированного штамма *E. coli* (штамм 3) с интеграцией гена *fred* и без нее в соответствии с SEQ ID NO: 2, кодирующего экспрессию белка-транспортера Fred MFS в соответствии с SEQ ID NO: 1, штамм 4 и штамм 3, соответственно. Данные получены анализом культуры с подпиткой способом с глубокими лунками.

На фиг. 5 показана относительная продукция LNT модифицированным штаммом *E. coli* (штамм 5) с интеграцией и без интеграции гена *fred* в соответствии с SEQ ID NO: 2, кодирующего экспрессию белка-транспортера Fred MFS в соответствии с SEQ ID NO: 1, штаммы 6 и 7 по сравнению со штаммом 5, соответственно. Данные получены анализом культуры с подпиткой способом с глубокими лунками.

На фиг. 6 показано относительное распределение LNT внутри и снаружи клеток модифицированного штамма *E. coli* (штамм 5) с интеграцией и без интеграции гена *fred* в соответствии с SEQ ID NO: 2, кодирующего экспрессию белка-транспортера Fred MFS. согласно SEQ ID NO: 1, штаммы 6 и 7 по сравнению со штаммом 5, соответственно. Данные получены анализом культуры с подпиткой способом с глубокими лунками.

На фиг. 7 показан процент (%) относительных концентраций LNT в общих образцах для семи штаммов, которые экспрессируют гены *GlcNacT* и *Gal4T* и ген гетерологичного транспортера, *yberC0001_9420*, *bad*, *fred*, *marc*, *vag*, *nec* или *yabM*, соответственно. Штамм MP3672 экспрессирует гены гликозилтрансферазы и не экспрессирует гены транспортера;

На фиг. 8 показан процент (%) относительных концентраций LNT в супернатанте клеток, экспрессирующих *vag* или *yabM*.

Осуществление изобретения

Далее воплощения изобретения будут описаны более подробно. Каждая конкретная вариация признаков может быть применена к другим воплощениям изобретения, если специально не указано иное.

Как правило, все применяемые в настоящем документе термины следует интерпретировать в соответствии с их обычным значением в технической области и как применимые ко всем аспектам и воплощениям изобретения, если явно не определено или не указано иное.

Все ссылки на «a/an/the [клетка, последовательность, ген, транспортер, стадия и т.д.]» нужно интерпретировать открыто как относящиеся, по меньшей мере, к одному экземпляру указанной клетки, последовательности, гена, транспортера, стадии и т.д. если

прямо не указано иное.

Этапы любого раскрытого в настоящем документе способа не обязательно выполнять в точном раскрытом порядке, если это не указано явно.

Настоящее изобретение в целом относится к генетически модифицированной клетке для эффективного производства олигосахаридов и к применению указанной генетически модифицированной клетки в способе получения олигосахаридов. В особенности, настоящее изобретение относится к генетически модифицированной клетке, способной синтезировать олигосахарид, предпочтительно, гетерологичный олигосахарид, в особенности, олигосахарид грудного молока (ОГМ).

Соответственно, генетически модифицированная клетка по изобретению модифицируют для экспрессии набора рекомбинантных нуклеиновых кислот, которые необходимы для синтеза клетками одного или нескольких ОГМ (которые позволяют клетке-хозяину синтезировать один или несколько ОГМ), таких как гены, кодирующие один или несколько ферментов с гликозилтрансферазной активностью, как описано ниже. Продуцирующую олигосахарид рекомбинантную клетку по изобретению дополнительно модифицируют для включения гетерологичной последовательности рекомбинантной нуклеиновой кислоты, предпочтительно, последовательности ДНК, кодирующей предположительный транспортный белок MFS (major facilitator superfamily, суперсемейство основных посредников), происходящий из бактерии *Yersinia frederiksenii*.

Более конкретно, изобретение относится к генетически модифицированной клетке, оптимизированной для продукции одного или нескольких конкретных олигосахаридов, в особенности, одного или нескольких конкретных ОГМ, включающей рекомбинантную нуклеиновую кислоту, кодирующую белок Fred.

Последовательность нуклеиновой кислоты, кодирующая белок Fred, имеющая последовательность нуклеиновой кислоты SEQ ID NO: 2, идентифицируют в настоящем документе как «нуклеиновая кислота/ДНК, кодирующая Fred», или «ген *fred*», или «*fred*».

Белок-транспортёр MFS, обозначенный в настоящем документе как «белок Fred» или «транспортёр Fred» или «Fred», взаимозаменяемо, имеет аминокислотную последовательность SEQ ID NO: 1; Аминокислотная последовательность, обозначенная в настоящем документе как SEQ ID NO: 1, представляет собой аминокислотную последовательность, которая имеет 100% идентичность с аминокислотной последовательностью, имеющей номер доступа GenBank ID WP_087817556.1.

Соответственно, один аспект изобретения относится к генетически модифицированной клетке, способной продуцировать один или несколько олигосахаридов грудного молока (ОГМ), в котором указанная генетически модифицированная клетка

включает гетерологичную последовательность нуклеиновой кислоты, кодирующую полипептид, способный транспортировать сахар, причем указанная последовательность нуклеиновой кислоты, имеющая, по меньшей мере, 70%, более предпочтительно, по меньшей мере, 80%, более предпочтительно, по меньшей мере, 90%, более предпочтительно, по меньшей мере, 95% и еще более предпочтительно, по меньшей мере, 99% идентичности с последовательностью SEQ ID NO: 2.

Кроме того, изобретение относится к генетически модифицированной клетке, оптимизированной для продукции одного или нескольких конкретных олигосахаридов, в особенности, одной или нескольких конкретных ОГМ, включающей рекомбинантную нуклеиновую кислоту, кодирующую белок, содержащий более 95,4%, например, по меньшей мере, 95,5% идентичности последовательности, предпочтительно, по меньшей мере, на 96%, более предпочтительно, по меньшей мере, на 97%, более предпочтительно, по меньшей мере, на 98% и еще более предпочтительно, по меньшей мере, на 99% идентичности последовательности с аминокислотной последовательностью SEQ ID NO: 1.

Соответственно, один аспект изобретения относится к генетически модифицированной клетке, способной продуцировать один или несколько олигосахаридов грудного молока (ОГМ), где указанная генетически модифицированная клетка включает гетерологичную последовательность нуклеиновой кислоты, кодирующую полипептид, способный транспортировать сахар, причем указанная последовательность нуклеиновой кислоты, имеющая, по меньшей мере, 70%, более предпочтительно, по меньшей мере, 75%, более предпочтительно, по меньшей мере, 80%, более предпочтительно, по меньшей мере, 90% и еще более предпочтительно, по меньшей мере, 95%, например, на 99% идентичности с последовательностью SEQ ID NO: 2.

Соответственно, другой аспект изобретения относится к генетически модифицированной клетке, способной продуцировать одну или несколько ОГМ, при этом указанная клетка включает рекомбинантную нуклеиновую кислоту, кодирующую белок SEQ ID NO: 1, или его функциональный гомолог, аминокислотная последовательность которого идентична, по меньшей мере, на 95,4%, предпочтительно, по меньшей мере, на 96%, более предпочтительно, по меньшей мере, на 97%, более предпочтительно, по меньшей мере, на 99% идентична SEQ ID NO: 1.

Под термином «функциональный гомолог» в настоящем контексте подразумевается белок, аминокислотная последовательность которого более чем на 95,4%, например, на 95,5%-99,9% идентична SEQ ID NO: 1, и имеет полезную функцию для достижения, по меньшей мере, одного полезного эффекта изобретения, т.е. увеличение общей продукции ОГМ или выборки ОГМ генетически модифицированной

клеткой, облегчают восстановление продуцированного(продуцированных) ОГМ, эффективность продукции ОГМ и/или жизнеспособность клетки, продуцирующей ОГМ.

Под термином «суперсемейства основных посредников (MFS)» подразумевают большое и исключительно разнообразное семейство класса вторичных активных транспортеров, которые отвечают за транспорт ряда различных субстратов, включая сахара, лекарства, гидрофобные молекулы, пептиды, органические ионы и т.д. Специфичность белков-переносчиков сахара крайне непредсказуема, и идентификация нового белка-переносчика со специфичностью, например, в отношении олигосахаридов требует необременительных лабораторных экспериментов (более подробную информацию смотри в обзоре Reddy V.S. et al., (2012), FEBS J. 279(11): 2022–2035). Термин «переносчик MFS» в данном контексте означает белок, который облегчает транспорт олигосахарида, предпочтительно ОГМ, через клеточную мембрану, предпочтительно транспорт ОГМ/олигосахарида, синтезированного генетически модифицированной клеткой, из цитозоля клетки во внешнюю среду клетки, предпочтительно ОГМ/олигосахарида, включающего три или четыре звена сахара, например, 2'-FL, 3'-FL, LNT-2, LNT, LNT, 3'-SL или 6'-SL. Дополнительно или альтернативно переносчик MFS может также способствовать оттоку молекул, которые не считаются ОГМ или олигосахаридами согласно настоящему изобретению, таких как лактоза, глюкоза, клеточные метаболиты или токсины.

Термин «идентичность последовательности [определенного]» в контексте двух или более последовательностей нуклеиновых кислот или аминокислот означает, что две или более последовательностей имеют общие нуклеотиды или аминокислотные остатки в данном проценте при сравнении и выравнивании для максимального соответствия в окне сравнения или обозначенных последовательностях нуклеиновых кислот или аминокислот (т.е. последовательности имеют, по меньшей мере, 90-процентную (%) идентичность). Процентную идентичность последовательностей нуклеиновых кислот или аминокислот можно определить с помощью алгоритма сравнения последовательностей BLAST 2.0 с параметрами по умолчанию или путем ручного выравнивания и визуального контроля (см., например, <http://www.ncbi.nlm.nih.gov/BLAST/>). Это определение также применимо к комплементу тестируемой последовательности и к последовательностям, которые имеют делеции и/или добавления, а также к тем, которые имеют замены. Примером алгоритма, подходящего для определения процента идентичности, сходства последовательностей и выравнивания, служит алгоритм BLAST 2.2.20+, описанный в Altschul et al. *Nucl. Acids Res.* 25, 3389 (1997). BLAST 2.2.20+ применяют для определения процента идентичности последовательностей нуклеиновых кислот и белков по изобретению. Программное

обеспечение для проведения анализов BLAST общедоступно через Национальный центр биотехнологической информации. (<http://www.ncbi.nlm.nih.gov/>). CLUSTAL Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>), EMBOSS Needle (http://www.ebi.ac.uk/Tools/psa/emboss_needle/), MAFFT (<http://mafft.cbrc.jp/alignment/server/>) or MUSCLE (<http://www.ebi.ac.uk/Tools/msa/muscle/>) представляют собой примеры алгоритмов выравнивания последовательностей.

В контексте изобретения термин «олигосахарид» означает сахаридный полимер, содержащий ряд моносахаридных звеньев. В некоторых воплощениях предпочтительными олигосахаридами представляют собой полимеры сахаридов, состоящие из трех или четырех моносахаридных звеньев, т.е. трисахариды или тетрасахариды. Предпочтительные олигосахаридами изобретения представляют собой олигосахаридами грудного молока (ОГМ).

ОГМ

Термин «олигосахарид грудного молока» или «ОГМ» в данном контексте означает сложный углевод, обнаруженный в грудном молоке человека (для отсылки смотри Urashima et al.: *Milk Oligosaccharides*. Nova Science Publisher (2011); или Chen, *Adv. Carbohydr. Chem. Biochem.* 72, 113 (2015)). ОГМ имеют структуру ядра, включающую единицу лактозы на восстанавливаемом конце, которая может быть удлинена одной или несколькими единицами β -N-ацетиллактозаминила и/или одной или несколькими единицами β -лакто-N-биозила, и эта структура ядра может быть замещенная α -L-фукопиранозильной и/или α -N-ацетилнейраминильной (сиалильной) частью. В связи с этим неокислотные (или нейтральные) ОГМ лишены остатка сиалила, а кислые ОГМ имеют в своей структуре хотя бы один остаток сиалила. Неокислотный (или нейтральный) ОГМ может быть фукозилированным или нефукозилированным. Примеры таких нейтральных нефукозилированных ОГМ включают лакто-N-триозу 2 (LNT-2), лакто-N-тетраозу (LNT), лакто-N-неотетраозу (LNnT), лакто-N-неогексаозу (LNnH), паралакто-N-неогексаозу (pLNnH), паралакто-N-гексаозу (pLNH) и лакто-N-гексаозу (LNH). Примеры нейтральных фукозилированных ОГМ включают 2'-фукозиллактозу (2'-FL), лакто-N-фукопентаозу I (LNFP-I), лакто-N-дифукогексаозу I (LNDFH-I), 3-фукозиллактозу (3-FL), дифукозиллактозу (DFL), лакто-N-фукопентаозу II (LNFP-II), лакто-N-фукопентаозу III (LNFP-III), лакто-N-дифукогексаозу III (LNDFH-III), фукозиллакто-N-гексаозу II (FLNH-II), лакто-N-фукопентаозу V (LNFP-V), лакто-N-дифукогексаозу II (LNDFH-II), фукозиллакто-N-гексаозу I (FLNH-I), фукозилпаралакто-N-гексаозу I (FpLNH-I), фукозилпаралакто-N-неогексаозу II (F-pLNnH II) и фукозиллакто-N-неогексаозу (FLNnH). Примеры кислотных ОГМ включают 3'-сиалиллактозу (3'-SL), 6'-сиалиллактозу (6'-SL),

3-фукозил-3'-сиалиллактозу (FSL), 3'-О-сиалиллакто-N-тетраозу а (LST а), фукозил-LST а (FLST а), 6'-О-сиалиллакто-N-тетраозу b (LST b), фукозил-LST b (FLST b), 6'-О-сиалиллакто-N-неотетраозу (LST с), фукозил-LST с (FLST с), 3'-О-сиалиллакто-N-неотетраозу (LST d), фукозил-LST d (FLST d), сиалиллакто-N-гексаозу (SLNH), сиалиллакто-N-неогексаозу I (SLNH-I), сиалиллакто-N-неогексаозу II (SLNH-II) и дисиалиллакто-N-тетраозу (DSLNT). В контексте настоящего изобретения лактозу не рассматривают как разновидность ОГМ.

В некоторых воплощениях изобретения предпочтительными могут быть три-ОГМ и тетра-ОГМ, например, трисахариды 2'-FL, 3-FL, LNT-2, 3'-SL, 6'-SL, и тетрасахариды DFL, LNT, LNnT, FSL.

В предпочтительном в настоящее время аспекте изобретения предпочтительные ОГМ представляют собой три-ОГМ, в особенности, трисахариды, выбранные из 2'-FL и 3-FL, и тетрасахариды, выбранные из DFL и LNT, такие как, в особенности, 3-FL.

2'-FL

2'-Фукозиллактоза (2'-FL или 2'О-фукозиллактоза) представляет собой трисахарид, точнее, фукозилированный, нейтральный трисахарид, состоящий из единиц L-фукозы, D-галактозы и D-глюкозы ($\text{Fuc}\alpha 1\text{-2Gal}\beta 1\text{-4Glc}$). Это наиболее распространенный олигосахарид грудного молока (ОГМ), естественным образом присутствующий в человеческом грудном молоке, составляя около 30% всех ОГМ. В генетически модифицированной клетке или в ферментативной реакции 2'-FL продуцируется главным образом в результате ферментативной реакции α -1,2-фукозилтрансферазы с лактозой и донором фукозила.

3-FL

3-Фукозиллактоза (3-FL) представляет собой трисахарид, точнее, фукозилированный, нейтральный трисахарид, состоящий из D-галактозы, L-фукозы и D-глюкозы ($\text{Gal}\beta 1\text{-4(Fuc}\alpha 1\text{-3)Glc}$). Он естественным образом присутствует в грудном молоке. В генетически модифицированной клетке или в ферментативной реакции 3-FL продуцируется, главным образом, ферментативной реакцией α -1,3-фукозилтрансферазы или α -1,3/4-фукозилтрансферазы с лактозой и донором фукозила.

LNT

Лакто-N-тетраоза (LNT) представляет собой тетрасахарид, точнее, нейтральный тетрасахарид, состоящий из галактозы, N-ацетилглюкозамина, галактозы и глюкозы ($\text{GlcNAc}\beta 1\text{-3Gal}\beta 1\text{-4Glc}$). Он естественным образом присутствует в грудном молоке.

DFL

Дифукозиллактоза (DFL или 2',3-ди-О-фукозиллактоза) представляет собой

олигосахарид, точнее, фукузилированный нейтральный тетрасахарид, состоящий из L-фукозы, D-галактозы, L-фукозы и D-глюкозы ($\text{Fuc}\alpha 1\text{-2Gal}\beta 1\text{-4(Fuc}\alpha 1\text{-3)Glc}$). Он естественным образом присутствует в грудном молоке. В генетически модифицированной клетке или в ферментативной реакции DFL продуцируется, главным образом, ферментативной реакцией α -1,2-фукозилтрансферазы, α -1,3-фукозилтрансферазы и/или α -1,3/4-фукозилтрансферазы с лактозой и двумя донорами фукозилов.

Функциональные ферменты

Чтобы иметь возможность синтезировать один или несколько ОГМ, рекомбинантная клетка по изобретению включает, по меньшей мере, одну рекомбинантную нуклеиновую кислоту, которая кодирует функциональный фермент с гликозилтрансферазной активностью. Ген галактозилтрансферазы может быть интегрирован в геном (посредством хромосомной интеграции) генетически модифицированной клетки, или, альтернативно, он может быть включен в плазмидную ДНК и экспрессироваться как переносимый плазмидой. Если для производства ОГМ необходимы две или более гликозилтрансфераз, например, LNT или LNTnT, то две или более рекомбинантных нуклеиновых кислоты, кодирующих разные ферменты с гликозилтрансферазной активностью, могут быть интегрированы в геном и/или экспрессированы из плазмиды, например β -1,3-N-ацетилглюкозаминилтрансферазу (первая рекомбинантная нуклеиновая кислота, кодирующая первую гликозилтрансферазу) в комбинации с β -1,3-галактозилтрансферазой (вторая рекомбинантная нуклеиновая кислота, кодирующая вторую гликозилтрансферазу) для продукции LNT, где первая и вторая рекомбинантные нуклеиновые кислоты могут быть независимо друг от друга интегрированы в хромосомно или на плазмиде.

В одном предпочтительном воплощении, как первая, так и вторая рекомбинантные нуклеиновые кислоты интегрированы в хромосому клетки-продуцента; в другом воплощении, по меньшей мере, одна из первой и второй гликозилтрансфераз представляет собой переносимую плазмидой. Белок/фермент с гликозилтрансферазной активностью (гликозилтрансфераза) может быть выбран в различных воплощениях из ферментов, обладающих активностью α -1,2-фукозилтрансферазы, α -1,3-фукозилтрансферазы, α -1,3/4-фукозилтрансферазы, α -1,4-фукозилтрансферазы, α -2,3-сиалилтрансферазы, α -2,6-сиалилтрансферазы, β -1,3-N-ацетилглюкозаминилтрансферазы, β -1,6-N-ацетилглюкозаминилтрансферазы, β -1,3-галактозилтрансферазы и β -1,4-галактозилтрансферазы. Например, для производства 2'-FL требуется, чтобы модифицированная клетка экспрессировала активный фермент α -1,2-фукозилтрансферазу; для продукции 3-FL модифицированная клетка нуждается в экспрессии активного

фермента α -1,3-фукозилтрансферазы; для продукции LNT модифицированная клетка должна экспрессировать по меньшей мере две гликозилтрансферазы, β -1,3-N-ацетилглюкозаминилтрансферазу и β -1,3-галактозилтрансферазу; для продукции 6'-SL модифицированная клетка должна экспрессировать активный фермент α -2,6-сиалилтрансферазу и путь для синтеза CMP-сиаловой кислоты; для производства 3'SL модифицированная клетка должна экспрессировать активный фермент α -2,3-сиалилтрансферазу и путь для синтеза CMP-сиаловой кислоты. Некоторые неограничивающие воплощения белков, обладающих гликозилтрансферазной активностью, которые могут кодироваться рекомбинантными генами, включенными в клетку-продуцент, могут быть выбраны из неограничивающих примеров таблицы 1.

Таблица 1

Ген	ID белковой последовательности (GenBank)	Описание	Пример ОГМ
<i>lgtA_Nm</i>	<i>WP_002248149.1</i>	β -1,3-N-ацетилглюкозамин-трансфераза	LNT, LNnT, LNFP-I, LNFP-II, LNFP-III, LNFP-V, LNFP-VI, LNDFH-I, LNDFH-II, pLNH, F-pLNH I, pLNnH
<i>lgtA_Nm_MC58</i>	<i>AAF42258.1</i>	β -1,3-N-ацетилглюкозамин-трансфераза	LNT, LNnT, LNFP-I, LNFP-II, LNFP-III, LNFP-V, LNFP-VI, LNDFH-I, LNDFH-II, pLNH, F-pLNH I, pLNnH
<i>lgtA_Hd</i>	<i>AAN05638.1</i>	β -1,3-N-ацетилглюкозамин-трансфераза	LNT, LNnT, LNFP-I, LNFP-II, LNFP-III, LNFP-V, LNFP-VI, LNDFH-I, LNDFH-II, pLNH, F-pLNH I, pLNnH
<i>lgtA_Ng_PID2</i>	<i>AAK70338.1</i>	β -1,3-N-ацетилглюкозамин-трансфераза	LNT, LNnT, LNFP-I, LNFP-II, LNFP-III, LNFP-V, LNFP-VI, LNDFH-I, LNDFH-II, pLNH, F-pLNH I, pLNnH
<i>lgtA_Ng_NCCP1 1945</i>	<i>ACF31229.1</i>	β -1,3-N-ацетилглюкозамин-трансфераза	LNT, LNnT, LNFP-I, LNFP-II, LNFP-III, LNFP-V, LNFP-VI, LNDFH-I, LNDFH-II, pLNH, F-pLNH I, pLNnH
<i>lgtA_Past</i>	<i>AAK02595.1</i>	β -1,3-N-ацетилглюкозамин-трансфераза	LNT, LNnT, LNFP-I, LNFP-II, LNFP-III, LNFP-V, LNFP-VI, LNDFH-I, LNDFH-II, pLNH, F-pLNH I, pLNnH
<i>lgtA_Nc</i>	<i>EEZ72046.1</i>	β -1,3-N-ацетилглюкозамин-трансфераза	LNT, LNnT, LNFP-I, LNFP-II, LNFP-III, LNFP-V, LNFP-VI, LNDFH-I, LNDFH-II, pLNH, F-pLNH I, pLNnH
<i>lgtA_Nm_87255</i>	<i>ELK60643.1</i>	β -1,3-N-ацетилглюкозамин-трансфераза	LNT, LNnT, LNFP-I, LNFP-II, LNFP-III, LNFP-V, LNFP-VI, LNDFH-I, LNDFH-II, pLNH, F-pLNH I, pLNnH
<i>galT_Hp/HP0826</i>	<i>NP_207619.1</i>	β -1,4-галактозилтрансфераза	LNnT, LNFP-III, LNFP-VI, pLNH I, F-pLNH I, pLNnH

Ген	ID белковой последовательности (GenBank)	Описание	Пример ОГМ
<i>galT_Nm/lgtB</i>	<i>AAF42257.1</i>	β -1,4-галактозилтрансфераза	LNnT, LNFP-III, LNFP-VI, pLNH I, F-pLNH I, pLNnH
<i>wbgO</i>	<i>WP_000582563.1</i>	β -1,3-галактозилтрансфераза	LNT, LNFP-I, LNFP-II, LNFP-V, LNDFH-I, LNDFH-II, pLNH, F-pLNH I
<i>cpsIBJ</i>	<i>AB050723.1</i>	β -1,3-галактозилтрансфераза	LNT, LNFP-I, LNFP-II, LNFP-V, LNDFH-I, LNDFH-II, pLNH, F-pLNH I
<i>jhp0563</i>	<i>AEZ55696.1</i>	β -1,3-галактозилтрансфераза	LNT, LNFP-I, LNFP-II, LNFP-V, LNDFH-I, LNDFH-II, pLNH, F-pLNH I
<i>galTK</i>	<i>homologous to BD182026</i>	β -1,3-галактозилтрансфераза	LNT, LNFP-I, LNFP-II, LNFP-V, LNDFH-I, LNDFH-II, pLNH, F-pLNH I
<i>futC</i>	<i>WP_080473865.1</i>	α -1,2-фукозилтрансфераза	2'-FL, DFL, LNFP-I, LNDFH-I
<i>FucT2_HpUA802</i>	<i>AAC99764.1</i>	α -1,2-фукозилтрансфераза	2'-FL, DFL, LNFP-I, LNDFH-I
<i>FucT2_EcO126t</i>	<i>ABE98421.1</i>	α -1,2-фукозилтрансфераза	2'-FL, DFL, LNFP-I, LNDFH-I
<i>FucT2_Hm12198</i>	<i>CBG40460.1</i>	α -1,2-фукозилтрансфераза	2'-FL, DFL, LNFP-I, LNDFH-I
<i>FucT2_Pm9515</i>	<i>ABM71599.1</i>	α -1,2-фукозилтрансфераза	2'-FL, DFL, LNFP-I, LNDFH-I
<i>FucT2_HpF57</i>	<i>BAJ59215.1</i>	α -1,2-фукозилтрансфераза	2'-FL, DFL, LNFP-I, LNDFH-I
<i>FucT6_3_Bf</i>	<i>CAH09151.1</i>	α -1,3-фукозилтрансфераза	2'-FL, 3-FL, DFL, LNFP-I, LNFP-III, LNFP-V, LNFP-VI, LNDFH-II, F-pLNH I
<i>FucT7_3_Bf</i>	<i>CAH09495.1</i>	α -1,3-фукозилтрансфераза	2'-FL, 3-FL, DFL, LNFP-I, LNFP-III, LNFP-V, LNFP-VI, LNDFH-II, F-pLNH I
<i>FucT_3_Am</i>	<i>ACD04596.1</i>	α -1,3-фукозилтрансфераза	2'-FL, 3-FL, DFL, LNFP-I, LNFP-III, LNFP-V, LNFP-VI, LNDFH-II, F-pLNH I
<i>MAMA_R764</i>	<i>AGC02224.1</i>	α -1,3-фукозилтрансфераза	2'-FL, 3-FL, DFL, LNFP-I, LNFP-III, LNFP-V, LNFP-VI, LNDFH-II, F-pLNH I
<i>Mg791</i>	<i>AEQ33441.1</i>	α -1,3-фукозилтрансфераза	2'-FL, 3-FL, DFL, LNFP-I, LNFP-III, LNFP-V, LNFP-VI, LNDFH-II, F-pLNH I
<i>Moumou_00703</i>	<i>YP_007354660</i>	α -1,3-фукозилтрансфераза	2'-FL, 3-FL, DFL, LNFP-I, LNFP-III, LNFP-V, LNFP-VI, LNDFH-II, F-pLNH I
<i>futA</i>	<i>NP_207177.1</i>	α -1,3-фукозилтрансфераза	2'-FL, 3-FL, DFL, LNFP-I, LNFP-III, LNFP-V, LNFP-VI, LNDFH-II, F-pLNH I
<i>fucT</i>	<i>AAB81031.1</i>	α -1,3-фукозилтрансфераза	2'-FL, 3-FL, DFL, LNFP-I, LNFP-III, LNFP-V, LNFP-VI, LNDFH-II, F-pLNH I
<i>fucTIII</i>	<i>AY450598.1</i>	α -1,4-фукозилтрансфераза	LNDFH-I, LNDFH-II

Ген	ID белковой последовательности (GenBank)	Описание	Пример ОГМ
<i>fucTa</i>	<i>AF194963.1</i>	α -1,3/4-фукозилтрансфераза	LNFP-II, LNDFH-I, LNDFH-II
<i>Pd2,6ST</i>	<i>BAA25316.1</i>	α -2,6-сиалилтрансфераза	6'-SL
<i>PspST6</i>	<i>BAF92026.1</i>	α -2,6-сиалилтрансфераза	6'-SL
<i>PiST6_145</i>	<i>BAF91416.1</i>	α -2,6-сиалилтрансфераза	6'-SL
<i>PiST6_119</i>	<i>BAI49484.1</i>	α -2,6-сиалилтрансфераза	6'-SL
<i>NST</i>	<i>AAC44541.1</i>	α -2,3-сиалилтрансфераза	3'-SL

Гетерологичная последовательность нуклеиновой кислоты

Один из аспектов настоящего изобретения представляет собой обеспечение конструкции нуклеиновой кислоты, включающей гетерологичную(гетерологичные) последовательность(последовательности) нуклеиновой кислоты, кодирующую(кодирующие) полипептид, способный к транспортировке сахара, который представляет собой полипептид суперсемейства основных посредников (MFS), как показано в SEQ ID NO: 1, или его функциональный гомолог, аминокислотная последовательность которого более чем на 95,4% идентична последовательности SEQ ID NO: 1, при этом последовательность нуклеиновой кислоты, кодирующая полипептид транспортер сахаров, имеет, по меньшей мере, 70% идентичности с последовательностью SEQ ID NO: 2.

Под термином «гетерологичная последовательность нуклеиновой кислоты», «кодирующая рекомбинантный ген/нуклеиновая кислота/ДНК» или «кодирующая последовательность нуклеиновой кислоты» подразумевают искусственную последовательность нуклеиновой кислоты (т.е. полученную *in vitro* с применением стандартных лабораторных способов получения последовательностей нуклеиновых кислот), которая включает набор последовательных, неперекрывающихся триплетов (кодонов), которые транскрибируются в мРНК и транслируются в полипептид, когда они находятся под контролем соответствующих контрольных последовательностей, т.е. промотора. Границы кодирующей последовательности обычно определяются сайтом связывания рибосомы, расположенным прямо перед открытой рамкой считывания на 5'-конце мРНК, стартовым кодоном транскрипции (AUG, GUG или UUG) и стоп-кодоном трансляции (UAA, UGA или UAG). Кодированная последовательность может включать, но не ограничиваться ими, геномную ДНК, кДНК, синтетические и рекомбинантные последовательности нуклеиновых кислот.

Термин «нуклеиновая кислота» включает молекулы РНК, ДНК и кДНК. Понятно,

что в результате вырождения генетического кода может быть получено множество нуклеотидных последовательностей, кодирующих данный белок. Термин «нуклеиновая кислота» применяют взаимозаменяемо с термином «полинуклеотид».

«Олигонуклеотид» представляет собой молекулу нуклеиновой кислоты с короткой цепью.

«Праймер» представляет собой олигонуклеотид, встречающийся в природе в виде очищенного фрагмента рестрикции или полученный синтетическим путем, который способен действовать как точка инициации синтеза при помещении в условия, в которых индуцируется синтез продукта удлинения праймера, комплементарного цепи нуклеиновой кислоты (т.е. в присутствии нуклеотидов и индуцирующего агента, такого как ДНК-полимераза, и при подходящей температуре и pH). Праймер предпочтительно применяют одноцепочечный для максимальной эффективности амплификации, но альтернативно он может быть двухцепочечным. Если праймер двухцепочечный, то его сначала обрабатывают, чтобы разделить его нити, прежде чем применять для приготовления продуктов для удлинения. Предпочтительно праймер представляет собой дезоксирибонуклеотидную последовательность. Праймер должен быть достаточно длинным, чтобы инициировать синтез продуктов удлинения в присутствии индуцирующего агента. Точная длина праймеров будет зависеть от многих факторов, включая температуру, источник праймера и применение способа.

Рекомбинантная нуклеиновая последовательность по изобретению может представлять собой кодирующую последовательность ДНК, например, ген или не кодирующая последовательность ДНК, например, регуляторную ДНК, такую как промоторная последовательность. Один аспект изобретения относится к получению рекомбинантной клетки, включающей последовательности рекомбинантной ДНК, кодирующие ферменты, необходимые для продукции одного или нескольких ОГМ, и последовательность ДНК, кодирующую переносчик Fred. Соответственно, в одном воплощении изобретение относится к конструкции нуклеиновой кислоты, включающей кодирующую последовательность нуклеиновой кислоты, т.е. последовательность рекомбинантной ДНК представляющего интерес гена, например, гена гликозилтрансферазы или гена *fred*, и не кодирующую последовательность ДНК, например, последовательность промоторной ДНК, например, рекомбинантную промоторную последовательность, полученную из промотора оперона *lac* или оперона *glp*, или промоторную последовательность, полученную из другой последовательности ДНК геномного промотора, или синтетическая промоторная последовательность, где кодирующая и промоторная последовательности функционально связаны. Термин

«функционально связанный» относится к функциональной взаимосвязи между двумя или более сегментами нуклеиновой кислоты (например, ДНК). Как правило, это относится к функциональной связи регуляторной последовательности транскрипции с транскрибируемой последовательностью. Например, промоторная последовательность функционально связана с кодирующей последовательностью, если она стимулирует или модулирует транскрипцию кодирующей последовательности в подходящей клетке-хозяине или другой системе экспрессии. Как правило, промоторные регуляторные последовательности транскрипции, которые функционально связаны с транскрибируемой последовательностью, физически примыкают к транскрибируемой последовательности, т.е. они представляют собой *cis*-действующие последовательности.

В одном воплощении конструкция нуклеиновой кислоты по изобретению может представлять собой частью векторной ДНК, в другом воплощении конструкция представляет собой экспрессионную кассету/картридж, интегрированную в геном клетки-хозяина. Соответственно, термин «конструкция нуклеиновой кислоты» означает искусственно сконструированный сегмент нуклеиновой кислоты, в особенности, сегмент ДНК, который предназначен для «трансплантации» в клетку-мишень, например, в клетку-мишень, например, в бактериальную клетку, для модификации экспрессии гена генома или экспрессии последовательности гена/кодирующей ДНК, которая может быть включена в конструкцию. В контексте изобретения конструкция нуклеиновой кислоты включает последовательность рекомбинантной ДНК, включающую две или более последовательностей рекомбинантной ДНК: по существу, некодирующую последовательность ДНК, включающую последовательность промоторной ДНК, и кодирующую последовательность ДНК, кодирующую представляющий интерес ген, например белок Fred, гликозилтрансферазу или другой ген, пригодный для продукции ОГМ в клетке-хозяине. Предпочтительно конструкция включает дополнительные некодирующие последовательности ДНК, которые либо регулируют транскрипцию, либо трансляцию кодирующей ДНК конструкции, например, последовательность ДНК, облегчающую связывание рибосомы с транскриптом, лидирующую последовательность ДНК, которая стабилизирует транскрипт.

Интеграция интересующего рекомбинантной нуклеиновой кислоты, включенной в конструкции (в кассету экспрессии), в бактериальный геном может быть достигнута обычными способами, например, с помощью линейных картриджей, содержащих фланкирующие последовательности, гомологичные определенному участку на хромосоме, как описано для attTn7-сайта (Waddell C.S. and Craig N.L., *Genes Dev.* (1988) Feb;2(2):137-49.); способами геномной интеграции последовательностей нуклеиновых кислот, в

которых рекомбинация опосредована рекомбиназной функцией Red фага λ или рекомбиназной функцией RecE/RecT профага Rac (Murphy, *J Bacteriol.* (1998);180(8):2063-7; Zhang et al., *Nature Genetics* (1998)20: 123-128 Muyrers et al., *EMBO Rep.*(2000) 1(3): 239–243); способами, основанными на рекомбинации Red/ET (Wenzel et al., *Chem Biol.* (2005), 12(3):349-56.; Vetcher et al., *Appl Environ Microbiol.* (2005);71(4):1829-35); или с помощью положительных клонов, т.е. клонов, которые несут кассету экспрессии, могут быть отобраны, например, с помощью маркерного гена, или потери или усиления функции гена.

Одной копии кассеты экспрессии, включающей интересующий ген, может быть достаточно для обеспечения продукции желаемого ОГМ и достижения желаемых эффектов согласно изобретению. Соответственно, в некоторых предпочтительных воплощениях изобретение относится к рекомбинантной ОГМ-продуцирующей клетке, которая включает одну, две или три копии представляющего интерес гена, интегрированного в геномную ДНК клетки. В некоторых воплощениях предпочтительна единственная копия гена.

В одном предпочтительном воплощении рекомбинантная кодирующая последовательность нуклеиновой кислоты конструкции нуклеиновой кислоты по изобретению представляет собой гетерологичную последовательность по отношению к промотору, что означает, что эквивалентная нативная кодирующая последовательность в геноме вида происхождения транскрибируется под контролем другой промоторной последовательности (т.е. не промоторной последовательности конструкции). Тем не менее, что касается клетки-хозяина, кодирующая ДНК может быть либо гетерологичной (т.е. полученной из другого биологического вида или рода), такой как, например, последовательность ДНК, кодирующая белок Fred, экспрессированная в клетках-хозяевах *Escherichia coli*, или гомологичной (т.е. полученной из клетки-хозяина), такой как, например, гены оперона толстой кишки, гены *wca*.

Предпочтительно конструкция по настоящему изобретению, включающая ген, связанный с биосинтетической продукцией ОГМ, промоторную последовательность ДНК и другие регуляторные последовательности, такие как последовательность сайта связывания рибосомы (например, последовательность Шайна-Дальгарно), экспрессируемая в генетически модифицированной клетке, обеспечивает ОГМ на уровне не менее 0,03 г /OD (оптическая плотность) на 1 л ферментационной среды, включающей суспензию генетически модифицированных клеток, например, на уровне от около 0,05 г/л/OD до около 0,5 г/л/OD. Для целей изобретения последний уровень продукции ОГМ считается «достаточным», а генетически модифицированная клетка, способная

продуцировать желаемый уровень ОГМ, считается «подходящей генетически модифицированной клеткой», т.е. клетка может быть дополнительно модифицирована для экспрессии белка-переносчика ОГМ, например, Fred, для достижения хотя бы одного из описанных в настоящем документе эффектов, выгодных для производства ОГМ.

Генетически модифицированная клетка или конструкция нуклеиновой кислоты по настоящему изобретению включает последовательность нуклеиновой кислоты, такую как гетерологичный ген, кодирующий предполагаемый белок-переносчик MFS (суперсемейство основных посредников).

Белок Fred – это MFS-транспортер, представляющий особый интерес в настоящем изобретении. Таким образом, конструкция нуклеиновой кислоты по настоящему изобретению включает последовательность нуклеиновой кислоты, имеющую, по меньшей мере, 70% идентичности последовательности с геном, *fred*, SEQ ID NO: 2.

Последовательность нуклеиновой кислоты, содержащаяся в генетически модифицированной клетке или в конструкции нуклеиновой кислоты, кодирует белок последовательности SEQ ID NO: 1 или его функциональный гомолог, аминокислотная последовательность которого, по меньшей мере, на более чем 95,4% идентична SEQ ID NO: 1.

Функциональный гомолог белка последовательности SEQ ID NO: 1 может быть получен путем мутагенеза. Функциональный гомолог должен иметь оставшуюся функциональность не менее 50%, такую как 60%, 70%, 80%, 90% или 100% по сравнению с функциональностью аминокислотной последовательности SEQ ID NO: 1. Функциональный гомолог может иметь более высокую функциональность по сравнению с функциональностью аминокислотной последовательности SEQ ID NO: 1. Функциональный гомолог последовательности SEQ ID NO: 1 должен быть способен усиливать продукцию ОГМ генетически модифицированной клетки по изобретению.

Генетически модифицированная клетка

Термин «генетически модифицированная клетка», применяемый в настоящем документе, понимают как клетку, которая была трансформирована или трансфицирована гетерологичной полинуклеотидной последовательностью. В данном контексте термины «генетически модифицированная клетка» и «клетка-хозяин» применяют взаимозаменяемо.

Соответственно, термин «генетически модифицированная клетка» или «генетически модифицированная клетка» в данном контексте понимают как клетку-хозяина, которая была трансформирована или трансфицирована экзогенной полинуклеотидной последовательностью.

Генетически модифицированная клетка предпочтительно представляет собой

прокариотическую клетку. Подходящие микробные клетки, которые могут функционировать в качестве клетки-хозяина, включают дрожжевые клетки, бактериальные клетки, клетки архебактерий, клетки водорослей и клетки грибов.

Клетка-хозяин

Генетически модифицированная клетка (клетка-хозяин или рекомбинантная клетка) может представлять собой, например, бактериальную или дрожжевую клетку. В одном предпочтительном воплощении генетически модифицированная клетка представляет собой бактериальную клетку.

Что касается бактериальных клеток-хозяев, то, в принципе ограничений нет; они могут представлять собой эубактерии (грамм-положительные или грамм-отрицательные) или архебактерии, если они допускают генетические манипуляции для вставки интересующего гена и их можно культивировать в производственных масштабах. Предпочтительно клетка-хозяин обладает свойством культивирования до высокой плотности клеток. Неограничивающими примерами бактериальных клеток-хозяев, подходящих для рекомбинантного промышленного производства ОГМ согласно изобретению, могут представлять собой *Erwinia herbicola* (*Pantoea agglomerans*), *Citrobacter freundii*, *Pantoea citrea*, *Pectobacterium carotovorum* или *Xanthomonas campestris*. Также могут быть применены бактерии рода *Bacillus*, в том числе *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus coagulans*, *Bacillus thermophilus*, *Bacillus laterosporus*, *Bacillus megaterium*, *Bacillus mycoides*, *Bacillus pumilus*, *Bacillus lentus*, *Bacillus cereus* и *Bacillus circulans*. Аналогично, бактерии родов *Lactobacillus* и *Lactococcus* могут быть модифицированы с применением способов по данному изобретению, включая, но не ограничиваясь ими, *Lactobacillus acidophilus*, *Lactobacillus salivarius*, *Lactobacillus plantarum*, *Lactobacillus helveticus*, *Lactobacillus delbrueckii*, *Lactobacillus rhamnosus*, *Lactobacillus bulgaricus*, *Lactobacillus crispatus*, *Lactobacillus gasseri*, *Lactobacillus casei*, *Lactobacillus reuteri*, *Lactobacillus jensenii* и *Lactococcus lactis*. *Streptococcus thermophiles* и *Propionibacterium freudenreichii* также представляют собой подходящие виды бактерий для описанного в настоящем документе изобретения. Также часть этого изобретения представляют собой штаммы, модифицированные, как описано в настоящем документе, из родов *Enterococcus* (например, *Enterococcus faecium* и *Enterococcus thermophiles*), *Bifidobacterium* (например, *Bifidobacterium longum*, *Bifidobacterium infantis* и *Bifidobacterium bifidum*), *Sporolactobacillus* spp., *Micromonospora* spp., *Micrococcus* spp., *Rhodococcus* spp. и *Pseudomonas* (например, *Pseudomonas fluorescens* и *Pseudomonas aeruginosa*).

Бактерии, включающие характеристики, описанные в настоящем документе,

культивируют в присутствии лактозы, и олигосахарид, такой как ОГМ, продуцируемый клеткой, выделяют либо из самой бактерии, либо из культурального супернатанта бактерии. В одном предпочтительном воплощении генетически модифицированная клетка по изобретению представляет собой клетку *Escherichia coli*.

В другом предпочтительном воплощении клетка-хозяин представляет собой дрожжевую клетку, например, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Pichia pastoris*, *Kluveromyces lactis*, *Kluveromyces marxianus* и т.п.

Генетически модифицированные клетки по изобретению могут быть получены с применением стандартных способов в данной области техники, например, описанных в руководствах *Sambrook et al.*, *Wilson & Walker*, *Maniatis et al.* и *Ausubel et al.*

Хозяин, подходящий для производства ОГМ, например *E. coli*, может включать эндогенный ген β -галактозидазы или экзогенный ген β -галактозидазы, например, *E. coli* включает эндогенный ген *lacZ* (например, номер доступа в GenBank V00296 (GI:41901)). Для целей изобретения клетка-хозяин, продуцирующая ОГМ, подвергается генетическим манипуляциям либо для включения любого гена β -галактозидазы, либо для включения гена, который инактивирован. Ген может быть инактивирован путем полной или частичной делеции соответствующей последовательности нуклеиновой кислоты из бактериального генома, или последовательность гена может быть мутирована так, как она транскрибируется, или, если транскрибируется, транскрипт не транслируется, или если транслируется в белок (т.е. β -галактозидазу), то белок не обладает соответствующей ферментативной активностью. Таким образом, бактерия, продуцирующая ОГМ, накапливает повышенный внутриклеточный пул лактозы, что полезно для продукции ОГМ.

В некоторых воплощениях сконструированная бактерия включает дефицитный катаболический путь сиаловой кислоты. Под «катаболическим путем сиаловой кислоты» подразумевается последовательность реакций, обычно контролируемая и катализируемая ферментами, которая приводит к разложению сиаловой кислоты. Примерный путь катаболизма сиаловой кислоты, описанный в настоящем документе, представляет собой путь *E. coli*. В этом пути сиаловая кислота (Neu5Ac; N-ацетилнейраминная кислота) расщепляется ферментами NanA (лиаза N-ацетилнейраминной кислоты), NanK (N-ацетилманнозаминкиназа) и NanE (N-ацетилманнозамин-6-фосфатэпимераза), все кодируемые из оперона *nanATEK-yhcH*, и подавляется NanR (<http://ecocyc.org/ECOLI>). Недостаточный катаболический путь сиаловой кислоты воспроизводится у хозяина *E. coli* путем внесения мутации в эндогенные гены *nanA* (N-ацетилнейраминалиаза) (например, номер доступа в GenBank D00067.1 (GL216588)) и/или *nanK* (N-ацетилманнозаминкиназа)

(например, номер доступа в GenBank (аминокислота) BAE77265.1 (GL85676015)), и/или *nanE* (N-ацетилманнозамин-6-фосфатэпимераза, GI: 947745, включена сюда посредством отсылки). Необязательно, ген *nanT* (транспортер N-ацетилнейрамината) также инактивируют или мутируют. Другие промежуточные продукты метаболизма сиаловой кислоты включают: (ManNAc-6-P) N-ацетилманнозамин-6-фосфат; (GlcNAc-6-P) N-ацетилглюкозамин-6-фосфат; (GlcN-6-P) глюкозамин-6-фосфат и (Fruс-6-P) фруктозо-6-фосфат. В некоторых предпочтительных воплощениях, *nanA* мутирован. В других предпочтительных воплощениях, *nanA* и *nanK* мутированы, тогда как *nanE* остается функциональным. В другом предпочтительном воплощении, *nanA* и *nanE* мутированы, тогда как *nanK* не был мутирован, инактивирован или удален. Мутация представляет собой одно или несколько изменений в последовательности нуклеиновой кислоты, кодирующей продукт гена *nanA*, *nanK*, *nanE* и/или *nanT*. Например, мутация может представлять собой 1, 2, до 5, до 10, до 25, до 50 или до 100 изменений в последовательности нуклеиновой кислоты. Например, *nanA*, *nanK*, *nanE* и/или *nanT* гены мутируют путем нулевой мутации. Нулевые мутации, как описано в настоящем документе, включают аминокислотные замены, добавления, делеции или вставки, которые либо вызывают потерю функции фермента (т.е. снижение или отсутствие активности), либо потерю фермента (т.е. отсутствие генного продукта). Под «удаленными» подразумевают, что кодирующая область удалена полностью или частично, так что не образуется (функциональный) продукт гена. Под инактивацией подразумевается, что кодирующая последовательность была изменена таким образом, что полученный генный продукт будет функционально неактивным или будет кодировать генный продукт с активностью, менее чем на 100%, например, на 90%, 80%, 70%, 60%, 50%, 40%, 30% или 20% по сравнению с активностью нативного, встречающегося в природе, эндогенного генного продукта. «Немутированный» ген или белок не отличается от нативной, встречающейся в природе или эндогенной кодирующей последовательности на 1, 2, вплоть до 5, вплоть до 10, вплоть до 20, вплоть до 50, вплоть до 100, вплоть до от 200 или вплоть до 500 или более кодонов или до соответствующей закодированной аминокислотной последовательности.

Кроме того, бактерия (например, *E. coli*) также обладает способностью к синтезу сиаловой кислоты. Например, бактерия обладает способностью синтезировать сиаловую кислоту за счет экзогенной UDP-GlcNAc 2-эпимеразы (например, *neuC* из *Campylobacter jejuni* (GenBank AAK91727.1) или ее эквивалента (например, (GenBank CAR04561.1), синтазы Neu5Ac (например, *neuB* из *C. jejuni* (GenBank AAK91726.1) или ее эквивалента (например, синтазы сиаловой кислоты из *Flavobacterium limnosediminis*, GenBank

WP_023580510.1), и/или синтазы CMP-Neu5Ac (например, *neuA* из *C. jejuni* (GenBank AAK91728.1) или ее эквивалента (например, CMP-синтазы сиаловой кислоты из *Vibrio brasiliensis*, GenBank WP_006881452.1).

Получение нейтрального ОГМ, содержащего N-ацетилглюкозамин, в модифицированных бактериях также известно в данной области техники (смотри, например, Gebus C et al (2012) Carbohydrate Research 363 83-90).

Для производства ОГМ, содержащих N-ацетилглюкозамин, таких как лакто-N-триоза 2 (LNT-2), лакто-N-тетраоза (LNT), лакто-N-неотетраоза (LNnT), лакто-N-фукопентаоза I (LNFP-I), лакто-N-фукопентаоза II (LNFP-II), лакто-N-фукопентаоза III (LNFP-III), лакто-N-фукопентаоза V (LNFP-V), лакто-N-дифукогексаоза I (LDFH-I), лакто-N-дифукогексаоза II (LDFH-II) и лакто-N-неодифукогексаоза II (LNDFH-III), как описано выше, и она модифицирована так, чтобы включать экзогенный ген UDP-GlcNAc:Gal α / β -R β -3-N-ацетилглюкозаминилтрансферазы или его функциональный вариант или фрагмент. Этот экзогенный ген UDP-GlcNAc:Gal α / β -R β -3-N-ацетилглюкозаминилтрансферазы может быть получен из любого из ряда источников, например, из гена *lgtA*, описанного в *N. meningitidis* (номер доступа белка в GenBank AAF42258.1) или *N. gonorrhoeae* (номер доступа белка в GenBank ACF31229.1). Необязательно дополнительный ген экзогенной гликозилтрансферазы может быть коэкспрессирован в бактерии, включающей экзогенный ген UDP-GlcNAc:Gal α / β -R β -3-N-ацетилглюкозаминилтрансферазы. Например, ген β -1,4-галактозилтрансферазы коэкспрессируется с геном UDP-GlcNAc:Gal α / β -R β -3-N-ацетилглюкозаминилтрансферазы. Этот экзогенный ген β -1,4-галактозилтрансферазы может быть получен из любого источника, например, описанного в *N. meningitidis*, ген *lgtB* (номер доступа белка в GenBank AAF42257.1) или из *H. pylori*, ген *HP0826/galT* (номер доступа белка в GenBank NP_207619.1). Необязательно дополнительный ген экзогенной гликозилтрансферазы, коэкспрессируемый в бактерии, включающей экзогенный ген UDP-GlcNAc:Gal α / β -R β -3-N-ацетилглюкозаминилтрансферазы, представляет собой ген P-1,3-галактозилтрансферазы, например, тот, что описан из *E. coli* 055:H7, ген *wbgO* (номер доступа белка в GenBank WP_000582563.1), или из *H. pylori*, ген *jhp0563* (номер доступа белка в GenBank AEZ55696.1), или из *Streptococcus agalactiae* тип *Ib* O12 ген *cpsIBJ* g (номер доступа белка в GenBank AB050723). Раскрытое изобретение также охватывает функциональные варианты и фрагменты любого из ферментов, описанных выше.

Ген N-ацетилглюкозаминилтрансферазы и/или ген галактозилтрансферазы также может быть функционально связан с *Pglp* и экспрессироваться из соответствующей интегрированной в геном кассеты. В одном воплощении ген, интегрированный в геном,

представляет собой ген, кодирующий галактозилтрансферазу, например, ген *HP0826*, кодирующий фермент GalT из *H. pylori* (номер доступа белка в GenBank NP_207619.1); в другом воплощении ген, интегрированный в геном, представляет собой ген, кодирующий β -1,3-N-ацетилглюкозаминилтрансферазу, например, ген *lgtA* из *N. meningitidis* (номер доступа белка в GenBank AAF42258.1). В этих воплощениях второй ген, т.е. ген, кодирующий β -1,3-N-ацетилглюкозаминилтрансферазу или галактозилтрансферазу, соответственно, может экспрессироваться либо из интегрированной в геном, либо из переносимой плазмидой кассеты. Второй ген необязательно может экспрессироваться либо под контролем промотора *glp* или под контролем любого другого промотора, подходящего для системы экспрессии, например, *Plac*.

Клетки-хозяева, продуцирующие ОГМ, обычно включают функциональный ген *lacY* и дисфункциональный ген *lacZ*.

ОГМ, продуцируемые рекомбинантными клетками по изобретению, могут быть очищены с применением подходящей процедуры, доступной в данной области техники (например, такой, как описана в WO 2015188834, WO 2017182965 или WO 2017152918).

Кодируемый полипептид, способный транспортировать сахар

Транспорт сахара относится к транспорту сахара, такого как, но не ограничиваясь этим, олигосахарида, а в связи с изобретением к транспорту притока и/или оттока одного и/или нескольких ОГМ из цитоплазмы или периплазмы генетически модифицированной клетки в продукционную среду и/или из продукционной среды в цитоплазму или периплазму генетически модифицированной клетки. Таким образом, экспрессируемый в генетически модифицированной клетке полипептид, способный транспортировать ГМО из цитоплазмы или периплазмы в продукционную среду и/или из продукционной среды в цитоплазму или периплазму генетически модифицированной клетки, представляет собой полипептид, способный к транспорту сахара. Таким образом, в настоящем изобретении транспортировка сахара может означать транспортировку оттока и/или притока сахара, такого как, но, не ограничиваясь этим, олигосахарида.

В этом отношении полипептид, способный транспортировать сахар, представляет собой полипептид, принадлежащий к суперсемейству основных переносчиков (MFS). В особенности, полипептид имеет более чем 95,4%, например, по меньшей мере, 95,5%, например, по меньшей мере, 96%, 97%, 98% или 99% идентичности последовательности с SEQ ID NO:1, или он представляет собой его функциональный вариант, как описано в настоящем документе. SEQ ID NO:1 представляет собой аминокислотную последовательность белка Fred.

Генетически модифицированная клетка или конструкция нуклеиновой кислоты по

настоящему изобретению включает последовательность нуклеиновой кислоты, такую как гетерологичный ген, кодирующий белок Fred. Указанная последовательность нуклеиновой кислоты имеет, по меньшей мере, 70% идентичности последовательности с геном *fred*, как показано в SEQ ID No: 2, например, по меньшей мере, 75%, 80%, 85%, 90%, 95% или 99%.

Конструкция последовательности нуклеиновой кислоты кодирует белок SEQ ID NO: 1 или его функциональный гомолог, аминокислотная последовательность которого более чем на 95,4%, например, по меньшей мере, на 95,5%, 96%, 97%, 98 или 99% идентична SEQ ID NO: 1.

Функциональный гомолог белка SEQ ID NO: 1 может быть получен путем мутагенеза. Функциональный гомолог должен иметь остаточную функциональность по меньшей мере, 50%, например, 60%, 70%, 80%, 90% или 100%, по сравнению с функциональностью аминокислотной последовательности SEQ ID NO: 1. Функциональный гомолог может иметь более высокую функциональность по сравнению с функциональностью аминокислотной последовательности SEQ ID NO: 1. Функциональный гомолог SEQ ID NO: 1 должен быть способен усиливать продукцию ОГМ генетически модифицированной клеткой согласно изобретению.

Генетически модифицированная клетка или конструкция нуклеиновой кислоты может содержать одну или несколько последовательностей нуклеиновой кислоты, кодирующих полипептид, способный к транспортировке сахара. Чаще генетически модифицированная клетка или конструкция нуклеиновой кислоты по изобретению кодирует единственную копию полипептида, способного транспортировать сахар.

Одной копии экспрессионной кассеты, включающей интересующий ген, может быть достаточно для обеспечения продукции желаемого ОГМ и достижения желаемых эффектов согласно изобретению. Соответственно, в некоторых предпочтительных вариантах осуществления изобретение относится к клетке, продуцирующей рекомбинантный ОГМ, которая включает одну, две или три копии интересующего гена или генов, интегрированных в геномную ДНК клетки. В некоторых воплощениях предпочтительна единственная копия гена или генов.

Генетически модифицированная клетка или конструкция нуклеиновой кислоты может также содержать один или несколько регуляторных элементов для регуляции экспрессии последовательности нуклеиновой кислоты, кодирующей полипептид, транспортирующий сахар, и при этом указанная последовательность нуклеиновой кислоты имеет, по меньшей мере, 70% идентичности последовательности с SEQ ID. NO: 2.

Регуляторный элемент

Как указано выше, генетически модифицированная клетка или конструкция нуклеиновой кислоты могут дополнительно включать последовательность нуклеиновой кислоты, включающую регуляторный элемент для регуляции экспрессии последовательности нуклеиновой кислоты, имеющей, по меньшей мере, 70% идентичности последовательности с SEQ ID NO: 2. Последовательность нуклеиновой кислоты регуляторной области может быть гетерологичной или гомологичной.

Термин «регуляторный элемент», или «промотор», или «промоторная область», или «промоторный элемент» представляет собой последовательность нуклеиновой кислоты, которая распознается и связывается ДНК-зависимой РНК-полимеразой во время инициации транскрипции. Промотор вместе с другими последовательностями нуклеиновых кислот, регулирующими транскрипцию и трансляцию (также называемыми «контрольными последовательностями»), необходим для экспрессии данного гена или группы генов (оперона). Как правило, транскрипционные и трансляционные регуляторные последовательности включают, но не ограничиваются ими, промоторные последовательности, сайты связывания рибосом, последовательности начала и остановки транскрипции, последовательности начала и остановки трансляции и энхансерные или активаторные последовательности. «Сайт начала транскрипции» означает первый транскрибируемый нуклеотид и обозначен как +1. Нуклеотиды ниже стартового сайта пронумерованы +2, +3, +4 и т.д., а нуклеотиды в 5' противоположном (выше) направлении пронумерованы -1, -2, -3 и т.д. Промотор конструкции может происходить из промоторной области любого гена, закодированного в геноме вида. Предпочтительна промоторная область геномной ДНК *E.coli*. Соответственно, любой промотор, способный связываться с РНК-полимеразой и инициировать транскрипцию, подходит для осуществления изобретения. В принципе, для контроля транскрипции рекомбинантного гена можно применять любой промотор, такой как транспортер MFS или гликозилтрансферазы по изобретению. При осуществлении изобретения могут быть применены разные или идентичные промоторные последовательности для запуска транскрипции различных представляющих интерес генов, интегрированных в геном клетки-хозяина или в ДНК вектора экспрессии. В одном примере промоторная последовательность А способствует экспрессии переносчика MFS, а другая промоторная последовательность В, или идентичная промоторная последовательность А, способствует экспрессии гликозилтрансферазы.

Для оптимальной экспрессии рекомбинантных генов, включенных в конструкцию, конструкция может включать дополнительные регуляторные последовательности,

например ведущая последовательность ДНК, такая как последовательность ДНК, полученная из 5'-нетранслируемой области (5'UTR) гена *glp* *E. coli*, последовательность для рибосомного связывания. Примеры более поздних последовательностей описаны в WO 2019123324 (включена в настоящий документ в качестве отсылки) и проиллюстрированы в настоящем документе в неограничивающих рабочих примерах.

В одном аспекте изобретения один или несколько регуляторных элементов могут быть вставлены в конструкцию ДНК, кодирующую ген *fred* согласно SEQ ID NO: 2, и/или в конструкцию ДНК, кодирующую одну или несколько гликозилтрансфераз согласно изобретению. В другом аспекте изобретения одна или несколько копий гена *fred* в соответствии с SEQ ID NO: 2 могут быть встроены в конструкцию ДНК согласно изобретению. В еще одном аспекте изобретения ген *fred* в соответствии с SEQ ID NO: 2 может быть встроен в одну или несколько идентичных или неидентичных конструкций ДНК согласно изобретению.

В одном аспекте изобретения регуляторным элементом для регуляции экспрессии рекомбинантного гена, включенного в конструкцию по изобретению, представляет собой промотор оперона *glpFKX*, *PglpF*. В другом аспекте изобретения промотор представляет собой промотор оперона *lac*, *Plac*. И в еще одном аспекте изобретения регуляторный элемент представляет собой *PglpF_SD4* и/или *PglpF_SD7*, которые представляют собой модифицированную версию последовательности *PglpF*, включающую модифицированную последовательность сайта связывания рибосомы ниже последовательности промотора. Однако любой промотор, обеспечивающий транскрипцию и/или регуляцию уровня транскрипции одной или нескольких рекомбинантных нуклеиновых кислот, кодирующих один или несколько полипептидов по изобретению, представляет собой подходящий промотор для практического применения изобретения.

Как правило, промоторы для экспрессии гетерологичного гена по настоящему изобретению выбирают из таблицы 2.

Таблица 2

Название промотора	Описание	Ссылка	SEQ ID NO:
PgatY_70UTR	Промотор <i>E. coli</i> для <i>gatYZABCD</i> ; тагатозо-1,6-бисP-альдолаза	NA	3
PglpF	Промотор <i>E. coli</i> для <i>glpFKX</i> ; поглощение глицерина	WO2019123324	4
PglpF_SD1	Промотор <i>E. coli</i> для оперона <i>glpFKX</i> ; поглощение глицерина	WO2019123324	
PglpF_SD10	Промотор <i>E. coli</i> для оперона <i>glpFKX</i> ; поглощение глицерина	WO2019123324	
PglpF_SD2	Промотор <i>E. coli</i> для оперона <i>glpFKX</i> ; поглощение глицерина	WO2019123324	
PglpF_SD3	Промотор <i>E. coli</i> для оперона <i>glpFKX</i> ; поглощение глицерина	WO2019123324	

	поглощение глицерина		
PglpF_SD4	Промотор <i>E. coli</i> для оперона glpFKX; поглощение глицерина	WO2019123324	5
PglpF_SD5	Промотор <i>E. coli</i> для glpFKX; поглощение глицерина	WO2019123324	
PglpF_SD6	Промотор <i>E. coli</i> для оперона glpFKX; поглощение глицерина	WO2019123324	
PglpF_SD7	Промотор <i>E. coli</i> для оперона glpFKX; поглощение глицерина	WO2019123324	6
PglpF_SD8	Промотор <i>E. coli</i> для оперона glpFKX; поглощение глицерина	WO2019123324	
PglpF_SD9	Промотор <i>E. coli</i> для оперона glpFKX; поглощение глицерина	WO2019123324	
Plac_16UTR	Промотор <i>E. coli</i> для оперона lacZYA: lac-оперон	WO2019123324	
Plac	Промотор <i>E. coli</i> для оперона lacZYA: lac-оперон	WO2019123324	19
PmglB_70UTR	Промотор <i>E. coli</i> для mglBAC; транспортер галактозы/метилгалактозидазы	NA	7
PmglB_70UTR SD4	Промотор <i>E. coli</i> для mglBAC; транспортер галактозы/метилгалактозидазы	NA	8

Предпочтительные регуляторные элементы, присутствующие в генетически модифицированной клетке или в конструкции нуклеиновой кислоты по настоящему изобретению, выбирают из группы, состоящей из: *PgatY_70UTR*, *PglpF*, *PglpF_SD1*, *PglpF_SD10*, *PglpF_SD2*, *PglpF_SD3*, *PglpF_SD4*, *PglpF_SD5*, *PglpF_SD6*, *PglpF_SD7*, *PglpF_SD8*, *PglpF_SD9*, *Plac_16UTR*, *Plac*, *PmglB_70UTR* и *PmglB_70UTR_SD4*.

Особенно предпочтительные регуляторные элементы, присутствующие в генетически модифицированной клетке или в конструкции нуклеиновой кислоты по настоящему изобретению, выбирают из группы, состоящей из *PglpF*, *PglpF_SD4* и *PglpF_SD7*.

Генетически модифицированные клетки для биосинтетического производства

В настоящем изобретении промоторы могут быть либо необходимыми, либо полезными для достижения оптимального уровня биосинтетического производства одной или нескольких ОГМ в генетически модифицированной клетке и обеспечения желаемых эффектов согласно изобретению. Таким образом, промоторная последовательность по настоящему изобретению обеспечивает транскрипцию и/или регулирует экспрессию полипептида, способного к транспорту сахара, и/или гликозилтрансфераз по изобретению, что приводит к оптимизации биосинтеза и транспорта ОГМ или предшественников ОГМ и/или деградации побочных продуктов получения ОГМ.

В генетически модифицированной клетке по изобретению конструкция нуклеиновой кислоты включена в генетически модифицированную клетку, которая кодирует, по меньшей мере, один ген, связанный с биосинтетическим продуцированием одной или нескольких ОГМ, последовательность промоторной ДНК и другие

регуляторные последовательности, такие как последовательность сайта связывания рибосомы (например, последовательность Шайна-Дальгарно). Экспрессия гена или генов, связанных с биосинтетической продукцией одной или нескольких ОГМ в генетически модифицированной клетке, делает возможным продуцирование ОГМ, делая клетку-хозяина «подходящей клеткой-хозяином» для осуществления изобретения, как описано. В генетически модифицированной клетке экспрессия гена или генов, связанных с биосинтетическим производством ОГМ, как упоминалось выше, позволяет производить одно или несколько ОГМ на уровне 0,03 г/л/OD (оптическая плотность) из 1 литра ферментационной среды, включавшей суспензию генетически модифицированных клеток. Таким образом, уровень ОГМ может быть от около 0,05 г/л/OD до около 0,5 г/л/OD, например, по меньшей мере, 0,4 г/л/OD. Для целей изобретения уровень продукции ОГМ рассматривают как «достаточный», и генетически модифицированная клетка, способная продуцировать желаемый уровень ОГМ или смеси ОГМ, считается подходящей генетически модифицированной клеткой для осуществления изобретения.

Таким образом, в свете изобретения «подходящая генетически модифицированная клетка» может быть дополнительно модифицирована, как описано выше, для экспрессии полипептида сахара, способного к транспортировке сахара семейства MFS, например, *fred*, для достижения, тем или иным образом, выгодного производства ОГМ, такого как, помимо прочего, более высокий уровень ОГМ в биосинтетическом производстве, более высокий уровень чистоты в биосинтетическом производстве, сокращение длительности производства и/или более эффективное биосинтетическое производство ГМО.

Способ получения

Второй аспект изобретения относится к способу получения одного или нескольких ОГМ, включающему следующие этапы:

(i) обеспечение генетически модифицированной клетки, способной продуцировать ОГМ, причем указанная клетка включает рекомбинантную нуклеиновую кислоту, кодирующую белок последовательности SEQ ID NO: 1, или ее функциональный гомолог, аминокислотная последовательность которого идентична более чем на 95,4%, идентична, по меньшей мере, на 95,5%, предпочтительно идентична, по меньшей мере, на 96%, более предпочтительно идентична, по меньшей мере, на 99,9% последовательности SEQ ID NO: 1;

(ii) культивирование клетки (i) в подходящей среде для культивирования клеток для обеспечения продукции ОГМ и экспрессии последовательности ДНК с получением белка, имеющего аминокислотную последовательность SEQ ID NO: 1, или его функциональную аминокислотную последовательность более чем на 95,4% идентична,

например, по меньшей мере, на 95,5% идентична, предпочтительно, по меньшей мере, на 96% идентична, более предпочтительно, по меньшей мере, на 99,9% идентична SEQ ID NO: 1;

(iii) сбор ОГМ, полученного на этапе (ii).

Согласно изобретению термин «культивирование» (или «культивирование» или «культивирование», также называемое «ферментация») относится к размножению бактериальных экспрессирующих клеток в контролируемом биореакторе в соответствии со способами, известными в промышленности.

Для получения одного или нескольких ОГМ бактерии, продуцирующие ОГМ, как описано в настоящем документе, культивируют в соответствии с процедурами, известными в данной области техники, в присутствии подходящего источника углерода, например глюкозы, глицерина, лактозы и т.д., и полученный ОГМ собирают из сред культивирования и микробной биомассы, образующейся в процессе культивирования. После этого ОГМ очищают в соответствии с процедурами, известными в данной области техники, например, такими, как описаны в WO2015188834, WO2017182965 или WO2017152918, и очищенный ОГМ применяют в качестве нутрицевтиков, фармацевтических препаратов или для любых других целей, например, для исследований.

Производство ОГМ обычно осуществляется путем культивирования в больших объемах. Термин «производство» и «производственный масштаб» в значении изобретения определяет ферментацию с минимальным объемом 5 л культурального бульона. Обычно процесс в «производственном масштабе» определяют способностью обрабатывать большие объемы препарата, содержащего продукт, и производить такие количества представляющего интерес белка, которые соответствуют, например, в случае терапевтического соединения или композиции, требованиям для клинических испытаний, а также для предложения на рынке. Помимо большого объема, способ в производственных масштабах, в отличие от простых способов в лабораторных масштабах, таких как культивирование во встряхиваемых колбах, характеризуется применением технической системы биореактора (ферментера), оснащенного устройствами для перемешивания, аэрации, подачи питательных веществ, слежения за параметрами процесса и контроля параметров процесса (pH, температура, давление растворенного кислорода, обратное давление и др.). В значительной степени поведение системы экспрессии в лабораторном масштабе, например, во встряхиваемых колбах, настольных биореакторах или формате глубоких лунок, описанных в примерах настоящего раскрытия, позволяет предсказать поведение этой системы этой системы в сложной среде биореактора.

Что касается подходящей среды для клеток, применяемой в процессе ферментации, ограничений нет. Культуральная среда может быть полуопределенной, то есть содержащей комплексные соединения среды (например, дрожжевой экстракт, соевый пептон, казаминовые кислоты и т.д.), или она может быть химически определенной, без каких-либо комплексных соединений.

Под термином «один или несколько ОГМ» подразумевают, что производственная ячейка ОГМ может быть способна производить одну структуру ОГМ (первый ОГМ) или несколько структур ОГМ (второй, третий и т.д. ОГМ). В некоторых воплощениях может быть предпочтительной генетически модифицированная клетка, которая продуцирует один ОГМ, в других предпочтительных воплощениях может быть предпочтительной генетически модифицированная клетка, продуцирующая несколько структур ОГМ. Клетки, продуцирующие 2'-FL, 3-FL, 3'-SL, 6'-SL или LNT-2 представляют собой неограничивающие примеры генетически модифицированных клеток, продуцирующих ОГМ единственной структуры. Неограничивающие примеры генетически модифицированных клеток, способных продуцировать несколько структур ОГМ, могут представлять собой клетки-продуценты DFL, FSL, LNT, LNnT, LNFP-I, LNFP-II, LNFP-III, LNFP-IV, LNFP-V, pLNnH, pLNH2.

В особенности, настоящее изобретение относится к генетически модифицированным клеткам, продуцирующим одну или несколько одиночных структур ОГМ, которые выбирают из группы, состоящей из клеток, продуцирующих 2'-FL, 3-FL, DFL и LNT.

Термин «сбор» в контексте изобретения относится к сбору полученных ОГМ после прекращения ферментации. В различных воплощениях это может включать сбор ОГМ, включенных как в биомассу (т.е. генетически модифицированные клетки), так и в среду культивирования, т.е. до/без отделения ферментационного бульона от биомассы. В других воплощениях полученный ОГМ можно собирать отдельно от биомассы и ферментационного бульона, т.е. после отделения /следом за отделением биомассы от среды культивирования (т.е. ферментационного бульона). Отделение клеток от среды может быть осуществлено любым из способов, хорошо известных специалисту в данной области техники, например, любым подходящим типом центрифугирования или фильтрации. Отделение клеток от среды может следовать сразу за сбором ферментационного бульона или осуществляться на более позднем этапе после хранения ферментационного бульона в соответствующих условиях. Извлечение произведенных ОГМ из оставшейся биомассы (или полная ферментация) включает их извлечение из биомассы (клеток-продуцентов). Это можно осуществить любыми подходящими

способами в данной области техники, например, обработкой ультразвуком, кипячением, гомогенизацией, ферментативным лизисом с применением лизоцима или замораживанием и измельчением.

После восстановления после ферментации ОГМ (доступен) доступны для дальнейшей обработки и очистки.

Очистку ОГМ, полученных ферментацией, можно проводить с применением подходящей процедуры, описанной в WO2016095924, WO2015188834, WO2017152918, WO2017182965, US20190119314 (все включены посредством отсылки).

В некоторых воплощениях изобретения генетически модифицированная клетка может продуцировать несколько ОГМ, где один ОГМ представляет собой ОГМ-«продукт», а некоторые/все прочие ОГМ представляют собой ОГМ-«побочный продукт». Как правило, побочные ОГМ представляют собой либо основных предшественников ОГМ, либо продукты дальнейшей модификации основных ОГМ. В некоторых воплощениях может быть желательно производить продукт ОГМ в больших количествах и побочный продукт ОГМ в малых количествах. Клетки и способы получения ОГМ, описанные в настоящем документе, обеспечивают контролируемое производство продукта ОГМ с заданным профилем ОГМ, например, в одном воплощении полученная смесь ОГМ, в которой продукт ОГМ представляет собой доминирующим ОГМ по сравнению с другими ОГМ (т.е. побочными ОГМ) смеси, т.е. ОГМ-продукт производится в большем количестве, чем другие ОГМ-побочные продукты; в других воплощениях клетка, производящая одну и ту же смесь ОГМ, может быть настроена на производство одного или нескольких побочных продуктов ОГМ в большем количестве, чем ОГМ-продукт. Например, при производстве 2'-FL и/или 3-FL ,ОГМ-продукта, часто образуется значительное количество DFL, ОГМ-побочного продукта. С генетически модифицированными клетками по настоящему изобретению уровень DFL в 2'-FL и/или 3-FL -продукте может быть значительно снижен.

Преимущественно изобретение обеспечивает как снижение отношения побочного продукта к продукту, так и увеличение общего выхода продукта (и/или ОГМ в целом). Это, сниженное образование побочных продуктов по сравнению с образованием продукта, способствует повышенному образованию продукта и повышает эффективность как производства, так и процесса извлечения продукта, обеспечивая превосходную технологию производства ОГМ.

В различных предпочтительных воплощениях могут быть выбраны различные генетически модифицированные клетки, продуцирующие 2'-FL, 3-FL, DFL и/или LNT, в качестве ОГМ продукта или побочного продукта. В одном предпочтительном воплощении

продукт представляет собой 3-FL. В другом предпочтительном воплощении продукт представляет собой 2'-FL, а побочный продукт представляет собой DFL. В другом предпочтительном воплощении продукт представляет собой LNT-2, а побочный продукт представляет собой LNT and LNFP I.

2'-FL, LNT и 3-FL представляют собой предпочтительные ОГМ-продукты по изобретению, как описано.

Применение

Настоящее изобретение также относится к применению клетки-хозяина, описанной в настоящем документе, для получения одного или нескольких олигосахаридов грудного молока (ОГМ). В особенности, изобретение относится к применению клетки-хозяина, как описано в настоящем документе, для продукции конкретного ОГМ, при этом клетка-хозяин выбирают с целью получения большей части одного конкретного ОГМ, предпочтительно выбранного из 2'-FL, 3-FL и LNT, в настоящее время наиболее предпочтителен, 3-FL.

Общее

Следует понимать, что любой признак и/или аспект, рассмотренный выше в связи с описываемым изобретением, применим по аналогии с описанными в настоящем документе способами.

Следующие фигуры и примеры представлены ниже для иллюстрации настоящего изобретения. Они предназначены для иллюстрации и никоим образом не должны рассматриваться как ограничивающие.

Примеры

Материалы и методы

Если не указано иное, для манипулирования с нуклеиновыми кислотами, трансформации и экспрессии нуклеиновых кислот применяют стандартные методики, векторы, элементы контрольной последовательности и другие элементы системы экспрессии, известные в области молекулярной биологии. Такие стандартные техники, векторы и элементы можно найти, например, в руководствах: Ausubel et al. (eds.), *Current Protocols in Molecular Biology* (1995) (John Wiley & Sons); Sambrook, Fritsch, & Maniatis (eds.), *Molecular Cloning* (1989) (Cold Spring Harbor Laboratory Press, NY); Berger & Kimmel, *Methods in Enzymology* 152: Guide to Molecular Cloning Techniques (1987) (Academic Press); Bukhari et al. (eds.), *DNA Insertion Elements, Plasmids and Episomes* (1977) (Cold Spring Harbor Laboratory Press, NY); Miller, J.H. *Experiments in molecular genetics* (1972.) (Cold spring Harbor Laboratory Press, NY)

Описанные ниже воплощения выбраны для иллюстрации изобретения и никоим

образом не ограничивают изобретение.

Штаммы

Штаммы, примененные в настоящих примерах, описаны в приведенной ниже таблице 3.

Таблица 3

ID штаммов	Продукт	Соответствующий генотип
DH1	-	<i>F⁻λ⁺endA1recA1relA1gyrA96thi-1glnV44hsdR17(r_K⁻m_K⁻)</i>
MDO	-	<i>E. coli DH1 ΔlacZ ΔlacA, ΔnanKETA, ΔmelA, ΔwcaJ, ΔmdoH wcaF::Plac</i>
Штамм 1	2'-FL	<i>MDO PglpF-CA PglpF-futC ΔlacI</i>
Штамм 2		<i>MDO PglpF-CA PglpF-futC ΔlacI PglpF-fred</i>
Штамм 3	3-FL	<i>MDO PglpF-CA PglpF-futA ΔlacI</i>
Штамм 4		<i>MDO PglpF-CA PglpF-futA ΔlacI PglpF-fred</i>
Штамм 5	LNT	<i>MDO PglpF-lgtA PglpF-galTK</i>
Штамм 6		<i>MDO PglpF-lgtA PglpF-galTK PglpF SD4-fred</i>
Штамм 7		<i>MDO PglpF-lgtA PglpF-galTK PglpF SD7-fred</i>
MP3672	LNnT	<i>MDO PglpF-GlcNAcT PglpF-Gal4T (version 1)*</i>
MP3708		<i>MDO PglpF-GlcNAcT PglpF-Gal4TPlac-yberC0001 9420</i>
MP3972		<i>MDO PglpF-GlcNAcT PglpF-Gal4TPlac-bad</i>
MP3979		<i>MDO PglpF-GlcNAcT PglpF-Gal4TPlac-fred</i>
MP4011		<i>MDO PglpF-GlcNAcT PglpF-Gal4TPlac-marc</i>
MP3994		<i>MDO PglpF-GlcNAcT PglpF-Gal4TPlac-vag(version 1)**</i>
MP3984		<i>MDO PglpF-GlcNAcT PglpF-Gal4TPlac-nec</i>
MP4362		<i>MDO PglpF-GlcNAcT PglpF-Gal4T Plac-yabM</i>
MP3020		<i>MDO PglpF-GlcNAcT PglpF-Gal4T(version 2)*</i>
MP4064		<i>MDO PglpF-GlcNAcT PglpF-Gal4TPglpF-vag</i>
MP4065		<i>MDO PglpF-GlcNAcT PglpF-Gal4TPlac-vag(version 2)**</i>
MP4473	LNT	<i>MDO PglpF-GlcNAcT PglpF-Gal3T</i>
MP4546		<i>MDO PglpF-GlcNAcT PglpF-Gal3T PglpF-vag</i>

CA = *gmd-wcaG-wcaH-wcaI-manC-manB*

*Штаммы MP3672 и MP3020 несут одни и те же гетерологичные *GlcNAcT* и *Gal4T*, но отличаются количеством копий соответствующего гена, кодирующего *GlcNAcT*.

**Штаммы MP3994 и MP4065 несут одни и те же гетерологичные *GlcNAcT* и *Gal4T*, но отличаются числом копий соответствующего гена, кодирующего *GlcNAcT*.

Культивирование

Если не указано иное, штаммы *E. coli* размножали в среде базовой минимальной среде, содержащей 0,2% глюкозы, при 37°C при перемешивании. Чашки с агаром инкубировали при 37°C в течение ночи.

Базовая минимальная среда имела следующий состав: NaOH (1 г/л), KOH (2,5 г/л), KН₂PO₄ (7 г/л), NH₄H₂PO₄ (7 г/л), лимонная кислота (0,5 г/л), раствор микроэлементов (5 мл/л). Исходный раствор микроэлементов содержал: ZnSO₄*7H₂O 0,82 г/л, лимонную кислоту 20 г/л, MnSO₄*H₂O 0,98 г/л, FeSO₄*7H₂O 3,925 г/л, CuSO₄*5H₂O 0,2 г/л. pH базовой минимальной среды доводили до 7,0 с помощью 5 н NaOH и автоклавировали.

Перед инокуляцией в базовую минимальную среду добавляли 1 мМ MgSO₄, 4 мкг/мл тиамина, 0,5% заданного источника углерода (глицерин («Carbosynth»)). Тиамин, антибиотики стерилизовали фильтрованием. Все процентные концентрации для глицерина выражены как объем/объем, а для глюкозы – масса/объем.

Химически компетентные клетки и трансформации

E. coli инокулировали из чашек LB в 5 мл LB, содержащего 0,2% глюкозы, при 37°C при встряхивании до OD₆₀₀ ~0,4. 2 мл культуры собирали центрифугированием в течение 25 секунд при 13000 g. Супернатант удаляли, а клеточный осадок ресуспендировали в 600 мкл холодных растворов TB (10 мМ PIPES, 15 мМ CaCl₂, 250 мМ KCl). Клетки инкубировали на льду в течение 20 минут с последующим осаждением в течение 15 секунд при 13000 g. Супернатант удаляли, а клеточный осадок ресуспендировали в 100 мкл холодного раствора TB. Трансформацию плазмид проводили с применением 100 мкл компетентных клеток и 1-10 нг плазмидной ДНК. Клетки и ДНК инкубировали на льду в течение 20 минут, а затем подвергали тепловому шоку при 42°C в течение 45 секунд. После 2-минутной инкубации на льду добавляли 400 мкл SOC (20 г/л триптона, 5 г/л дрожжевого экстракта, 0,5 г/л NaCl, 0,186 г/л KCl, 10 мМ MgCl₂, 10 мМ MgSO₄ и 20 мМ глюкозы) и культуру клеток инкубировали при 37°C при встряхивании в течение 1-го часа перед посевом на селективные чашки.

Плазмиду трансформировали в химически компетентные клетки TOP10 в условиях, рекомендованных поставщиком («ThermoFisher Scientific»).

ДНК-технологии

Плазмидную ДНК из *E. coli* выделяли с применением набора QIAprep Spin Miniprep («Qiagen»). Хромосомную ДНК из *E. coli* выделяли с применением набора QIAmp DNA Mini Kit («Qiagen»). Продукты ПЦР очищали с применением набора QIAquick PCR Purification Kit («Qiagen»). Мастер-микс для ПЦР DreamTaq («ThermoFisher»), мастер-микс для ПЦР с горячим стартом Phusion U («ThermoFisher»), USER Enzyme («New England Biolab») применяли в соответствии с рекомендациями поставщика. Праймеры были поставлены компанией «Eurofins Genomics», Германия. Фрагменты ПЦР и плазмиды были секвенированы компанией «Eurofins Genomics». ПЦР колоний проводили с применением DreamTaq PCR Master Mix в термоциклере T100TM («Bio-Rad»).

Гетерологичные белки, экспрессированные в генетически модифицированных клетках, продуцирующих ОГМ, по данному изобретению, описаны в таблице 4, промоторные элементы, которые применяли в приведенных ниже примерах изобретения, описаны в таблице 5, а олигонуклеотиды, примененные для амплификации остовов плазмиды, промотора, элементов и *fred*, описаны в таблице 6.

Таблица 4. Гетерологичные белки, экспрессированные в генетически модифицированных клетках, продуцирующих ОГМ

Ген	Происхождение гена	Номер доступа белка	Функция белка
<i>futC</i>	<i>Helicobacter pylori</i> 26695	WP_080473865.1	α -1,2-фукозилтрансфераза
<i>futA</i>	<i>Helicobacter pylori</i> 26695	WP_000487428.1	α -1,3- фукозилтрансфераза
<i>lgtA</i>	<i>Neisseria meningitidis</i> 053442	WP_002248149.1	β -1,3- ацетилглюкозаминилтрансфераза
<i>galTK</i>	<i>Helicobacter pylori</i> 43504	BD182026	β -1,3- галактозилтрансфераза
<i>fred</i>	<i>Yersinia frederiksenii</i>	WP_087817556.1	MFS-транспортер
<i>yberC0001_9420</i>	<i>Yersinia bercovieri</i>	EEQ08298.1	Суперсемейство основных переносчиков MFS_1
<i>nec*</i>	<i>Rosenbergiella nectarea</i>	WP_092672081.1	MFS-транспортер
<i>marc*</i>	<i>Serratia marcescens</i>	WP_060448169.1	MFS-транспортер
<i>bad*</i>	<i>Rouxiella badensis</i>	WP_017489914.1	MFS-транспортер
<i>vag*</i>	<i>Pantoea vagans</i>	WP_048785139.1	MFS-транспортер
<i>yabM</i>	<i>Erwinia pyrifoliae</i>	CAY73138.1	Предполагаемый транспортер сахаров MFS

*название гена дано для идентификации нуклеиновой кислоты, кодирующей белок, имеющий аминокислотную последовательность, соответствующую номеру доступа в GenBank, для целей настоящего изобретения

Таблица 5. Элементы синтетической ДНК, примененные для экспрессии *fred*.

Название последовательности	SEQ ID NO	Последовательность (5'-3')	Описание
<i>PglpF</i>	4	GCGGCACGCCTTGCAGATTACGGTTTG CCACACTTTTCATCCTTCTCCTGGTGAC ATAATCCACATCAATCGAAAATGTAA TAAATTTGTTGCGCGAATGATCTAACA AACATGCATCATGTACAATCAGATGGA ATAAATGGCGCGATAACGCTCATTTTA TGACGAGGCACACACATTTTAAGTTTCG ATATTTCTCGTTTTTGCTCGTTAACGAT AAGTTTACAGCATGCCTACAAGCATCG TGGAGGTCCGTGACTTTCACGCATACA ACAAACATTAACCAAGGAGGAAACAG CT	Элемент экспрессии ДНК из 300 нуклеотидов
<i>PglpF_SD4</i>	5	GCGGCACGCCTTGCAGATTACGGTTTG CCACACTTTTCATCCTTCTCCTGGTGAC ATAATCCACATCAATCGAAAATGTAA TAAATTTGTTGCGCGAATGATCTAACA AACATGCATCATGTACAATCAGATGGA ATAAATGGCGCGATAACGCTCATTTTA TGACGAGGCACACACATTTTAAGTTTCG ATATTTCTCGTTTTTGCTCGTTAACGAT	Элемент экспрессии ДНК из 300 нуклеотидов

		AAGTTTACAGCATGCCTACAAGCATCG TGGAGGTCCGTGACTTTTCACGCATACA ACAAACATTAACCAACTAGGAAACAG CT	
<i>PglpF_SD7</i>	6	GCGGCACGCCTTGCAGATTACGGTTTG CCACACTTTTCATCCTTCTCCTGGTGAC ATAATCCACATCAATCGAAAAATGTAA TAAATTTGTTGCGCGAATGATCTAACA AACATGCATCATGTACAATCAGATGGA ATAAATGGCGCGATAACGCTCATTTTA TGACGAGGCACACACATTTTAAGTTCG ATATTTCTCGTTTTTGCTCGTTAACGAT AAGTTTACAGCATGCCTACAAGCATCG TGGAGGTCCGTGACTTTTCACGCATACA ACAAACATTAACCAAGAGCAAAACAG CT	Элемент экспрессии ДНК из 300 нуклеотидов
<i>fred</i>	2	ATGAAGAGCGCGCTGACCTTCAGCCGT CGTATTAACCCGGTTTTCTGGCGTTCT TTGTGGTTGCGTTCCTGAGCGGTATTGC GGGTGCGCTGCAGGCGCCGACCCTGAG CCTGTTCTGAGCACCGAGGTGAAAGT TCGTCCGCTGTGGGTGGGCCTGTTCTA CACCGTTAACGCGATTGCGGGTATCAC CGTGAGCTTTGTTCTGGCGAAGCGTAG CGACCTGCGTGGCGATCGTCGTAAACT GATCCTGGTGTGCTACCTGATGGCGGT TGGTAACTGCCTGCTGTTTCGCGTTTAAAC CGTGACTATCTGACCCTGATTACCGCG GGCGTGCTGCTGGCGGGCGGTTGCGAAC ACCGCGATGCCGCAGATTTTCGCGCTG GCGCGTGAGTACGCGGATAACAGCGC GCGTGAAGTGTTATGTTTAGCAGCAT TATGCGTGCGCAACTGAGCCTGGCGTG GGTTATCGGTCCGCGCTGAGCTTCAT GCTGGCGCTGAACTATGGCTTCACCCT GATGTTTTGCATTGCGGCGGGTATCTTC GTGCTGAGCGCGCTGGTTGTGTGGTTT ATTCTGCCGAGCGTGCAGCGTGCGGAA CCGGTTATGGATGCGCCGACCGTGGCG CAAGGCAGCCTGTTTCGCGGACAAGGAT GTTCTGCTGCTGTTTATTGCGAGCATGC TGATGTGGACCTGCAACACCATGTACA TCATTGATATGCCGCTGTATATCACCG CGAGCCTGGGTCTGCCGGAGCGTCTGG CGGGTCTGCTGATGGGCACCGCGGCGG GTCTGGAAATCCCGATTATGCTGCTGG CGGGCTACAGCGTGCGTCGTTTTGGCA AGCGTAAAATCATGCTGTTTCGCGGTGC TGGCGGGCGTTCTGTTTTATACCGGTCT GGTTCTGTTCAAGTTTAAAAGCGCGCT GATGCTGCTGCAGATTTTCAACGCGAT CTTTATTGGTATCGTGGCGGGTATCGG CATGCTGTACTTCCAAGACCTGATGCC GGGTCGTGCGGGTGCGGCGACCACCCT GTTTACCAACAGCATTAGCACCGGCGT TATCCTGGCGGGCGTGCTGCAAGGTGT TCTGACCGAAACCTGGGGTCAACACAG	ДНК, кодирующая транспортер MFS

		CGTGTATGTTATGGCGATGATTCTGGC GATCCTGAGCCTGATCATTTGCGCGCG TGTGCGTGAAGCGTAA	
<i>Plac</i>	19	ATGCGCAAATTGTGAGTTAGCTCACTC ATTAGGCACCCCAGGCTTTACACTTTA TGCTTCCGGCTCGTATGTTGTGTGGAAT TGTGAGCGGATAACAATTTACACAGG AAACAGCTATGACCATGATTACGCCAA GCGCGCAATTAACCCTCACTAAAGGGA ACAAAAGCTGGGTACCTAAGGAGGAA ACAGCT	Элемент экспрессии ДНК из 195 нуклеотидов

Таблица 6. Олигонуклеотиды, применяемые для амплификации остова плазмиды, промоторных элементов и представляющих интерес генов, включая *fred*.

Название	SEQ ID NO	Олигонуклеотидная последовательность 5'-3'	Описание
O40	9	ATTAACCCUCCAGGCATCAAATAAAAC GAAAGGC	Остов.прямой
O79	10	ATTTGCGCAUCACCAATCAAATTCACG CGGCC	Остов.обратный
O261	11	ATGCGCAA AUGCGGCACGCCTTGCAGA TTACG	PglpF.прямой
O262	12	AGCTGTTUCCTCCTTGGTTAATGTTTGT TGTATGCG	PglpF.обратный
O459	13	AGCTGTTUCCTAGTTGGTTAATGTTTGT TGTATGCG	<i>PglpF</i> _SD4
O462	14	AGCTGTTUTGCTCTTGGTTAATGTTTGT TGTATGCG	<i>PglpF</i> _SD7
KABY733	15	AAACAGCUATGAAGAGCGCGCTGACCT TCAG	<i>fred</i> .прямой
KABY734	16	AGGGTTAAUTTACGCTTCACGCACACG CG	<i>fred</i> .обратный
O48	17	CCCAGCGAGACCTGACCGCAGAAC	<i>galK</i> .прямой
O49	18	CCCCAGTCCATCAGCGTGACTACC	<i>galK</i> .обратный
O68	20	ATGCGCAAUTGTGAGTTAGCTCACTC ATTAG	Plac.прямой
O113	21	AGCTGTTUCCTCCTTAGGTACCCAGCTT TTGTTCCC	Plac.обратный
KABY745	22	AAACAGCUATGAAGAGCCTGCTGACCC GTAAAC	vag.прямой
KABY746	23	AGGGTTAAUTTAAACGTTTTTCACACGC GCG	vag.обратный
KABY721	24	AAACAGCUATGAAGAGCGCGCTGACCT TTAGC	yberC0001_9420.прямой
KABY722	25	AGGGTTAAUTTACGCCTCACGCACACG CG	yberC0001_9420.обратный
KABY729	26	AAACAGCUATGAGCAGCCGTCGTCTGA GC	bad.прямой
KABY730	27	AGGGTTAAUTTACACGTTTTTAACACG GGTCATCAG	bad.обратный
KABY741	28	AAACAGCUATGCAGAGCTTCACCCCGC C	нес.прямой
KABY742	29	AGGGTTAAUTTACGCCTGCTCTTTAACA CGCAGC	нес.обратный
KABY737	30	AAACAGCUATGCAGCGTCTGAGCCGTC TGAG	marc.прямой

KABY738	31	AGGGTTAAUTTAAACTTCACGCACTTTC GCGC	marc.обратный
MP1217	32	AAACAGCUATGAAGGCGCTGTGGAGCC GTCG	yabM.прямой
MP1218	33	AGGGTTAAUCGCCAGCGGAACGCTCTT CACG	yabM.обратный

Конструирование плазмид

Были синтезированы плазмидные скелеты, содержащие два эндонуклеазных сайта I-SceI, разделенных двумя фрагментами ДНК, приспособленными для гомологичной рекомбинации в геном *E. coli*, и последовательностью терминатора транскрипции T1. Например, в одной плазмидной основе (pUC57::gal) был синтезирован оперон *gal* (необходимый для гомологичной рекомбинации в *galk*) и последовательность терминатора транскрипции T1 («GeneScript»). Последовательности ДНК, примененные для гомологичной рекомбинации в гал-опероне, охватывают пары оснований 3.628.621-3.628.720 и 3.627.572-3.627.671 в последовательности *Escherichia coli* K-12 MG155, полный геном GenBank: ID: CP014225.1. Вставка путем гомологичной рекомбинации привела бы к делеции 949 пар оснований *galK* и фенотипа *galK*-. Подобным образом могут быть синтезированы скелеты на основе pUC57 («GeneScript») или любого другого подходящего вектора, содержащего два сайта эндонуклеазы I-SceI, разделенных двумя фрагментами ДНК, приспособленными для гомологичной рекомбинации в геном *E. coli*, и последовательностью терминатора транскрипции T1. Для конструирования праймеров и амплификации специфических последовательностей ДНК хромосомной ДНК *Escherichia coli* K-12 DH1 применяли стандартные методики, хорошо известные в области молекулярной биологии. Такие стандартные приемы, векторы и элементы можно найти, например, в руководствах: Ausubel et al. (eds.), *Current Protocols in Molecular Biology* (1995) (John Wiley & Sons); Sambrook, Fritsch, & Maniatis (eds.), *Molecular Cloning* (1989) (Cold Spring Harbor Laboratory Press, NY); Berger & Kimmel, *Methods in Enzymology* 152: Guide to Molecular Cloning Techniques (1987) (Academic Press); Bukhari et al. (eds.).

Хромосомную ДНК, полученную из *E. coli* DH1, применяли для амплификации фрагмента ДНК размером 300 п.н., содержащего промотор *PglpF*, с применением олигонуклеотидов O261 и O262, *PglpF*_SD4 применяя олигонуклеотиды O261 и O459, и *PglpF*_SD7 применяя олигонуклеотиды O261 и O462 (таблица 6).

Фрагмент ДНК длиной 1.182 п.н., содержащий кодоноптоимизированную версию гена *fred*, SEQ ID NO: 2, происходящего из *Yersinia frederiksenii* синтезировали с помощью «GeneScript» (таблица 4). Ген *fred* амплифицировали с применением олигонуклеотидов KABY733 and KABY734.

Все ПЦР-фрагменты (основы плазмиды, элементы, содержащие промотор, и ген

fred) очищали и собирали основы плазмиды, элементы промотора (*Plac*, *PglpF* или *PglpF_SD4* или *PglpF_SD7*), и содержащий *fred* (или другой представляющий интерес ген, смотри таблицу 4) фрагмент ДНК. Плазмиды клонировали с помощью стандартного клонирования USER. Клонирование в любой подходящей плазмиде может быть осуществлено с помощью любых стандартных способов клонирования ДНК. Плазмиды трансформировали в клетки TOP10 и отбирали на чашках LB, содержащих 100 мкг/мл ампициллина (или любого подходящего антибиотика) и 0,2% глюкозы. Сконструированные плазмиды очищали, а последовательность промотора и 5'-конец гена *fred* подтверждали секвенированием ДНК («MWG Eurofins Genomics»). Таким образом конструировали генетическую кассету, содержащую любой представляющий интерес промотор, связанный с геном *fred* (или другим представляющим интерес геном, смотри таблицу 4).

Конструирование штаммов

Примененный бактериальный штамм, MDO, был сконструирован из *Escherichia coli* K-12 DH1. Гентип *E. coli* K-12 DH1 представлял собой: *F⁻*, Δ -, *gyrA96*, *recA1*, *relA1*, *endA1*, *thi-1*, *hsdR17*, *supE44*. В дополнение к генотипу *E. coli* K-12 DH1 MDO имеет следующие модификации: *lacZ*: делеция, размером 1,5 т.п.н., *lacA*: делеция, размером 0,5 т.п.н., *nanKETA*: делеция, размером 3,3 т.п.н., *melA*: делеция, размером 0,9 т.п.н., *wcaJ*: делеция, размером 0,5 т.п.н., *mdoH*: делеция, размером 0,5 т.п.н., и вставка промотора *Plac* выше гена *gmd*.

Вставку кассеты экспрессии, содержавшей промотор, связанный с геном *fred* Вставку кассеты экспрессии, содержавшей промотор, связанный с геном T1 в хромосомной ДНК *E.coli* K-12 DH1 MDO осуществляли с помощью Gene Gorging по существу, как описано в работе Herring et al. (Herring, C.D., Glasner, J.D. and Blattner, F.R. (2003). *Gene* (311). 153-163). Вкратце, донорную плазмиду и хелперную плазмиду котрансформировали в MDO и отбирали на чашках с LB, содержавших 0,2% глюкозы, ампициллин (100 мкг/мл) или канамицин (50 мг/мл) и хлорамфеникол (20 мкг/мл). Одиночную колонию инокулировали в 1 мл LB, содержащего хлорамфеникол (20 мкг/мл) и 10 мкл 20% L-арабинозы, и инкубировали при 37°C при встряхивании в течение 7–8 часов. Для интеграции в локусы *galk* клеток *E. coli* затем высевали на чашки M9-DOG и инкубировали при 37°C в течение 48 часов. Отдельные колонии, сформированные на чашках MM-DOG, повторно высевали штрихом на чашки с LB, содержавшие 0,2% глюкозы, и инкубировали в течение 24-х часов при 37°C. Ожидалось, что колонии, которые выглядели белыми на чашках с агаром MacConkey-галактоза и были чувствительны как к ампициллину, так и к хлорамфениколу, потеряли донорную и

вспомогательную плазмиду и содержали вставку в локусах *galK*. Отдельные колонии, сформированные на чашках MM-DOG, повторно высевали штрихом на чашки с LB, содержащие 0,2% глюкозы, и инкубировали в течение 24-х часов при 37°C. Ожидалось, что колонии, которые выглядели белыми на чашках с агаром MacConkey-галактоза и были чувствительны как к ампициллину, так и к хлорамфениколу, потеряли донорную и вспомогательную плазмиду и содержали вставку в локусах *galK*. 17) и O49 (SEQ ID NO: 18) а вставленную ДНК проверяли секвенированием («Eurofins Genomics», Германия).

Встраивание генетических кассет в другие локусы хромосомной ДНК *E. coli* осуществляли аналогичным образом с применением различных маркерных генов селекции.

Анализ в глубоких лунках

Анализ в глубоких лунках выполняли, как первоначально описано в работе Lv et al (Bioprocess Biosyst Eng (2016) 39:1737–1747) и оптимизировали для целей настоящего изобретения.

Более конкретно, штаммы, раскрытые в примерах, подвергали скринингу в 24 планшетах с глубокими лунками с применением 4-дневного протокола. В течение первых 24-х часов клетки культуры на шейкере выращивали при 34°C при встряхивании со скоростью 700 об/мин до высокой плотности, в то время как в следующие 48 часов для клеток, продуцирующих 2'-FL и 3-FL, и 72 часа для клеток, продуцирующих LNT, клетки переносили в среду, которая позволяла индуцировать экспрессию генов и образование продукта. В особенности, в течение 1 дня готовили свежие инокуляты с применением базовой минимальной среды с добавлением сульфата магния, тиамин и глюкозы. Через 24 часа инкубации клетки переносили на новую базовую минимальную среду (2 мл) с добавлением сульфата магния и тиамин с добавлением исходного болюса, состоящего из 20% раствора глюкозы (1 мкл) и 10% раствора лактозы (0,1 мл). Затем к клеткам добавляли 50% раствор сахарозы (0,04 мл) в качестве источника углерода, сопровождаемый добавлением сахарозогидролазы (инвертазы, 5 мкл раствора 0,1 г/л), так что глюкоза поступала с медленной скоростью для роста за счет расщепления сахарозы инвертазой. После инокуляции новой среды клетки встряхивали при 700 об/мин при 28°C в течение 48 часов. После денатурации и последующего центрифугирования супернатанты анализировали с помощью HPLC. Для анализа общих образцов клеточный лизат, приготовленный кипячением, осаждали центрифугированием в течение 10 минут при 4700 об/мин. Концентрацию ОГМ в супернатанте определяли с помощью HPLC или НРАС.

Пример 1. Получение 2'-FL

Разработка *Escherichia coli* для производства 2'-FL, экспрессирующих ген *fred*

Штамм *Escherichia coli* K-12 (DH1) MDO) можно изменять для экспрессии представляющих интерес гетерологичных генов. Например, штамм 1 представляет собой штамм, продуцирующий 2'-FL, со сверхэкспрессией гена α -1,2-фукозилтрансферазы, *futC*, и генов колановой кислоты (*gmd-wcaG-wcaH-wcaI-manC-manB*). Вставка экспрессионной кассеты, содержащей промоторный элемент (*PglpF*), связанный с *fred* в единственной копии, в фоновый штамм 1 привела к штамму 2. Результаты анализов в глубоких лунках показали, что экспрессия гена *fred* с применением *PglpF* i) увеличивает продукцию 2'-FL в 1,5 раза (увеличение общей продукции 2'-FL на 15%) и ii) улучшает распределение продукта 2'-FL за счет снижения количества 2'-FL в клеточной фракции и увеличения количества продукцию в среде.

Пример 2. Получение 3-FL

Конструирование *Escherichia coli*, экспрессирующих ген *fred*, для получения 3-FL

Штамм *Escherichia coli* K-12 (DH1) MDO) можно изменять для экспрессии представляющих интерес гетерологичных генов. Например, штамм представляет собой штамм-продуцент 3-FL со сверхэкспрессией гена альфа-1,3-фукозилтрансферазы, *futA* и генов колановой кислоты (*gmd-wcaG-wcaH-wcaI-manC-manB*). Вставка экспрессионной кассеты, содержащей промоторный элемент (*PglpF*) связанный с геном *fred* в единственной копии, в олов штамма 3 привела к штамму 4. Результаты анализов в глубоких лунках показали, что экспрессия гена *fred* с применением *PglpF* i) увеличивает продукцию 3-FL почти в 2 раза (увеличение общей продукции 3-FL на 90%) и ii) улучшает распределение продуктов 3-FL за счет снижения количества 3-FL в клеточной фракции и увеличения количества 3-FL в среде.

Пример 3. Получение LNT

Конструирование *Escherichia coli*, экспрессирующих ген *fred*, для получения LNT

Штаммы *Escherichia coli* K-12 (DH1) MDO) можно изменять для экспрессии представляющих интерес гетерологичных генов. Например, штамм 5 представляет собой штамм-продуцент LNT со сверхэкспрессией гена β -1,3-N-ацетилглюкозаминилтрансферазы, *lgtA*, и β -1,3- галактозилтрансферазы, *galTK*. Вставка экспрессионной кассеты, содержащей промоторный элемент (*PglpF_SD4*) связанный с геном *fred*, в единственной копии в олов штамма 5 привела к штамму 6. Вставка экспрессионной кассеты, содержащей промоторный элемент (*PglpF_SD7*), связанный с геном *fred*, в единственной копии в олов штамма 5 привела к штамму 7. Результаты анализов в глубоких лунках показали, что экспрессия гена *fred* с применением *PglpF_SD4*

или *PglpF_SD7* i) увеличение продукции LNT в 1,2-1,3 раза (увеличение общей продукции LNT на 20-30%) и ii) небольшое улучшение распределения продукта LNT за счет снижения количества LNT в клеточной фракции и увеличение количества LNT в среде.

Резюме примеров 1-3

Хромосомная экспрессия *fred* в штамме-продуcente 2'-FL, LNT или 3-FL увеличивала продукцию ОГМ, а в случае 2'-FL и 3-FL экспрессия *fred* снижала количество ОГМ внутри клеток (фракция осадка).

Следовательно, экспорт ОГМ белками суперсемейства основных посредников *fred* увеличивает способность штаммов-продуцентов продуцировать ОГМ.

Таблица 7

Исходный штамм*	ID нового штамма, экспрессирующего <i>fred</i>	ОГМ	Продукция ОГМ	Снижение ОГМ в осадках
Штамм 1	Штамм 2	2'-FL	115%	да
Штамм 3	Штамм 4	3-FL	190 %	да
Штамм 5	Штамм 6 Штамм 7	LNT	130% 120%	нет нет

*Основа штамма для продукции ОГМ

Пример 4. Транспортёр Vag представляет собой новый переносчик ядра LNT-2 с высокой селективностью в отношении LNT.

Таблица 6. Примеры хелперных и донорных плазмид, применяемых для конструирования штаммов

Плазмида	Соответствующий генотип	Маркерный ген
pACBSR	<i>Para-I-SceI-λ Red, p15A ori, cam*</i>	cam
pUC57	<i>pMB1, bla</i>	Bla
pUC57::gal	<i>pUC57::galTK' / T1-galKM'</i>	Bla

Последовательности ДНК гетерологичных генов, кодирующих представляющие интерес транспортеры или гликозилтрансферазы, были оптимизированы по кодонам и синтезированы с помощью GenScript. Интересующие гены, кодирующие белки-транспортеры, как показано в таблице 4, амплифицировали с помощью ПЦР с применением подходящих праймеров, охватывающих стартовый кодон ATG и стоп-кодон ТАА гена (таблица 3). Для конструирования донорных плазмид с любым представляющего интерес гетерологичного гена, применяли стандартное клонирование USER для объединения очищенных ПЦР-фрагментов соответствующего скелета плазмиды, промоторного элемента и представляющего интерес гена. Клонирование в соответствующей плазмиде может быть осуществлено с применением любого стандартного способа клонирования ДНК. После клонирования ДНК трансформировали в клетки TOP10 и отбирали на чашках LB, содержащих 100 мкг/мл ампициллина (или 50 мг/мл канамицина в зависимости от применяемой основы) и 0,2% глюкозы.

Сконструированные плазмиды очищали, а последовательность промотора и 5'-конец интересующего гена подтверждали секвенированием ДНК (MWG Eurofins Genomics).

Конструирование штаммов

Применяемый бактериальный штамм, MDO, был сконструирован из *Escherichia coli* K-12 DH1. Генотип *E. coli* K-12 DH1 представляет собой: *F⁻, λ , gyrA96, recA1, relA1, endA1, thi-1, hsdR17, supE44*. В дополнение к генотипу *E. coli* K-12 DH1 MDO имеет следующие модификации: *lacZ*: делеция, размером 1,5 т.п.н., *lacA*: делеция, размером 0,5 т.п.н., *nanKETA*: делеция, размером 3,3 т.п.н., *melA*: делеция, размером 0,9 т.п.н., *wcaJ*: делеция, размером 0,5 т.п.н., *mdoH*: делеция, размером 0,5 т.п.н., и вставка промотора *Plac* выше гена *gmd*.

Вставку кассеты экспрессии, содержащей промотор, связанный с геном Fred и последовательностью терминатора транскрипции T1, осуществляли с помощью Gene Gorging, по существу, как описано Herring et al (Herring, C.D., Glasner, J.D. and Blattner, F.R. (2003). 311). 153-163), а примеры хелперной и донорной плазмид, примененных для конструирования штаммов, представленных в настоящей заявке, приведены в таблице 6. Вкратце, донорную плазмиду и хелперную плазмиду котрансформировали в MDO и отбирали на чашках с LB, содержащих 0,2% глюкозы, ампициллин (100 мкг/мл) или канамицин (50 мг/мл) и хлорамфеникол (20 мкг/мл). Одиночную колонию инокулировали в 1 мл LB, содержащего хлорамфеникол (20 мкг/мл) и 10 мкл 20% L-арабинозы, и инкубировали при 37°C при встряхивании в течение 7–8 часов. Для интеграции в локусы *galk* клеток *E. coli* затем высевали на чашки M9-DOG и инкубировали при 37°C в течение 48 часов. Отдельные колонии, сформированные на чашках MM-DOG, повторно высевали штрихом на чашки с LB, содержащих 0,2% глюкозы, и инкубировали в течение 24-х часов при 37°C. Колонии, которые выглядели белыми на чашках с агаром MacConkey-галактоза и были чувствительны как к ампициллину, так и к хлорамфениколу, потеряли донорную и вспомогательную плазмиду и содержали вставку в локусах *galk*. Вставки в сайт *galk* идентифицировали с помощью ПЦР колоний с применением праймеров O48 (SEQ ID NO: 23) и O49 (SEQ ID NO: 24), а вставленную ДНК проверяли секвенированием («Eurofins Genomics», Германия).

Встраивание генетических кассет в другие локусы хромосомной ДНК *E. coli* осуществляли аналогичным образом с применением различных маркерных генов селекции.

Выравнивание последовательности

Эвристическое парное выравнивание последовательностей, реализованное в BLAST 2.1.2 (Basic Local Alignment Search Tool) (Altschul et al. 1990) в базе данных NCBI

(<http://www.ncbi.nlm.nih.gov>), выполняли для проверки гомологии последовательностей белков-транспортеров, упомянутых в настоящем изобретении. Все параметры сохраняли со значениями по умолчанию для каждого выравнивания BLAST.

В настоящем изобретении мы проверили способность выбранных гетерологичных транспортеров MFS экспортировать L_{NnT} из клетки-хозяина с применением эталонного штамма, а именно MP3672. Штамм имеет одну *P_{glpF}*-управляемую копию каждой гетерологичной гликозилтрансферазы, которая необходима для образования продукта.

Одна копия каждого из выбранных гетерологичных генов-транспортеров, т.е. *vag*, *yberC0001_9420*, *marc*, *bad*, *nec*, *yabM* и *fred* (таблица 4), индивидуально интегрировали в геном штамма MP3672 под контролем промотора *Plac*, который, как известно, обеспечивает умеренные уровни транскриптов. Сайт связывания рибосом промотора *Plac* (т.е. 16 п.н. выше сайта начала трансляции) был модифицирован (смотри таблица 5) для усиления рибосомного связывания с синтезированными транскриптами и, таким образом, позитивно регулирует управляемую *Plac* экспрессию на уровне трансляции.

Следуя описанному выше подходу к конструированию штамма и тестируя эффективность штамма в DWA, мы сообщаем в настоящем документе, что только Vag-экспрессирующие клетки, продуцирующие L_{NnT}, демонстрируют относительное увеличение как оттока продуцируемого L_{NnT}, так и общей продукции L_{NnT} по сравнению с эталонным штаммом.

Как показано на фиг. 7, *Plac*-управляемая экспрессия большинства генов-транспортеров (*marc*, *fred*, *bad*, *nec*, *yabM*, *yberC0001_9420*) либо не оказывает существенного влияния, либо снижает продукцию L_{NnT} (до 45%). Напротив, введение гена *vag* в генетический фон штамма MP3672 приводит к заметному увеличению титра L_{NnT} (фиг. 7).

Выравнивание аминокислотных последовательностей семи протестированных предполагаемых переносчиков MSF показало, что переносчик Vag имеет очень высокий охват последовательностей (от 99 до 100%) по сравнению с остальными шестью транспортерами, протестированными в настоящем изобретении, с идентичностью последовательностей от 65% до 75%. Самая высокая идентичность последовательности (75%) была оценена для последовательности Vag и последовательности транспортера MFS, кодируемого геном *yabM* (таблица 8). Интересно, что транспортер MFS YabM, описанный в WO 2017042382 в качестве предполагаемого эффективного экспортера LNT, не оказывал существенного влияния ни на окончательный общий титр L_{NnT}, ни на отток L_{NnT} (фиг. 7, 8). Таким образом, вопреки относительно высокому сходству его белковой последовательности с Vag, транспортер *YabM*, по-видимому, неэффективен в отношении

транспорта L_{NnT}, что ясно отражается в концентрациях продукта, обнаруженных во внеклеточной фракции соответствующих культур клеток-хозяев (фиг. 8). В особенности, как показал анализ супернатантов фракций бактериальных культур, внеклеточные концентрации L_{NnT} были намного выше для штамма MP3994 (клетки, экспрессирующие *vag*), чем для штаммов MP4362 (клетки, экспрессирующие *yabM*) и MP3672 (клетки сравнения, не экспрессирующие гетерологичный транспортер) (фиг. 8).

В совокупности эти данные свидетельствуют о том, что транспортер Vag представляет собой эффективный переносчик L_{NnT}. Этот вывод основан как на результатах процедуры скрининга семи генов-транспортеров, геномно экспрессированных в одном и том же эталонном штамме, в тех же условиях культивирования и под контролем одного и того же промотора, *Plac*, так и на анализе последовательностей белков-транспортеров, который показал, что белки-транспортеры с относительно высокой гомологией к Vag, такие как YabM и Marc (75% и 71%), вообще не экспортируют L_{NnT} или если экспортируют, то с очень низкой эффективностью.

Таблица 8. Гомология между различными гетерологичными транспортерами по настоящему изобретению:

Название белка	Идентификация	Из организма	Идентичен VagWP_048785139.1 (покрытие)
YberC0001_9420	EEQ08298.1 Суперсемейство основных переносчиков MFS 1	<i>Yersinia bercovieri</i>	68% (99%)
Fred	WP_087817556.1 транспортер MFS	<i>Yersinia frederiksenii</i>	69% (99%)
Bad	WP_017489914.1 транспортер MFS	<i>Rouxiella badensis</i>	65% (99%)
Nec	WP_092672081.1 транспортер MFS	<i>Rosenbergiella nectarea</i>	66% (99%)
Marc	WP_060448169.1 транспортер MFS	<i>Serratia marcescens</i>	71% (99%)
YabM	CAY73138.1 Предполагаемый переносчик оттока сахара MFS	<i>Erwinia pyrifoliae</i>	75% (99%)

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Генетически модифицированная клетка, способная продуцировать один или несколько олигосахаридов грудного молока (ОГМ), в которой указанная генетически модифицированная клетка включает последовательность гетерологичной нуклеиновой кислоты, кодирующую полипептид суперсемейства основных переносчиков (MFS), показанный в SEQ ID NO: 1, или его функциональный гомолог, аминокислотная последовательность которого более чем на 95,4% идентична SEQ ID NO: 1.

2. Генетически модифицированная клетка по п. 1, в котором указанный функциональный гомолог, по меньшей мере, на 96% идентичен SEQ ID NO: 1.

3. Генетически модифицированная клетка по п. 1, в котором клетка дополнительно включает последовательность нуклеиновой кислоты, включающую регуляторный элемент для регуляции экспрессии гетерологичной последовательности нуклеиновой кислоты.

4. Генетически модифицированная клетка по п. 3, в котором регуляторный элемент регулирует экспрессию полипептида MFS, показанного в SEQ ID NO: 1, или его функционального гомолога, аминокислотная последовательность которого более чем на 95,4 % идентична SEQ ID NO: 1.

5. Генетически модифицированная клетка по п. 3 или 4, в котором регуляторный элемент выбран из группы, состоящей из PglpF, PglpF_SD4 и PglpF_SD7.

6. Генетически модифицированная клетка по любому из предшествующих пунктов, в которых генетически модифицированная клетка представляет собой *Escherichia coli*.

7. Генетически модифицированная клетка по любому из предшествующих пунктов, в которых олигосахарид грудного молока выбирают из группы, состоящей из 2'-фукозиллактозы (2'-FL), 3-фукозиллактозы (3-FL), дифукозиллактозы (DFL) и лакто-N-тетраоза (LNT).

8. Конструкция нуклеиновой кислоты, включающая последовательность нуклеиновой кислоты, кодирующую полипептид MFS в соответствии с SEQ ID NO: 1, или его функциональный гомолог, имеющий более чем 95,4% идентичности последовательности с SEQ ID NO: 1, в котором последовательность нуклеиновой кислоты, кодирующая полипептид MFS, имеет, по меньшей мере, 70% идентичности последовательности с SEQ ID NO: 2.

9. Конструкция нуклеиновой кислоты по п. 8, в котором конструкция дополнительно включает последовательность нуклеиновой кислоты, включающую регуляторный элемент.

10. Конструкция нуклеиновой кислоты по п. 9, в котором регуляторный элемент регулирует экспрессию последовательности нуклеиновой кислоты, имеющей, по меньшей

мере, 70% идентичности последовательности с SEQ ID NO: 2.

11. Конструкция нуклеиновой кислоты по любому из пп. 9-10, в которых регуляторный элемент выбирают из группы, состоящей из PglpF, PglpF_SD4 и PglpF_SD7.

12. Способ биосинтетического производства одного или нескольких олигосахаридов грудного молока (ОГМ), включающий стадии:

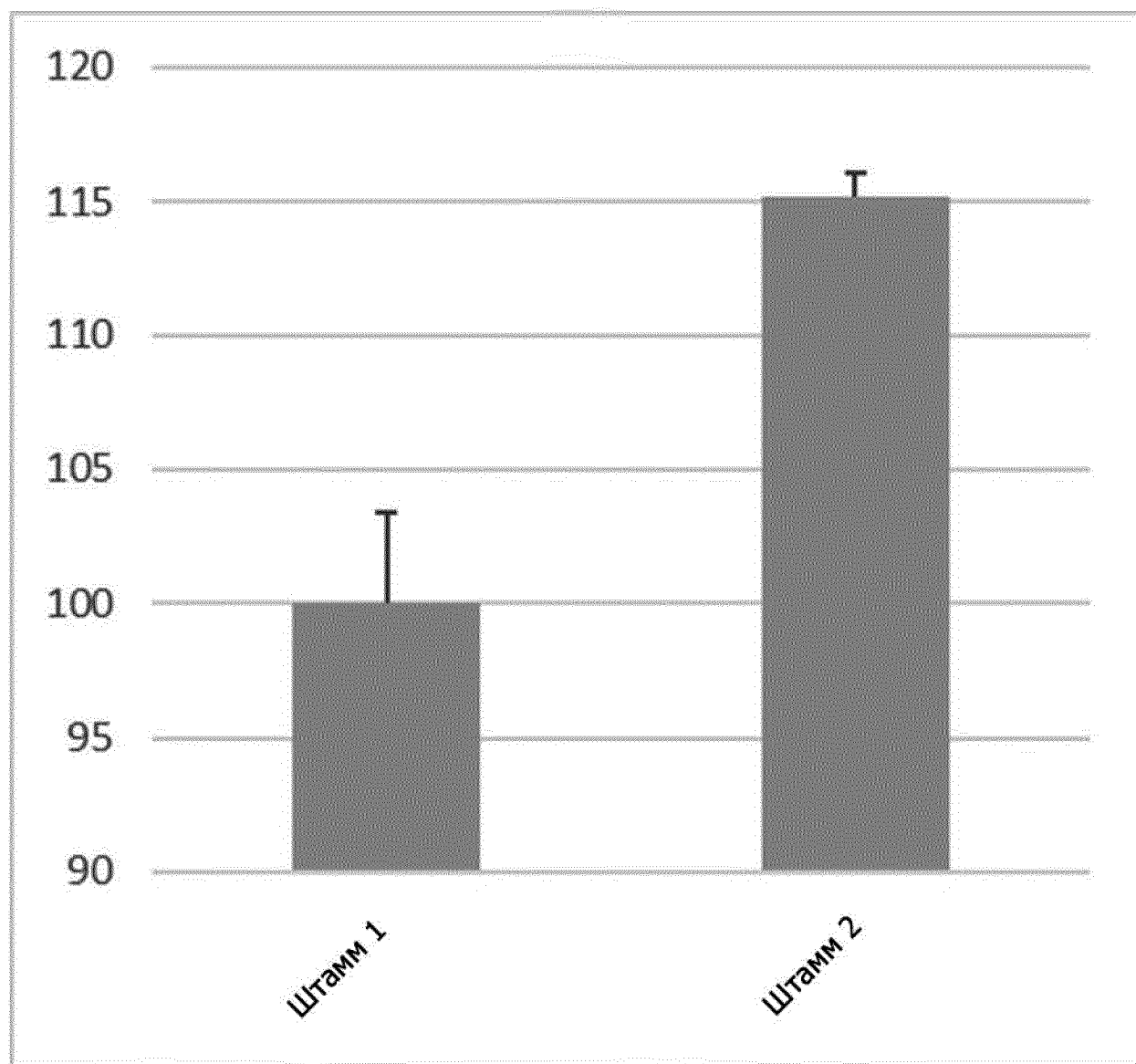
(i) обеспечение генетически модифицированной клетки по любому из пп.1-7;

(ii) культивирование генетически модифицированной клетки в соответствии с (i) в подходящей среде для культивирования клеток для экспрессии указанного полипептида, способного к транспортировке выходящего сахара и продуцированию одного или нескольких олигосахаридов грудного молока (ОГМ), и;

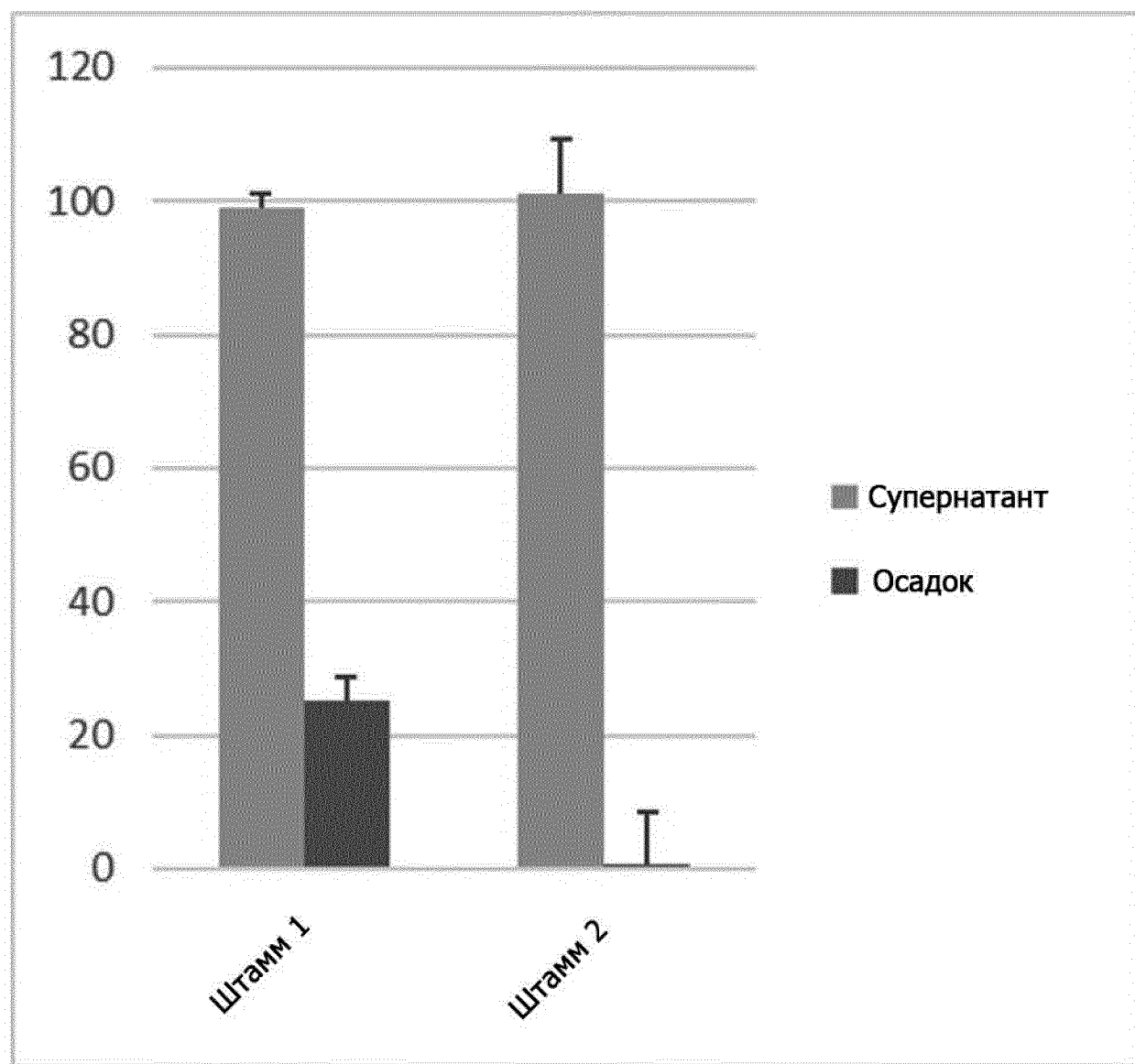
(iii) сбор одного или нескольких ОГМ, полученных на стадии (ii).

13. Применение генетически модифицированной клетки по любому из пп. 1-7 или конструкции нуклеиновой кислоты по любому из пп. 8-11 для биосинтетического производства одного или нескольких олигосахаридов грудного молока (ОГМ).

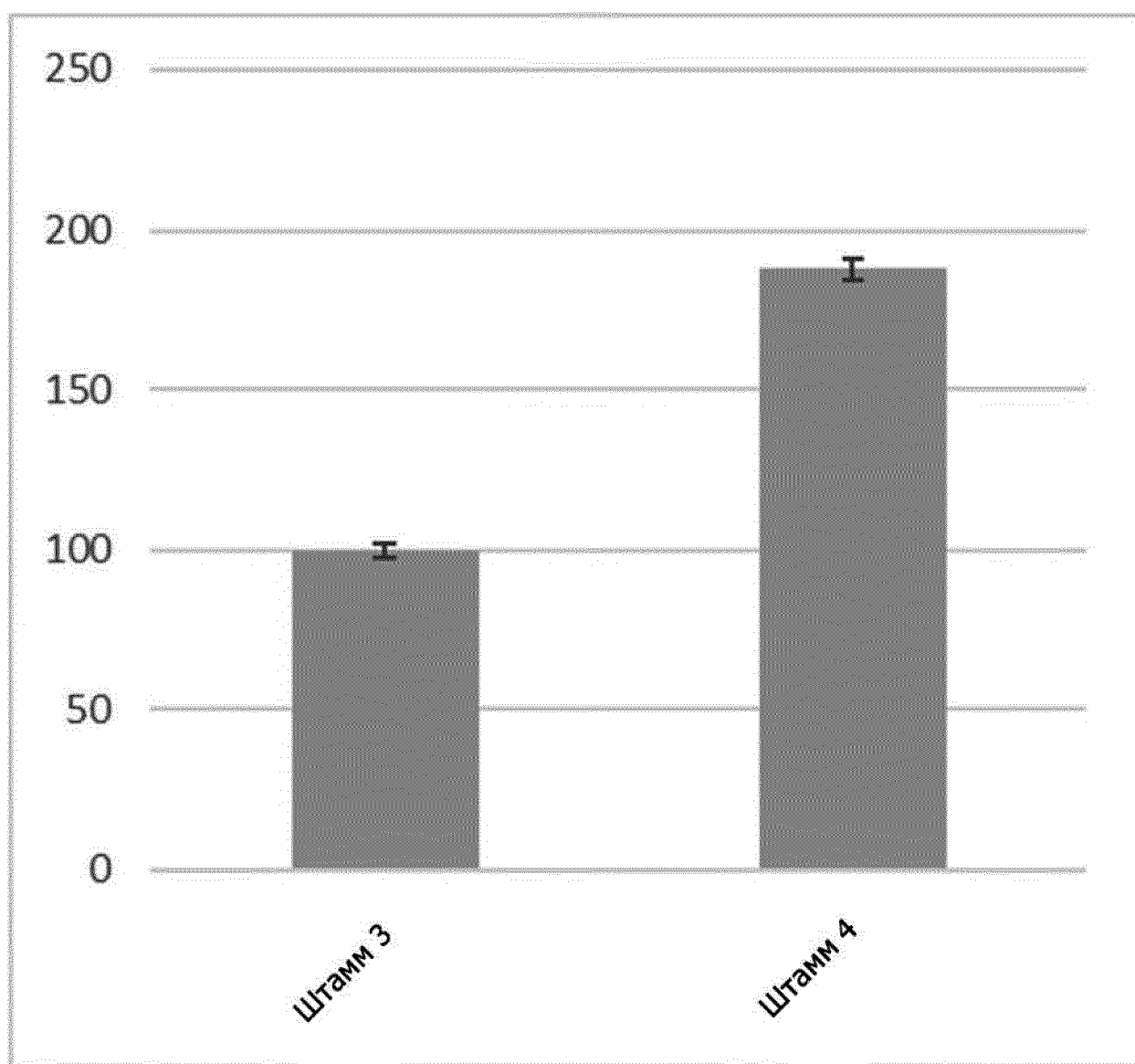
14. Применение генетически модифицированной клетки или конструкции нуклеиновой кислоты по п. 13 для биосинтетического производства конкретного ОГМ, выбранного из группы, состоящей из 2'-FL, 3-FL, DFL и LNT.



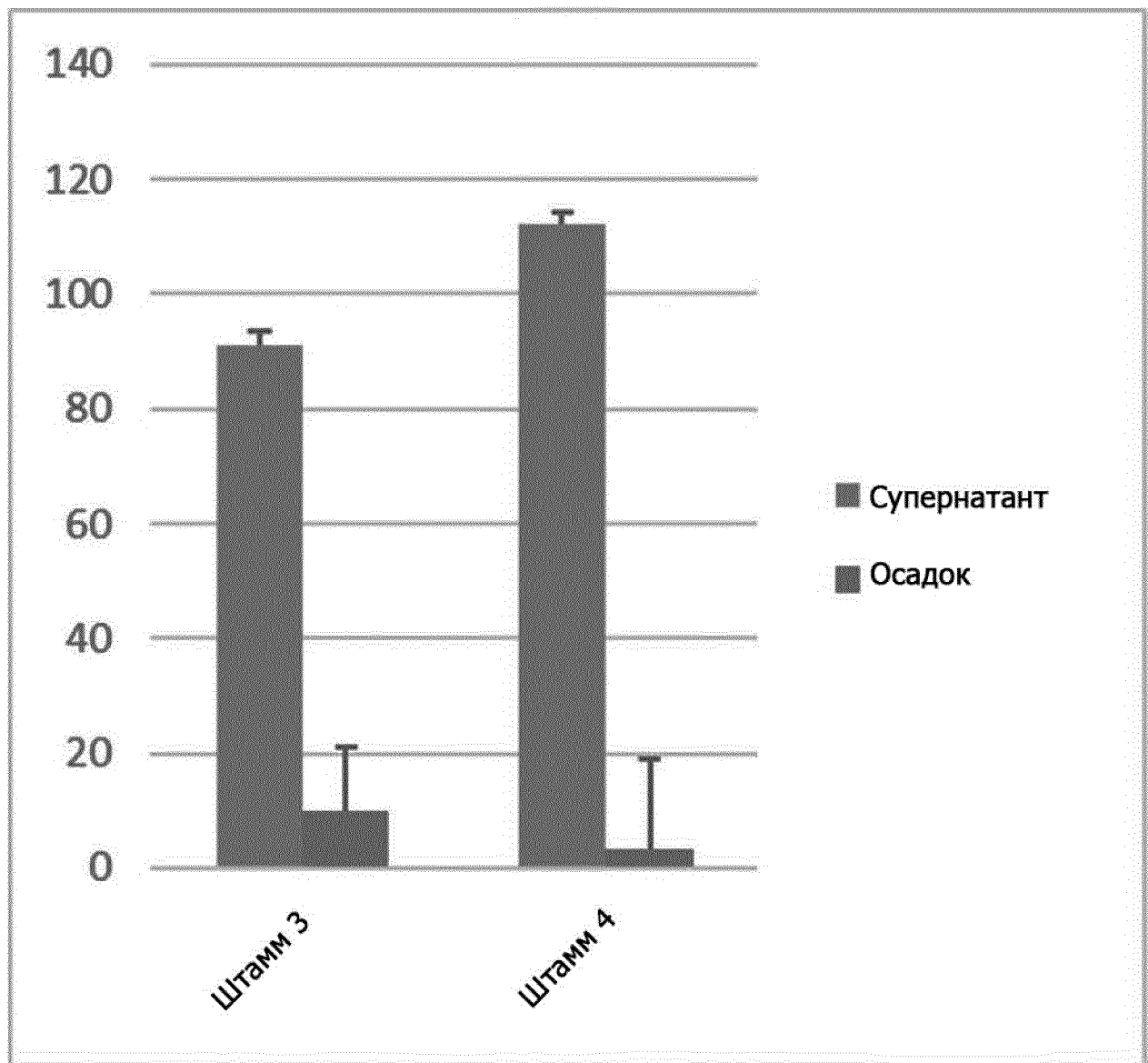
ФИГ. 1



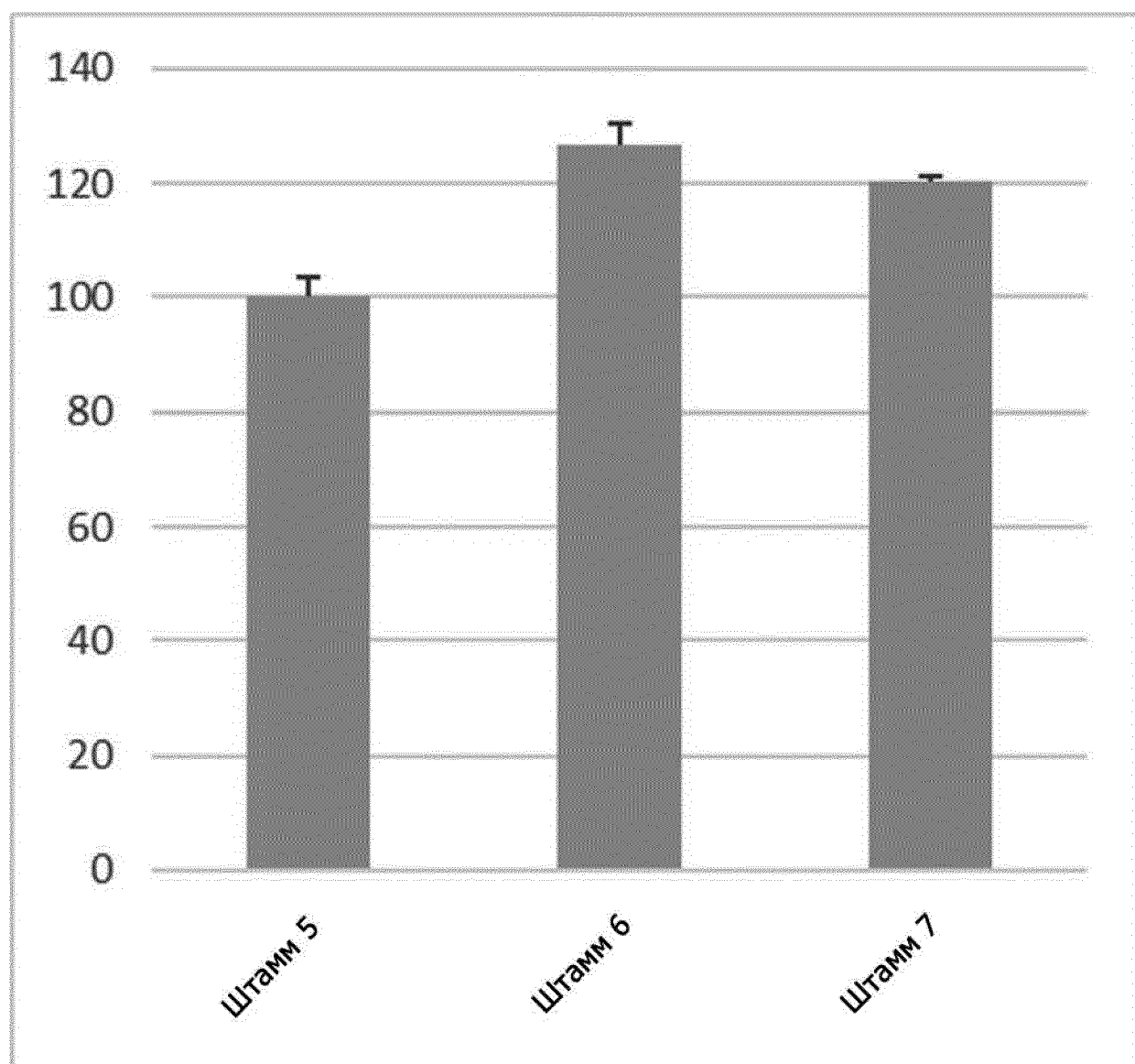
Фиг. 2



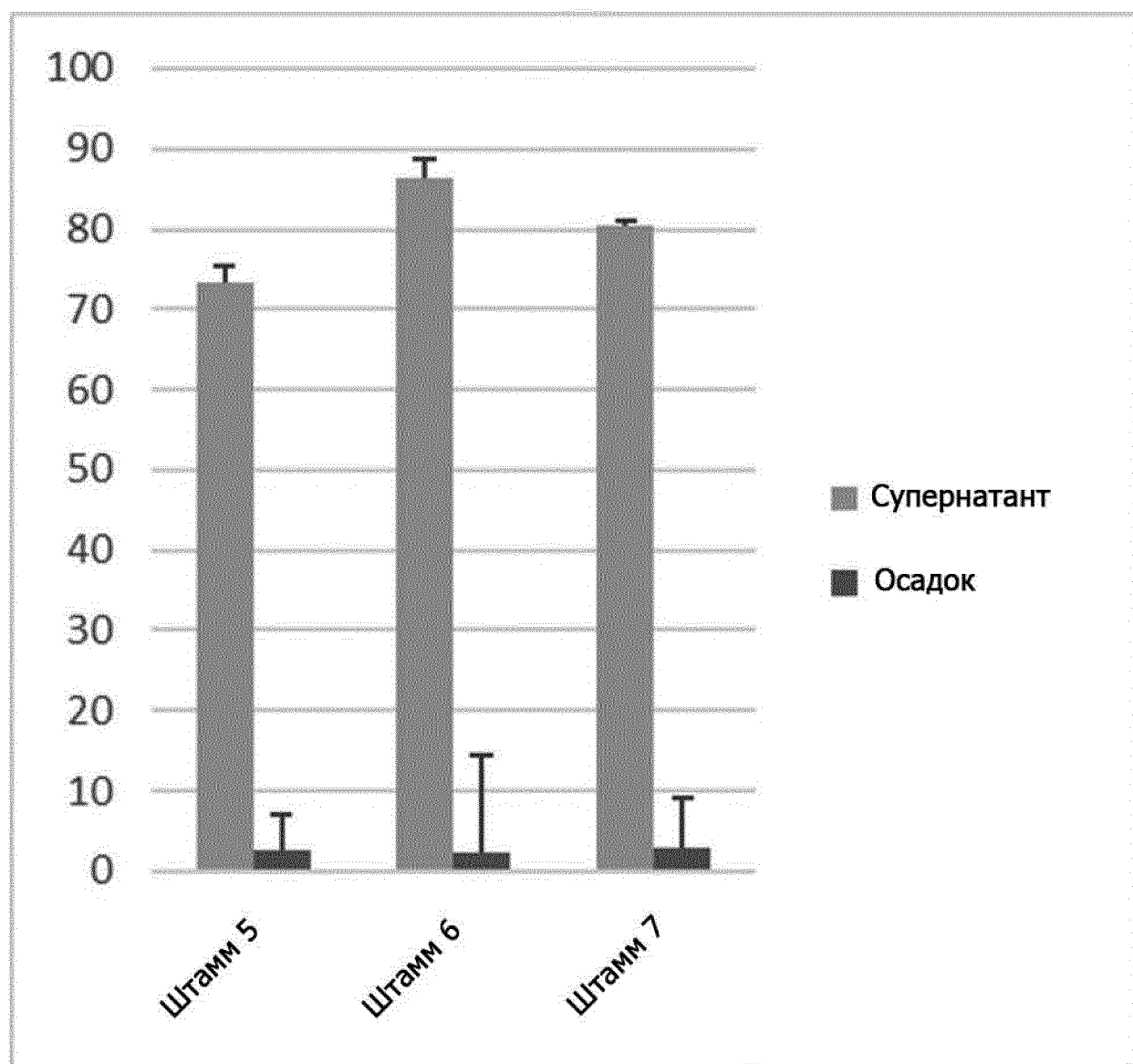
Фиг. 3



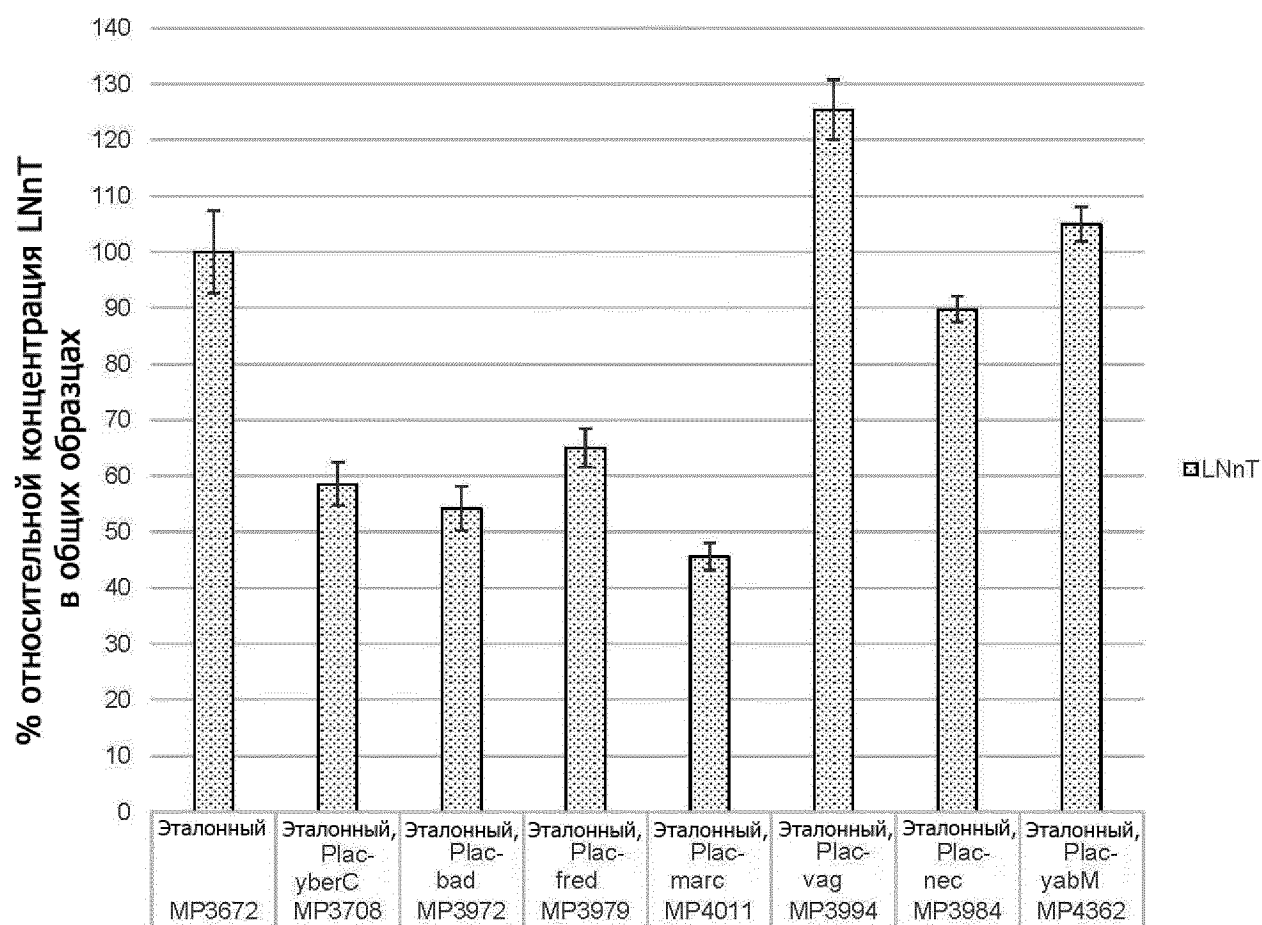
Фиг. 4



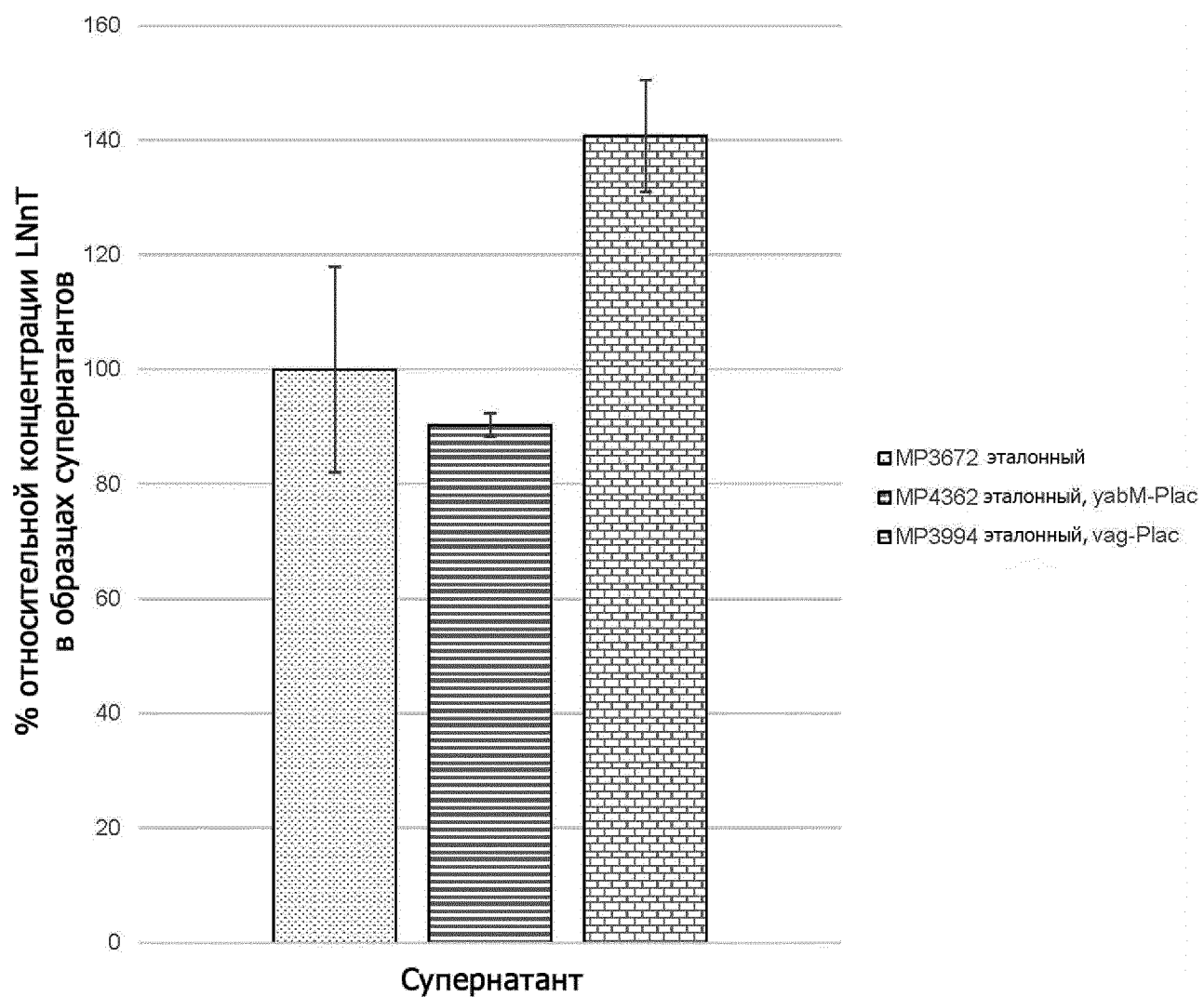
Фиг. 5



Фиг. 6



Фиг. 7



Фиг. 8