

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2019.11.29

(51) Int. Cl. *C12N 5/0735* (2010.01)
C12N 15/86 (2006.01)

(22) Дата подачи заявки
2017.11.24

(54) РЕГУЛИРУЕМАЯ ТРАНСКРИПЦИЯ

(31) 1619876.4

(32) 2016.11.24

(33) GB

(86) PCT/GB2017/053531

(87) WO 2018/096343 2018.05.31

(71) Заявитель:
**КЕМБРИДЖ ЭНТЕРПРАЙЗ
ЛИМИТЕД (GB)**

(72) Изобретатель:
Валльер Людовик, Коттер Марк,
Павловски Маттиас, Бертеро
Алессандро, Ортманн Дэниел (GB)

(74) Представитель:
Медведев В.Н. (RU)

(57) Настоящее изобретение относится к стабильному способу введения по меньшей мере одного индуцибельного кластера в клетку и достижения регулируемой транскрипции из этого индуцибельного кластера. Этот способ может быть применен для клеток любого типа, происходящих от любого эукариотического организма, а в частности он применяется для введения индуцибельных кластеров в плюрипотентные стволовые клетки, такие как плюрипотентные стволовые клетки животного или человека (hPSC). Индуцибельный кластер встраивают регулируемым образом для гарантии того, чтобы генетический материал, содержащий этот кластер, не подвергался сайленсингу или негативному воздействию в инсерционном сайте и при этом, чтобы достигалась регулируемая транскрипция генетического материала.



РЕГУЛИРУЕМАЯ ТРАНСКРИПЦИЯ

Настоящее изобретение относится к стабильному способу введения по меньшей мере одного индуцибельного кластера в клетку и достижения регулируемой транскрипции из этого индуцибельного кластера. Этот способ может быть применен для клеток любого типа, происходящих от любого эукариотического организма, а, в частности, он применяется для введения индуцибельных кластеров в плюрипотентные стволовые клетки, такие как плюрипотентные стволовые клетки животного или человека (hPSC). Индуцибельный кластер встраивают регулируемым образом для гарантии того, что генетический материал, содержащий этот кластер, не подвергался сайленсингу или негативному воздействию в инсерционном сайте и, при этом, достигалась регулируемая транскрипция генетического материала.

ПРЕДПОСЫЛКИ СОЗДАНИЯ ИЗОБРЕТЕНИЯ

Исследования стволовых клеток открывает большие перспективы для изучения эволюции человека, для регенеративной медицины, моделирования заболевания, разработки лекарственных средств и трансплантации клеток. Кроме того, клетки, происходящие от стволовых клеток, позволяют проводить исследования физиологических и патологических ответов человеческих клеточных популяций, которые не являются легко доступными. Это часто приводит к необходимости исследования генов (и других форм регуляторных механизмов, кодируемых РНК-нкРНК, которые не кодируют белок). К сожалению, регулируемая транскрипция или экспрессия генетической информации в человеческих клетках оказалась особенно проблематичной.

Кроме того, что касается нескольких ключевых аспектов регенеративной медицины, моделирования заболеваний, разработки лекарственных средств и трансплантации клеток, то они требуют проведения модификации и получения зрелых человеческих клеток различных типов из легко доступных источников. Регуляция экспрессии транскриптов в человеческих клетках является основой биологических исследований. Однако, в отношении человеческих

клеток, это оказалось трудной задачей. Кроме того, существует реальная необходимость в получении *in vitro* многих весьма желательных типов человеческих клеток, количество и качество которых было бы подходящим для разработки лекарственных средств и регенеративной медицины. Поскольку направленная дифференцировка стволовых клеток в клетки нужных типов часто является сложной задачей, то появились другие подходы, включая прямое перепрограммирование клеток в клетки нужных типов. В частности, прямое перепрограммирование, как метод прямого превращения плюрипотентных стволовых клеток, включая hPSC, в зрелые клетки различных типов, было признано как мощная стратегия получения человеческих клеток. Это перепрограммирование включает индуцированную экспрессию ключевых факторов транскрипции в линии дифференцировки (или не-кодирующих РНК, включая и-нкРНК и микроРНК) для того, чтобы превратить стволовые клетки в зрелые клетки конкретного типа. Кроме того, в этом контексте, регулируемая экспрессия генетической информации в клетках человека является сложной задачей. Доступные в настоящее время протоколы прямого перепрограммирования в значительной степени основаны на трансдукции клеток лентивирусами, что приводит к изменению экспрессии или полному сайленсингу случайно встроенных индуцибельных кластеров. В результате, для выделения субпопуляции, экспрессирующей необходимые факторы транскрипции, были проведены дополнительные стадии очистки. Таким образом, совершенно необходимы дальнейшие усовершенствования этих методов.

Помимо индуцируемой экспрессии транскрипционных факторов было бы очень желательно иметь возможность регулировать нокдаун и нокаут генов или других кодирующих последовательностей в клетках в целях исследования потери их функций. Исследования потери функций стволовых клеток и зрелых клеток дают уникальную возможность изучить механизмы, которые регулируют эволюцию человека, его заболевания и физиологию. Однако, существующие методы не позволяют легко и эффективно манипулировать экспрессией генов. Существующие методы введения материала, такого как индуцибельные короткие шпилечные РНК (кшРНК), в стволовые клетки для запуска

нокдауна гена, имеют множество недостатков, наблюдаемых при обсуждаемом выше прямом перепрограммировании, таком как, сайленсинг трансгена и эффекты положений, ограничивающие активность. Таким образом, существует необходимость в индуцировании нокаута и нокдауна гена в стволовых клетках, что позволило бы проводить исследования потери функций стволовых клеток.

Любые усовершенствования вышеуказанных методов должны гарантировать, что будет достигаться стабильная транскрипция генетического материала, содержащегося в индуцибельном кластере, например в качестве трансгена, и что такая транскрипция будет резистентной к сайленсингу и к другим негативным эффектам, ассоциированным с сайтом интеграции. Сайленсинг может быть вызван множеством эпигенетических механизмов, включая метилирование ДНК или модификации гистонов. С применением известных методов, основанных на трансдукции лентивирусов, получают клетки, представляющие собой гетерогенную популяцию с трансгенами, которые экспрессируются полностью, частично или вообще не экспрессируются. Очевидно, что во многих случаях это нежелательно. Вирусные векторы имеют тенденцию интегрировать свой генетический материал в транскрипционно активные области генома, что приводит к повышению потенциала онкогенных событий из-за инсерционного мутагенеза.

Для многих применений желательно регулировать транскрипцию встроенного генетического материала в клетке так, чтобы индуцибельный кластер мог включаться при необходимости и транскрибироваться на конкретных уровнях, включая высокие уровни. Это не может быть достигнуто, если встраивание индуцибельного кластера в геном происходит случайно.

Таким образом, авторами настоящего изобретения был разработан способ стабильного введения индуцибельного кластера в геном клетки и регуляции транскрипции этого индуцибельного кластера. Это имеет преимущества для клеток любого типа, в которые желательно вводить индуцибельный кластер и регулировать транскрипцию встроенного генетического материала, а в частности, в плюрипотентные стволовые клетки. Индуцибельный кластер может

включать любой генетический материал, способный транскрибироваться, например, трансген или некодирующую РНК (нкРНК). Материал, включенный в индуцибельный кластер, будет определяться эффектами, которые требуются от стволовых клеток, включая экспрессию трансгена или нокдаун или нокаут гена.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

Авторами настоящего изобретение было обнаружено, что можно встроить индуцибельный кластер и регулировать транскрипцию генетического материала в этом индуцибельном кластере с использованием системы, нацеленной на два геномных безопасных якорных сайта, описанных в настоящей заявке. Такой способ является весьма желательным, поскольку он позволяет уменьшить риск эпигенетического сайленсинга встроенного генетического материала и получить гомогенную популяцию клеток, транскрибирующих индуцибельный кластер.

Таким образом, настоящее изобретение относится к способу регуляции транскрипции генной последовательности в клетке, где указанный способ включает следующие стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

б) нацеленного встраивания индуцибельного кластера во второй генетически безопасный якорный сайт; где указанный индуцибельный кластер содержит указанную генную последовательность, функционально присоединенную к индуцибельному промотору, и где указанный промотор регулируется белком-регулятором транскрипции;

где указанные первый и второй генетически безопасные якорные сайты являются различными.

Специфическая интеграция индуцибельного кластера в геномные безопасные якорные сайты (GSH) является предпочтительной по сравнению со случайным встраиванием в геном. GSH были определены ранее как «внутригенные или межгенные области человеческого генома, которые способны обеспечивать предсказуемую экспрессию новой интегрированной ДНК без какого-либо негативного воздействия на клетку или организм хозяина. Подходящий

безопасный якорный сайт должен обеспечивать транскрипцию встроенной генной последовательности, достаточную для достижения желаемых уровней белка (посредством дополнительной трансляции) или некодирующей РНК. GSH также не должен сообщать клеткам предрасположенность к опухолевой трансформации, а также изменять клеточные функции» (Sadelain et al., 2012, Nature Reviews Cancer, 12(1), 51-8. doi:10.1038/nrc3179).

Первый генетически безопасный якорный сайт используется для введения гена, кодирующего по меньшей мере белок-регулятор транскрипции. Белок-регулятор транскрипции (или фактор транскрипции) повышает уровень транскрипции гена. Большинство регуляторов транскрипции представляют собой ДНК-связывающие белки, которые связываются с элементами, расположенными поблизости от энхансеров или промоторов и функционально присоединенными к гену.

В некоторых аспектах, белок-регулятор транскрипции конститутивно и перманентно экспрессируется в клетке. Таким образом, белок-регулятор транскрипции может быть функционально присоединен к конститутивному промотору. Конститутивные промоторы регулируют экспрессию гена равномерно в большинстве тканей и клеток на всех стадиях роста и развития. Конститутивные промоторы, если они используются в способах согласно изобретению, сообщают высокие уровни экспрессии гена.

Дополнительный генетический материал, включая гены, может быть встроен в первый GSH с белком-регулятором транскрипции. Такие гены могут включать один или более маркеров, таких как белок, флуоресцирующий в зеленом диапазоне спектра (GFP), который может быть использован, например, что подтверждения успешного встраивания белка-регулятора транскрипции. Другими вариантами являются гены, которые обеспечивают редактирование гена, например Cas9 и его производные или CasL и его производные, и репортерные последовательности, которые могут быть использованы для анализа эндогенной или экзогенной экспрессии специфических генов в клетке.

Второй GSH используется для введения индуцибельного кластера, в котором желаемая генная последовательность

функционально присоединена к индуцибельному промотору. Такой промотор обеспечивает транскрипцию, только если он правильно индуцируется белком-регулятором транскрипции. Белок-регулятор транскрипции может регулироваться веществом, которое является экзогенным для клетки. Таким образом, присутствие экзогенного вещества может разрешать или блокировать экспрессию от индуцибельного промотора. Примером такой регулируемой экспрессии является система Tet-ON, которая описана ниже в настоящей заявке.

Кроме того, индуцибельный(е) кластер(ы) может (могут) быть встроен(ы) в дополнительные GSH, где указанные GSH отличаются от первого и второго GSH, упомянутых выше.

Одна или более генетических последовательностей могут регулируемо транскрибироваться во втором и/или другом GSH. Действительно, индуцибельный кластер может содержать 1, 2, 3, 4, 5, 6, 7, 8, 9 или 10 генных последовательностей, которые являются желательными для встраивания в GSH, и транскрипция которых должна быть регулируемо индуцирована.

Генная последовательность или генные последовательности, которые желательно встраивать в GSH, присутствуют в индуцибельном кластере, функционально присоединенном к индуцибельному промотору. Эти генные последовательности могут представлять собой любую подходящую последовательность, способную транскрибироваться в РНК после индуцирования активности промотора. Подходящими генными последовательностями являются, но не ограничиваются ими, транскены (белок-кодирующие гены, в которых продуцируемой РНК являются матричная РНК (мРНК) транслируемая в полипептид), некодирующая РНК (нкРНК, включая, но не ограничиваясь ими, кшРНК, антисмысловая РНК (асРНК), руководящая РНК (рРНК), микроРНК (миРНК), короткая интерферирующая РНК (киРНК), транс-действующая РНК (тдРНК), антагомиры, аптамеры, миРНК-губки и любая другая функциональная РНК).

Индукцибельные кластеры могут включать дополнительный генетический материал, встраиваемый во второй или последующий GSH. Такой дополнительный генетический материал может включать

один или более маркеров, таких как белок, флуоресцирующий в зеленом диапазоне спектра (GFP), для подтверждения наличия транскрипции. Альтернативно или дополнительно, гены, такие как гены резистентности к антибиотикам или лекарственным средствам, позволяют осуществлять отбор правильно встроенных индуцибельных кластеров. Кроме того, может оказаться желательной индуцибельная экспрессия конкретного гена для исследования его функции или последовательностей, которые негативно влияют на его функцию. Аналогичным образом, может оказаться желательной экспрессия генов для усиления или подавления биологических функций клетки или влияния клеток в другой части организма, включая экспрессию факторов роста, пептидных гормонов, в том числе инсулина и т.д.

С технической точки зрения, встраивание в первый и/или во второй GSH может быть осуществлено на одной хромосоме или на обеих хромосомах. GSH присутствует в одних и тех же генетических локусах на обеих хромосомах диплоидных организмов. Инсерция в обе хромосомы является предпочтительной, поскольку это может повысить уровень транскрипции из введенного генетического материала в индуцибельном кластере и достичь особенно высоких уровней транскрипции.

Инсерции в GSH могут представлять собой регулируемую специфическую инсерцию генетического материала в конкретный GSH после того, как будет достигнута специальная сайт-специфическая генерация двухцепочечных разрывов ДНК (DSB) в GSH. Затем, генетический материал может быть введен по любому подходящему механизму, такому как гомологичная рекомбинация. При этом может быть применен любой способ получения специфических DSB в геноме, но предпочтительными системами являются CRISPR/Cas9 и модифицированные варианты, ZFN и система TALEN.

Кроме того, введение регулятора транскрипции и/или индуцибельного кластера может быть осуществлено так, чтобы оно было обратимым и чтобы встроенный генетический материал мог быть удален и/или заменен, если это необходимо, альтернативным регулятором транскрипции/индуцибельным кластером. Способы замены регулятора транскрипции и/или индуцибельного кластера являются частью изобретения. Такая замена может быть полезной, если

клеточные культуры были успешно модифицированы одним регулятором транскрипции и/или одним индуцибельным кластером, при этом может оказаться желательной замена регулятора транскрипции и/или индуцибельного кластера. Это имеет преимущество в случае уже успешной проведенной инсерции и позволяет вводить более крупные вставки. Для осуществления этого аспекта изобретения, во вставки могут быть включены отщепляемые последовательности для удаления всей или части вставки из GSH, например, части вставки. Предпочтительные способы удаления или замены включают рекомбинантные способы.

Кроме того, настоящее изобретение относится к векторам, подходящим для введения регулятора транскрипции и/или индуцибельного кластера в GSH.

В одном из своих аспектов, настоящее изобретение относится к способу регуляции экспрессии трансгена в клетке, где указанный способ включает следующие стадии:

a) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

b) нацеленного встраивания трансгена, функционально присоединенного к индуцибельному промотору, во второй генетически безопасный якорный сайт; где указанный индуцибельный промотор регулируется белком-регулятором транскрипции;

где указанные первый и второй генетически безопасные якорные сайты являются различными.

В этом аспекте изобретения, описанный ранее индуцибельный кластер содержит трансген, функционально присоединенный к индуцибельному промотору. В этом аспекте изобретения, нужной генной последовательностью, включенной в индуцибельный кластер, является трансген, а предпочтительно ген, кодирующий белок. Таким образом, транскрипция и трансляция (экспрессия) трансгена может регулироваться в клетке. Преимуществом настоящего способа является то, что он позволяет осуществлять сверхэкспрессию трансгена, если это необходимо.

Кроме того, в этом аспекте изобретения, в другой GSH может быть введен дополнительный, идентичный или отличающийся

трансен, где указанный GSH отличается от первого и второго GSH. Такой трансен функционально присоединяют к индуцибельному промотору, как описано выше.

В одном из своих аспектов, настоящее изобретение относится к способу регуляции транскрипции некодирующей РНК в клетке, где указанный способ включает следующие стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

б) нацеленного встраивания индуцибельного кластера во второй генетически безопасный якорный сайт; где указанный индуцибельный кластер содержит ДНК, кодирующую некодирующую последовательность РНК, функционально присоединенную к индуцибельному промотору, и где указанный промотор регулируется белком-регулятором транскрипции;

где указанные первый и второй генетически безопасные якорные сайты являются различными.

Кроме того, в этом аспекте изобретения, в другой GSH, который отличается от первого и второго GSH, может быть введен дополнительный идентичный или отличающийся индуцибельный кластер. Такой индуцибельный кластер может содержать ДНК, кодирующую некодирующую последовательность РНК или любую другую генную последовательность, функционально присоединенную к индуцибельному промотору, где указанный промотор регулируется белком-регулятором транскрипции.

Более конкретно, этот способ позволяет осуществлять нокдаун эндогенного гена в клетке. Таким образом, настоящее изобретение относится к способу снижения уровня транскрипции и/или трансляции эндогенного гена в клетке, где указанный способ включает следующие стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

б) нацеленного встраивания индуцибельного кластера во второй генетически безопасный якорный сайт; где указанный индуцибельный кластер содержит ДНК, кодирующую некодирующую

последовательность РНК, функционально присоединенную к индуцибельному промотору, и где указанный промотор регулируется белком-регулятором транскрипции, а указанная некодирующая последовательность РНК подавляет транскрипцию или трансляцию эндогенного гена;

где указанные первый и второй генетически безопасные якорные сайты являются различными.

Кроме того, в этом аспекте изобретения, в другой GSH, который отличается от первого и второго GSH, может быть введен дополнительный идентичный или отличающийся индуцибельный кластер. Такой индуцибельный кластер может содержать ДНК, кодирующую некодирующую последовательность РНК или любую другую генную последовательность, функционально присоединенную к индуцибельному промотору, где указанный промотор регулируется белком-регулятором транскрипции.

В любом аспекте или варианте осуществления изобретения, эндогенный ген может кодировать белок или некодирующую РНК.

В вышеприведенных двух аспектах изобретения, индуцибельный(е) кластер(ы) содержит(ат) ДНК, кодирующую некодирующую РНК, то есть РНК, которая является функциональной, но не транслируется в белок. Эта некодирующая РНК может представлять собой любую подходящую РНК, такую, как РНК, обсуждаемую ранее, а, предпочтительно, короткую шпилечную РНК (кшРНК). В последнем аспекте изобретения, некодирующая РНК может влиять на нокдаун гена любым подходящим способом, посредством блокирования транскрипции или трансляции гена или предотвращения экспрессии в целом. В конечном счете, экспрессия указанного гена снижается или блокируется, но сам ген остается интактным.

Альтернативно, некодирующая РНК, содержащаяся в последовательности индуцибельного кластера, может включать РНК, которые могут быть использованы для нокаута эндогенного гена в клетке, а в частности, для замены или дизрупции самого гена. Подходящими некодирующими РНК, которые могут быть использованы в этом аспекте изобретения, являются элементы каркаса CRISPR/Cas9, а более конкретно, направляющие РНК (гРНК), которые нацелены на эндогенный ген.

Таким образом, в одном из своих аспектов, настоящее изобретение относится к способу нокаута эндогенного гена в клетке, включающему следующие стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, и гена, кодирующего Cas9 или его производное, в первый генетически безопасный якорный сайт; и

б) нацеленного встраивания индуцибельного кластера во второй генетически безопасный якорный сайт; где указанный индуцибельный кластер содержит руководящую РНК, функционально присоединенную к индуцибельному промотору, и где указанный промотор регулируется белком-регулятором транскрипции, а указанная последовательность рРНК нацелена на эндогенный ген;

где указанные первый и второй генетически безопасные якорные сайты являются различными.

Кроме того, в этом аспекте изобретения, в другой GSH, который отличается от первого и второго GSH, может быть введен дополнительный идентичный или отличающийся индуцибельный кластер. Такой индуцибельный кластер может содержать генную последовательность, функционально присоединенную к индуцибельному промотору, где указанный промотор регулируется белком-регулятором транскрипции.

Таким образом, в вышеприведенном аспекте изобретения, транскрипция рРНК является регулируемо индуцированной.

В другом своем аспекте, настоящее изобретение относится к способу снижения уровня транскрипции и/или трансляции эндогенного гена в клетке, где указанный способ включает следующие стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, в первый аллель генетически безопасного якорного сайта; и

б) нацеленного встраивания индуцибельного кластера во второй аллель генетически безопасного якорного сайта, где указанный индуцибельный кластер содержит ДНК, кодирующую некодирующую последовательность РНК, функционально присоединенную к индуцибельному промотору, и где указанный промотор регулируется белком-регулятором транскрипции, а

указанная некодирующая последовательность РНК подавляет транскрипцию или трансляцию эндогенного гена.

Кроме того, настоящее изобретение относится к способу нокаута эндогенного гена в клетке, включающему следующие стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, и гена, кодирующего Cas9 или его производное, в первый аллель генетически безопасного якорного сайта; и

б) нацеленного встраивания индуцибельного кластера во второй аллель генетически безопасного якорного сайта; где указанный индуцибельный кластер содержит руководящую РНК, функционально присоединенную к индуцибельному промотору, и где указанный промотор регулируется белком-регулятором транскрипции, а указанная последовательность рРНК нацелена на эндогенный ген.

Такой одностадийный нокаут или нокдаун являются новыми и могут составлять часть изобретения.

В одном из своих аспектов, настоящее изобретение относится к способу прямого перепрограммирования плюрипотентных стволовых клеток, включающему следующие стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

б) нацеленного встраивания индуцибельного кластера во второй генетически безопасный якорный сайт, где указанный индуцибельный кластер содержит генную последовательность, кодирующую ключевой фактор транскрипции линии дифференцировки, функционально присоединенный к индуцибельному промотору, где указанный индуцибельный промотор регулируется белком-регулятором транскрипции; и

где указанные первый и второй генетически безопасные якорные сайты являются различными.

В другой GSH, который отличается от первого и второго GSH, может быть введен другой(ие) или дополнительный(е) индуцибельный(е) кластер(ы).

Прямое перепрограммирование плюрипотентных стволовых клеток в зрелые клетки конкретных типов является в высокой степени

желательным, и может быть достигнуто с применением стратегии двойного нацеливания согласно изобретению. Конкретные способы получения клеток определенных типов описаны ниже.

В одном из своих аспектов, настоящее изобретение относится к способу продуцирования миоцитов из плюрипотентных стволовых клеток, включающему следующие стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

б) нацеленного встраивания гена *MYOD1*, функционально присоединенного к индуцибельному промотору, во второй генетически безопасный якорный сайт, где указанный индуцибельный промотор регулируется белком-регулятором транскрипции;

где указанные первый и второй генетически безопасные якорные сайты являются различными; и

культивирования указанных клеток в присутствии ретиноевой кислоты.

Ген *MYOD1* представляет собой ген, кодирующий миогенный блок дифференцировки 1. Предпочтительной ретиноевой кислотой (RA) являются все транс-RA.

В другом своем аспекте, настоящее изобретение относится к способу продуцирования миоцитов из плюрипотентных стволовых клеток, экспрессирующих *MYOD1*, где указанный способ включает культивирование указанных клеток в присутствии ретиноевой кислоты.

Предпочтительно, RA представляет собой все транс-RA. Предпочтительно, клетки сверхэкспрессируют *MYOD1*.

В другом своем аспекте, настоящее изобретение относится к способу продуцирования олигодендроцитов из плюрипотентных стволовых клеток, где указанный способ включает стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

б) нацеленного встраивания любой комбинации генов *SOX 10*, *OLIG2*, *NKX2.2* и *NKX6.2*, функционально присоединенных к индуцибельному промотору, во второй генетически безопасный

якорный сайт, где указанный индуцибельный промотор регулируется белком-регулятором транскрипции; и где указанные первый и второй генетически безопасные якорные сайты являются различными.

Гены *SOX-10*, *OLIG2*, *NKX2.2*, *NKX6.2* кодируют фактор транскрипции *SOX-10*, *OLIG2*, *NKX2.2* и *NKX6.2*, соответственно.

Описание чертежей

Фигура 1 (a-d): Подтверждение оптимизированной системы сверхэкспрессии, нацеленной на два генетически безопасных якорных сайта. На фигуре 1(a) представлена конструкция векторов, нацеленных на гены локусов *hROSA26* и *AAVS1*. *HAR*: гомологичное плечо, *SA*: акцептор сплайсинга, *T2A*: сигнал удаления рибосомы *T2A*; *Neo*: ген резистентности к неомицину; *Puro*: ген резистентности к пурамицину. *pA*: сигнал полиаденилирования; *CAG*: конститутивно активный промотор *CAG*; *rtTA*: *rtTA* третьей генерации; *TRE*: индуцибельный Tet-чувствительный элемент; *EGFP*: активированный белок, флуоресцирующий в зеленом диапазоне спектра. На фигуре 1(b) показана кинетика индуцирования и «спасения» *EGFP* (1(c)) в *EGFP*-экспрессирующих *hESC*, детектируемая с помощью проточной цитометрии (медианная интенсивность флуоресценции, *MFI*). Результаты были получены для двух биологических дубликатов на определенный момент времени и выражены как среднее $\pm$ ср. кв. ош. Все величины были нормализованы на максимальную интенсивность флуоресценции через 5 дней после обработки доксициклином (на этой фигуре показаны на день 0). На фигуре 1(d) показан ответ на дозу доксициклина для сверхэкспрессии *EGFP* в *EGFP*-экспрессирующих *hESC* после индуцирования доксициклином в течение 5 дней. Результаты были получены для двух биологических дубликатов на определенное условие и выражены как среднее $\pm$ ср. кв. ош. Все величины были нормализованы на максимальную интенсивность флуоресценции, измеренную в этом эксперименте. Уровни экспрессии *EGFP* в *hPSC* с конститутивным *CAG-EGFP*, нацеленным на *GSH*, и в *hPSC* с индуцибельным *TRE-EGFP*, нацеленным на два *GSH*, после индуцирования доксициклином. *hPSC* дикого типа и не-индуцированные *TRE-EGFP*-клетки были включены как негативный

контроль.

Фигура 2 (a-d): Общий обзор экспериментального подхода и результаты быстрого одностадийного превращения hPSC в нейроны (i-нейроны) после обработки доксициклином (dox). На фигуре 2(a) схематически представлено такое превращение, в котором клетки, трансформированные NGN2 согласно изобретению, индуцируют для их дифференцировки в нейроны после Dox-обработки. На фигуре 2(b) указано время прямого перепрограммирования генерации i-нейронов из hESC, подтвержденного с помощью количественного ОТ-ПЦР-анализа, который продемонстрировал временной паттерн экспрессии маркерных генов всех нейронов (MAP2, SYP), переднего мозга (BRN2, FOXG1) и глутаматергических нейронов (VGLUT2, GRIA4). Клетки анализировали в указанные дни обработки доксициклином. Величины представлены по отношению к эндогенному гену PBGD «домашнего хозяйства» и нормализованы на условия сообщения плюрипотентности. Результаты были получены для трех биологических дубликатов на определенный момент времени и выражены как среднее $\pm$ ср. кв. ош. На фигуре 2(c) проиллюстрирована количественная оценка нейронов, позитивных по β III-тубулину (TUBB3), посредством иммунологического окрашивания i-нейронов, происходящих от hESC после индуцирования в течение одной недели. Недифференцированные клетки были использованы в качестве негативного контроля (Контроль), а номера были представлены для генерации i-нейронов в только что выделенных NGN2-экспрессирующих hESC и после 25 пассажей (+P25). На фигуре 2(d) представлены фотографии клеток, где указано время прямого перепрограммирования генерации i-нейронов из hESC на серийных фазово-контрастных изображениях, на которых проиллюстрированы морфологические изменения.

Фигура 3 (a-d): Прямое перепрограммирование hPSC в скелетные миоциты. На фигуре 3a схематически представлено быстрое одностадийное превращение hPSC в скелетные миоциты посредством индуцибельной сверхэкспрессии MYOD1 и обработки ретиноевой кислотой. На фигуре 3b проиллюстрирован количественный ОТ-ПЦР-анализ временного паттерна экспрессии маркерных генов миоцитов в

процессе получения i-миоцитов из hESC. Все величины даны по отношению к hPSC. Результаты были получены для трех биологических дубликатов на определенный момент времени и выражены как среднее $\pm$ ср. кв. ош. На фигурах 3 (c) и (d) показана количественная оценка MHC-позитивных клеток, проведенная с помощью проточной цитометрии через десять дней после индуцирования, и такая оценка показала, что OPTi-MYOD1-hPSC сохраняют свою миогенную активность даже после длительных периодов культивирования и пассажа (p) в результате нацеленной интеграции системы MYOD1. Недифференцированные клетки были использованы в качестве негативного контроля (Контроль), а номера были представлены для генерации i-миоцитов в только что выделенных OPTi-MYOD1-hESC или в тех же самых клетках после 50 пассажей (+P50).

Фигура 4 (a-f): Стратегия нацеливания системы сверхэкспрессии Tet-ON на два GSH. На фигуре 4(a) проиллюстрирована экспериментальная рабочая схема последовательного нацеливания на локусы hROSA26 и AAVS1 в hPSC. Обозначения: Cas9n: Cas9-эндонуклеаза, мутантная по D10A-никазе от *S. Pyogenes*; ZFN: нуклеаза «цинковые пальцы»; Neo: неомицин; Puro: пурамицин; rtTA: обратный транс-активатор тетрациклина третьей генерации. На этой фигуре показана индуцибельная система экспрессии EGFP (i-EGFP). На фигуре 4(b) схематически показана стратегия нацеливания на hROSA26. На фигуре 4(c) показана стратегия нацеливания на AAVS1. Обозначения на фигурах 4(b) и (c): R26-prom: промотор локуса ROSA26 (гена THUMPD3-AS1); AAV-prom: промотор локуса AAVS1 (гена PPP1R12C); ZFN: нуклеазы «цинковые пальцы»; 5'-HAR/3'-HAR: вышерасположенное/нижерасположенное гомологичное плечо. SA: акцептор сплайсинга; T2A: пептид T2A; pA: сигнал полиаденилирования; CAG: ранний энхансер CMV, гибридный промотор куриного β -актина и кроличьего β -глобина; TRE: Tet-чувствительный элемент; EGFP: активированный белок, флуоресцирующий в зеленом диапазоне спектра. На фигуре 4(d) схематически представлена стратегия генотипирования, используемая для идентификации линий

hPSC, в которые была осуществлена правильная доставка в локусы hROSA26 и AAVS1; GSH-prom: промотор GSH (hROSA26 и AAVS1, соответственно); WT: дикого типа; индуцибельный кластер: полноразмерная экзогенная последовательность, интегрированная после нацеливания. ПЦР локуса: ПЦР, охватывающая локус-мишень с обоими праймерами, связывающимися исключительно с геномной ДНК, находящейся за пределами геномной последовательности, соответствующей гомологичным плечам. Примечательно то, что из-за высокого содержания GC, промотор CAG не может быть амплифицирован с помощью стандартной ПЦР. Следовательно, правильная инсерция CAG-содержащего экспрессионного кластера приводит к потере ПЦР-ампликона. Присутствие полосы дикого типа указывает на присутствие аллелей, не являющихся мишенями; отсутствие полосы дикого типа указывает на гомозиготное нацеливание. 5'-INT/3'-INT: ПЦР: ПЦР, охватывающие 5'- и 3'-сайты инсерции, соответственно. ПЦР-ампликоны, имеющие нужные размеры, указывают на правильную интеграцию. 3'BB-PCR: ПЦР, охватывающая стык каркаса гомологичного плеча/таргетирующего вектора. Присутствие ПЦР-продукта указывает на неспецифическую нежелательную интеграцию донорной плазмиды. На фигуре 4(e) показана фотография геля, где указаны результаты генотипирования для отобранных гетерозиготных (HET) и гомозиготных (HOM) H9 hESC с локусами-мишенями hROSA26-CAG-rtTA. На фигуре 4(f) показана фотография геля, где указаны результаты генотипирования для отобранных гетерозиготных (HET) и гомозиготных (HOM) H9 hESC с локусами-мишенями AAVS1-TRE-EGFP. 1kb+: 1 т.п.о. плюс ДНК-лэддер; WT: hESC дикого типа; PL: таргетирующая плазида; H₂O: вода, контроль.

Фигура 5(a-e): Разработка оптимизированной стратегии индуцибельной сверхэкспрессии (OPTi-OX) на основе нацеливания на два GSH в hPSC. На фигуре 5(a) указанные hESC с индуцибельными EGFP H9, нацеленными на два GSH, были распределены по четырем экспериментальным группам независимо от того, было ли проведено успешное нацеливание на один или оба аллеля локусов hROSA26 и AAVS1, соответственно. На фигуре 5(b) проиллюстрировано

детектирование белка rtTA с помощью вестерн-блот-анализа гетеро- и гомозиготных H9-hESC, на hROSA26-CAG-rtTA-локусы которых было проведено успешное нацеливание. Человеческие ESC, несущие rtTA второй генерации в случайном положении генома, были включены в качестве контрольного образца. α -Тубулин: загрузочный контроль. На фигуре 5(с) проиллюстрирован проточный цитометрический анализ репрезентативных примеров различных hESC с индуцибельными EGFP, нацеленными на два GSH, как описано на фигуре 5(а). На фигуре 5(д) проиллюстрирована медианная интенсивность флуоресценции (MFI) EGFP, экспрессирующегося в различных hESC с индуцибельными EGFP, нацеленными на два GSH, как описано на фигуре 5(а). Клетки анализировали с помощью проточной цитометрии в контрольных условиях (без доксициклина, CTR) или после 5-дневной обработки доксициклином (DOX). Каждая исходная точка соответствует отдельной клональной линии. CAG-EGFP-hESC и hESC дикого типа (WT) были включены для сравнения. Статистический анализ групп, обработанных доксициклином ($n=4-5$, как показано на фигуре), продемонстрировал, что уровни экспрессии EGFP были самыми высокими в клонах, гомозиготных по двум локусам (односторонний ANOVA с *post-hoc* критерием Дюнетта; $F(2, 10)=25,34$, $p=0,0001$; **** $p<0,0001$; ** $p=0,0026$). Это условие было отобрано для последующих экспериментов. На фигуре 5(е) указан процент EGFP⁺-клеток в различных hESC с индуцибельными EGFP, нацеленными на два GSH, как описано на фигуре 5(а).

Фигура 6(а-д): Характеристика стратегии OPTi-OX в hPSC и во время дифференцировки в зародышевый слой. На фигуре 6(а) проиллюстрирован проточный цитометрический анализ на уровне EGFP в живых hPSC с успешной доставкой и после их дифференцировки в три зародышевых слоя после обработки доксициклином в течение пяти дней. Параметры доступа были выбраны так, чтобы они включали высокие уровни экспрессии индуцированного EGFP (DOX). Популяции неиндуцированного контроля (CTR) указаны непосредственно слева на оси y. На фигурах 6(б) и 6(с) систематизированы графики данных проточной цитометрии на 6(а), включая медианную интенсивность флуоресценции (MFI) и процент

EGFP⁺-клеток. На фигуре 6(d) показана гистограмма, на которой отложены результаты количественного ОТ-ПЦР-анализа на уровне экспрессии мРНК EGFP в гомозиготных плюрипотентных стволовых клетках и после их дифференцировки в три зародышевых слоя. WT: дикого типа.

Фигура 7: Характеристика человеческих i-нейронов. Результаты количественного ОТ-ПЦР-анализа продемонстрировали быстрое ингибирование факторов плюрипотентности *NANOG* и *OCT4* после обработки доксициклином.

Фигура 8: Передача RA-сигнала во время индуцирования миоцитов. На этой фигуре проиллюстрирован кол.ПЦР-анализ шести ретиноидов и ретиноидных рецепторов во время индуцирования миоцитов, и этот анализ продемонстрировал экспрессию *RAR α* , *RAR β* и всех трех изоформ *RXR*, но не *RAR γ* , во время индуцирования i-миоцитов. А означает α , В означает β и G означает γ .

Фигуры 9(a)-9(c): Характеристика развития OPTi-MYOD1-hESC в человеческие i-миоциты. На фигуре 9(a) показана прямая последовательность операций по перепрограммированию OPTi-MYOD1 hPSC в индуцированные миоциты. Морфологические изменения были задокументированы с использованием автоматизированного устройства для получения фазово-контрастных изображений, которые регистрировали через каждые 30 минут на системе Nikon Biostation IM с временным запаздыванием. Масштаб: 200 мкм. На фигуре 9(b) показаны результаты кол.ПЦР-анализа, которые продемонстрировали быстрое ингибирование факторов плюрипотентности *NANOG* и *OCT4* после обработки OPTi-NGN2 hESC доксициклином (левый график). Все пять основных человеческих миоцитов скелетных мышц, специфичных к изоформам тяжелой цепи миоцитов (кодируемых геном семейства *MYH*), подвергались сильной активации в процессе прямого перепрограммирования в миоциты (правый график). Эти миоциты включают две изоформы, которые экспрессируются в процессе развития эмбриона и развития мышц после рождения (эмбриональную изоформу MYH3; неонатальную изоформу MYH8), и три этих изоформы, которые обычно экспрессируются в скелетной мышце взрослого человека [MYH7 в волокнах, ответственных за слабые судороги

(типа I); MYH2 в волокнах, резистентных к быстрой усталости и к судорогам (типа IIa), и MYH1 в волокнах, ответственных за быструю утомляемость (типа IIx)]. В противоположность этому, MYH4, который представляет МНС-изоформу в волокнах миоцитов, ответственных за быстрые судороги и быструю утомляемость у кошек, не экспрессируется в значительных количествах у человека (<1%), а также не индуцируется в процессе прямого перепрограммирования. На фигуре 9(с) показано, что индуцированные миоциты скелетной мышцы экспрессируют типичные маркерные белки широкого ряда, включая F-актин (визуализированный посредством AlexaFluor488-конъюгированного фаллоидинового токсина), адгезивную молекулу нервных клеток (NCAM), десмин (DES), тяжелую цепь миозина (MYH), титин (TTN), α -актинин (ACTN2) и тропонин Т (TNNT), но не маркеры предшественников миобластов PAX3 и PAX7. Все образцы были подвергнуты контрастному окрашиванию миогенином (MYOG). Масштаб: 50 мкм. DAPI: окрашивание ядра.

Фигура 10: Эти три графика иллюстрируют результаты кол.ПЦР-анализа на общий MYOD1, эндогенный MYOD1 и MYOG через 2 дня после индуцирования OPTi-MYOD1 hPSC различными концентрациями доксициклина. Результаты кол.ПЦР-анализа были представлены через 48 часов после индуцирования различными концентрациями доксициклина. Уровни экспрессии отложены на графике по отношению к уровням экспрессии эндогенного гена PBGD домашнего хозяйства.

Фигура 11: Описание системы Tet-ON. Tet-ON состоит из двух компонентов: В верхней части представлен кластер-активатор, в котором конститутивный промотор (сР) запускает экспрессию rtTA (обратного трансаактиватора тетрациклина). RtTA представляет собой гибридный белок, который состоит из мутантной формы прокариотического репрессора Tet (TetR) и домена трансаактиватора транскрипции VP16 (происходящих от вируса простого герпеса). В нижней части представлен домен-респондент. Он состоит из индуцибельного промотора (TRE, элемента, чувствительного к Tet) и представляющего интерес гена. TRE представляет собой искусственный промотор, чувствительный к rtTA. Он состоит из 7

последовательно расположенных оперонов tet (tetO7) и сильного минимального промотора CMV (mCMV), который сам по себе не является активным, а осуществляет лишь рекрутинг механизма транскрипции после связывания rtTA с семью оперонами tet. Доксициклин, производное тетрациклина, необходим для связывания мутантного TetR с TRE, что приводит к экспрессии индуцибельного кластера, а в данном случае EGFP (pA: сигнал полиаденилирования).

Фигура 12 (a-d). Прямое превращение hPSC в олигодендроциты. На фигуре 12(a) схематически представлена экспериментальная стратегия быстрого превращения OPTi-OLIG2-SOX10-hPSC в клетки линии дифференцировки олигодендроцитов (i-OPC и i-OL). На фигуре 12(b) проиллюстрирована количественная оценка BrdU-позитивных клеток после 3 последовательных пассажей каждые 4 дня и последовательной подачи BrdU-импульсов каждые последние 4 дня (P=число пассажей). На фигуре 12(c) проиллюстрирован количественный ОТ-ПЦР-анализ временного паттерна экспрессии генов, кодирующих белки, ассоциированные с миелином (CNP, MAG, MBP, MOG и PLP) в процессе образования i-олигодендроцитов из hPSC. OPTi-OLIG2-SOX10-hPSC индуцировали в среде для культивирования олигодендроцитов, в которую добавляли PDGF $\alpha\alpha$ и FGF2. После индуцирования в течение одной недели, митогены удаляли для инициации конечной стадии дифференцировки. Все величины представлены по отношению к эндогенному гену PBGD домашнего хозяйства и нормализованы на условия сообщения плюрипотентности. Результаты были получены для 2-3 биологических дубликатов на определенный момент времени и выражены как среднее $\pm$ ср. кв. ош. На фигуре 12(d) проиллюстрирована количественная оценка CNP- и PLP-позитивных клеток посредством иммунологического окрашивания i-олигодендроцитов, происходящих от OPTi-OLIG2-SOX10-hPSC после индуцирования в течение 20 дней. Недифференцированные клетки были использованы в качестве негативного контроля, а цифрами отмечены i-олигодендроциты в только что выделенных OPTi-NGN2-hPSC и после 50 пассажей (+P50).

На фигуре 13 схематически представлены принципы настоящего

изобретения. По существу, на этой фигуре представлена инсерция в два различных генетически безопасных якорных сайта согласно изобретению. Одна такая инсерция регулирует экспрессию генной последовательности в индуцибельном кластере во второй вставке. В полицистронные векторные конструкции может быть включен дополнительный генетический материал, как показано на фигуре. Кроме того, может быть осуществлено нацеливание на более, чем два генетически безопасных якорных сайта так, чтобы множество индуцибельных кластеров или других генетических материалов могли находиться под контролем модулятора, расположенного в первом GSH-сайте.

На фигуре 14(a-f) представлены результаты, иллюстрирующие получение индуцибельной системы нокдауна на основе hSPC, в которых было осуществлено нацеливание на два GSH. На фигуре 14a проиллюстрирован экспериментальный метод: H1 - промотор H1, TO - оперон tet, tetR - репрессор тетрациклина. На фигуре 14b схематически представлены трансгенные аллели, используемые для получения hESC, экспрессирующих репортерный трансген EGFP, который может быть подвергнут сайленсингу с использованием индуцибельной кшРНК EGFP. На фигуре 14c показана экспрессия EGFP в отсутствии или в присутствии тетрациклина в течение 5 дней в hESC, в которые были доставлены указанные комбинации индуцибельных кшРНК и tetR EGFP (STD=стандарт дикого типа, OPT=оптимизированный по кодонам). hESC с двумя сайтами-мишенями, которые не несут кшРНК EGFP, были использованы в качестве негативного контроля. n.s.= $p>0,05$ (не значимые), **= $p>0,01$, ***= $p>0,001$ по сравнению с той же самой линией tetR, которая не содержит tet и кшРНК. На фигуре 14d представлен репрезентативный вестерн-блот-анализ на tetR в hESC с ROSA26-мишенями, экспрессирующих STD или OPT tetR. HET=гетерозиготный таргетинг, HOM=гомозиготный таргетинг. hESC со случайной интергацией STD tetR показаны как позитивный контроль, а h9-hESC дикого типа показаны как негативный контроль. TUB4A4A представляет собой загрузочный контроль. Различные количества белка загружали для облегчения количественного сравнения. Фигура 14(E): кинетика нокдауна и спасения EGFP в EGFP-OPTiKD-hESC, оцененная с помощью

проточной цитометрии (MFI) и кол.ПЦР (мРНК). Результаты представлены для 2 независимых культур на определенный момент времени. Фигура 14(F): Кривая «доза тетрациклина-ответ» на нокдаун EGFP в EGFP-OPTiKD-hESC. Представлена полумаксимальная ингибирующая концентрация (IC50). Результаты представлены для двух независимых культур на дозу и указано как среднее.

Фигура 15(a, b и c). Подтверждение локусов ROSA26 и AAVS1, в которые были встроены GSH. На фигуре 15a проиллюстрирован экспериментальный метод получения hPSC с репортерным EGFP GSH для тестирования экспрессии в GSH в процессе дифференцировки. Нейроны, олигодендроциты и астроциты были получены в виде общих культур, содержащих смесь этих линий дифференцировки, а все остальные клетки были получены отдельно. На фигуре 15b схематически представлены трансгенные аллели с репортерным EGFP ROSA26 и AAVS1. R26-prom: промотор локуса ROSA26; AAV-prom: промотор локуса AAVS1; 5'-HAR/3'-HAR: вышерасположенное/нижерасположенное гомологичное плечо; SA: акцептор сплайсинга; T2A: самоотщепляющийся пептид T2A; Neo: резистентность к неомицину; PuRo: резистентность к пурамицину; pA: сигнал полиаденилирования; CAG: промотор CAG; EGFP: активированный белок, флуоресцирующий в зеленом диапазоне спектра. Фигура 15(C): экспрессия EGFP в отсутствии или в присутствии тетрациклина в течение 5 дней в hESC, в которые были доставлены указанные комбинации индуцибельной кшРНК и tetR EGFP (стандартный tetR дикого типа, STDtetR, или оптимизированный по кодонам tetR, OPTtetR). hESC с двумя сайтами-мишенями, которые не несут кшРНК EGFP, были использованы в качестве негативного контроля. Результаты представлены для 2-3 отдельных линий на одно условие (таблица 1). n.s.= $p>0,05$ (не значимые), **= $p>0,01$, ***= $p>0,001$ по сравнению с той же самой линией tetR, которая не содержит tet и кшРНК (ANOVA с *post-hoc* сравнениями Хольма-Сидака).

Фигура 16(a-d). Получение hESC с репортерным EGFP ROSA26 и AAVS1. Фигура 16(A): схематически представлен метод нацеливания на ROSA26 и стратегии генотипирования, используемые для

идентификации линий с правильными мишенями. Cas9n: Cas9-эндонуклеаза, мутантная по D10A-никазе от *S. Pyogenes*; R26-prom: промотор локуса ROSA26 (гена THUMPD3-AS1); 5'-HAR/3'-HAR: вышерасположенное/нижерасположенное гомологичное плечо; трансген: область, интегрированная после доставки гена; ПЦР локуса: ПЦР-продукт локуса ROSA26 дикого типа (указывающий на аллель, не являющийся мишенью); ПЦР локуса/потеря аллеля: ПЦР-продукт аллеля мишени/ПЦР, которая не происходит в том случае, если трансген содержит GC-богатый промотор CAG (что указывает на ожидаемое нацеливание на трансген); 5' INT/3' INT PCR: ПЦР-продукт 5'-концевой/3'-концевой области интеграции трансгена (указывающий на ожидаемое нацеливание на трансген); 5' BB/3' BB PCR: ПЦР-продукт векторного остова с 5'-концом/3'-концом (указывающий на неспецифическую нежелательную интеграцию плазмиды). Следует отметить, что аналогичное нацеливание и аналогичные стратегии генотипирования были применимы и для нацеливания на locus AAVS1.

Фигура 16(B): схематически представлены трансгенные аллели ROSA26, полученные для проверки наилучшей стратегии экспрессии конститутивного EGFP (активированного белка, флуоресцирующего в зеленом диапазоне спектра). ENDO-EGFP: EGFP, индуцированный эндогенным промотором ROSA26 (R26-prom; таргетирующий вектор pR26-Puro_ENDO-EGFP); EF1 α -EGFP: EGFP, индуцированный промотором фактора элонгации 1 α (таргетирующий вектор pR26-Neo_EF1 α -EGFP); CAG-EGFP: EGFP, индуцированный промотором CAG (таргетирующий вектор pR26-Neo_CAG-EGFP); SA: акцептор сплайсинга; Puro: резистентность к пуромицину (пуромицин-N-ацетилтрансфераза); Neo: резистентность к неомицину (неомицин-фосфотрансфераза II); pA: сигнал полиаденилирования.

Фигура 16(C): Проточно-цитометрический количественный анализ для определения процента EGFP-позитивных клеток (EGFP+; окно возбуждения показано) и для оценки медианной интенсивности флуоресценции EGFP (MFI) в репрезентативных клональных линиях hESC с репортерным ROSA26-EGFP, или H9-hESC дикого типа. Фигура 16(D): процент EGFP-позитивных клеток в hESC с репортерным ROSA26-EGFP. Результаты представлены для 3 клонов с нацеливанием на гетерозиготные

ROSA26 на одно условие.

Фигура 17. Подтверждение оптимизированных индуцибельных стратегий нокдауна после дифференцировки hPSC. На графике показана экспрессия EGFP, оцененная с помощью кол.ПЦР в отсутствии (CTR) или в присутствии тетрациклина в течение 5 дней (TET) в клетках указанных типов, происходящих от hESC, содержащих EGFP OPTiKD (iKD) и sOPTiKD (siKD). Уровни EGFP представлены по сравнению с уровнями в контрольных условиях в той же самой линии для каждой отдельной линии дифференцировки. Аббревиатуры означают линии дифференцировки, описанные на фиг. 15 (pluri: недифференцированные). Результаты представлены для двух независимых культур на одно условие.

Фигура 18(a-d). Разработка оптимизированной индуцибельной стратегии нокаута CRISPR/Cas9 в hPSC. На фигуре 18a проиллюстрирован экспериментальный способ получения индуцибельных hPSC с нокаутом (iKO). На фигуре 18b схематически описана процедура клонирования для получения таргетирующих векторов AAVS1 с индуцибельным кластером pPHK. На фигуре 18c показаны трансгенные аллели, используемые для получения hESC, экспрессирующих репортерный трансген EGFPd2, который может быть нокаутирован по CRISPR/Cas9 с использованием индуцибельной pPHK EGFP (EGFP sOPTiKO hESC). Bsd: резистентность к бластицидину; EGFPd2: дестабилизированный EGFP. Фигура 18(d): Проточно-цитометрическая количественная оценка индуцибельной кинетики нокаута EGFPd2 в клетках sOPTiKO, как показано на фигуре 19c (pPHK 2-TO) и b (pPHK 3-2TO). Ежедневно проводили мониторинг процента EGFP-позитивных клеток после добавления тетрациклина. Результаты представлены для 2 независимых культур.

Фигура 19(a-e): Разработка оптимизированной индуцибельной стратегии нокаута CRISPR/Cas9 в hESC. (A-D) Представлен репрезентативный проточно-цитометрический анализ на экспрессию EGFPd2 в EGFPd2-гомозиготных sOPTiKO hESC, несущих указанные комбинации pPHK (2 или 3) и индуцибельный промотор (TO или 2TO, см. фиг. 19e). Таргетирующие векторы: pAAV-Puro_siKOEGFP-2 (19a), pAAV-Puro_siKO-2TO-EGFP-2 (19b), pAAV-Puro_siKO-EGFP-3 (19c), pAAV-Puro_siKO-2TO-EGFP-3 (19d). Клетки культивировали в

присутствии тетрациклина (TET) в течение 5 дней или поддерживали в контрольных условиях (CTR) в отсутствии тетрациклина. Следует отметить, что гистограммы были нормализованы так, чтобы площадь под кривой соответствовала 1 (100%) для всех представленных образцов в целях облегчения прямого визуального сравнения. Фигура 19(e): Нуклеотидные последовательности индуцибельных промоторов H1 Pol III для системы sOPTiKO, содержащей один или два оперона tet (H1-TO и H1-2TO, соответственно). Ключевые признаки последовательности выделены. Показаны сайты рестриктирующих ферментов, используемых для клонирования рРНК (фиг. 18B). DSE: элемент дистальной последовательности; PSE: элемент проксимальной последовательности; TETO2: оперон tet; +1: исходное положение транскрипции РНК.

На фигурах 20-33 представлены описания карт различных плазмид, используемых в примерах согласно изобретению. Такими плазмидами являются:

- 20) pSpCas9n(BB)_R26-R
- 21) pSpCas9n(BB)_R26-L
- 22) pR26_CAG_EGFP
- 23) pR26_CAG_rtTA
- 24) pZFN-AAVS1-L-ELD (нуклеаза «цинковые пальцы», слева)
- 25) pZFN-AAVS1-R-KKR (нуклеаза «цинковые пальцы», справа)
- 26) pAAV_CAG_EGFP (донор)
- 27) pR26-Neo_CAG-OPTtetR (нацеливание оптимизированного по кодонам tetR на hROSA26)
- 28) pAAV-Puro_iKD (нацеливание индуцибельной кшРНК на AAVS1)
- 29) pAAV-Neo_CAG-Cas9 (нацеливание Cas9 на AAVS1)
- 30) pAAV-Puro_siKO (нацеливание индуцибельной рРНК на AAVS1)
- 31) pAAV-Puro_siKO-2TO (нацеливание индуцибельной рРНК на AAVS1, вариант с 2 оперонами tet в промоторе)
- 32) pAAV_TRE-EGFP (индуцибельная сверхэкспрессия EGFP, сцепленная с геном)
- 33) pAAV_TRE-MYOD1 (индуцибельная сверхэкспрессия MYOD1 в мышцах)

ПОДРОБНОЕ ОПИСАНИЕ

Авторами настоящего изобретения был разработан способ, который является подходящим для индуцибельной транскрипции генных последовательностей, содержащихся в индуцибельных кластерах в эукариотических клетках, а в частности, в плюрипотентных стволовых клетках и в их потомстве.

Это особенно применимо к прямому продуцированию плюрипотентных стволовых клеток посредством сверхэкспрессии индуцибельных кластеров в указанных стволовых клетках, которые стимулируют развитие зрелых клеток конкретного типа. Кроме того, этот способ также применим к нокдауну или нокауту эндогенных функций в клетках для исследования потери функции или изменения клеточных функций или поведения этих клеток. Нокдаун или нокаут могут быть также осуществлены в белок-кодирующих генах или в последовательностях ДНК, кодирующих некодирующую РНК. Эти гены или последовательности могут быть таргетированы способами согласно изобретению посредством нокаута или нокдауна.

Этот способ основан по меньшей мере на нацеливании на два безопасных якорных сайта в геноме стволовых клеток с использованием системы индуцированной транскрипции, распределенной по двум или более GSH. Однако, этот способ не ограничивается стволовыми клетками и может быть применен для модификации генома клеток любого типа, например, в исследованиях или в генотерапии. В способах согласно изобретению, один GSH модифицируют так, чтобы он содержал регулятор транскрипции, который необходим для индуцирования транскрипции генной последовательности, содержащейся в индуцибельном кластере, встроенном в другой GSH, присутствующий в другом участке генома. Регулятор транскрипции предпочтительно конститутивно экспрессируется. Предпочтительно, чтобы экзогенное вещество/агент были введены для регуляции активности белка-регулятора транскрипции, и, таким образом, для регуляции экспрессии индуцибельного кластера. Поскольку в способе согласно изобретению используются по меньшей мере два отдельных GSH, то присутствуют всего четыре возможных локусов инсерции, поскольку каждый GSH присутствует на обеих хромосомах диплоидного

организма. Это способствует увеличению уровня возможной транскрипции из клеток, если все четыре локуса были модифицированы способом согласно изобретению. Пример различных результатов нацеленной инсерции показан на фигуре 5а. Кроме того, в способе согласно изобретению используются по меньшей мере два различных сайта GSH. Следует отметить, что дополнительные сайты GSH могут быть использованы для введения других регуляторов транскрипции, индуцибельных кластеров или любого другого генетического материала, включая, но не ограничиваясь ими, селективные маркеры, гены резистентности к антибиотикам или к лекарственным средствам, гены, ассоциированные с системой CRISPR/Cas9 или гены с неизвестной функцией.

Таким образом, настоящее изобретение относится к способу регуляции экспрессии встроенной генной последовательности в клетке, включающему следующие стадии:

а) нацеленного встраивания генной последовательности, кодирующей белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

б) нацеленного встраивания индуцибельного кластера во второй генетически безопасный якорный сайт, где указанный индуцибельный кластер содержит указанную генную последовательность, функционально присоединенную к индуцибельному промотору, где указанный промотор регулируется белком-регулятором транскрипции,

где указанные первый и второй генетически безопасные якорные сайты являются различными.

Кроме того, в этом аспекте изобретения, в другой GSH, который отличается от первого и второго GSH, может быть введен другой идентичный или другой отличающийся индуцибельный кластер. Такой индуцибельный кластер описан в настоящей заявке.

Специфические инсерции в генетически безопасные якорные сайты являются предпочтительными по сравнению со случайной интеграцией генома, поскольку предполагается, что такая модификация генома будет более безопасной и, менее вероятно, что она будет приводить к нежелательным побочным эффектам, таким как

сайленсинг экспрессии природного гена или мутации, которые приводят к образованию злокачественных клеток.

Генетически безопасный якорный сайт (GSH) представляет собой локус в геноме, в который могут быть встроены ген или другой генетический материал без каких-либо негативных воздействий на клетку или на встроенный генетический материал. Наиболее предпочтительным является GSH-сайт, в котором экспрессия встроенной генной последовательности не нарушается в результате какой-либо экспрессии сквозного прочитывания соседних генов, и такая экспрессия индуцибельного кластера минимизирует негативное влияние на последовательность событий эндогенной транскрипции. Были предложены и более формальные критерии, которые помогут определить в будущем, является ли конкретный локус GSH-сайтом (Papapetrou *et al*, 2011, *Nature Biotechnology*, 29(1), 73-8. doi:10.1038/nbt.1717). Такими критериями являются сайт, который находится: (i) на расстоянии 50 т.п.о. или более от 5'-конца любого гена, (ii) на расстоянии 300 т.п.о. или более от любого гена, ассоциированного с раком, (iii) на расстоянии 300 т.п.о. или более от любой микроРНК (миРНК), (iv) за пределами транскрипционной единицы и (v) за пределами ультраконсервативных областей (UCR). При этом необязательно, чтобы удовлетворялись все эти предложенные критерии, поскольку уже было обнаружено, что GSH не удовлетворяет всем этим критериям. Было высказано предположение, что подходящий GSH будет удовлетворять по меньшей мере 2, 3, 4 или всем этим критериям.

Другие сайты могут быть определены путем поиска сайтов, где вирусы естественным образом интегрируются без нарушения экспрессии природного гена.

Любой подходящий сайт GSH может быть использован в способе согласно изобретению, на том основании, что этот сайт позволит осуществлять инсерцию генетического материала без каких-либо негативных воздействий на клетку и позволит осуществлять транскрипцию встроенного генетического материала. Специалист в данной области может использовать такие упрощенные критерии для

определения подходящего GSH, и/или более формальные критерии, изложенные выше.

Для человеческого генома было идентифицировано несколько GSH-сайтов, и эти сайты включают локус AAVS1, локус hROSA26 и ген CLYBL. Ген CCR5 и ген HPRT также рассматриваются как возможные GSH, и дальнейшее исследование поможет идентифицировать один или более таких сайтов, как GSH в человеческом геноме.

Локус сайта интеграции аденоассоциированного вируса 1 (AAVS1) находится в гене регуляторной субъединицы 12C протеин-фосфатазы 1 (PPP1R12C) на человеческой хромосоме 19, который экспрессируется равномерно и повсеместно в тканях человека. Этот сайт служит в качестве специфического локуса интеграции для AAV серотипа 2, и, таким образом, он был идентифицирован как возможный GSH. Было показано, что AAVS1 представляет собой благоприятное окружение для транскрипции, поскольку он содержит открытую хроматиновую структуру и нативные хромосомные изоляторы, которые сообщают индуцибельным кластерам резистентность к сайленсингу. Не существует каких-либо известных негативных эффектов на клетку после разрушения гена PPP1R12C. Кроме того, индуцибельный кластер, встроенный в этот сайт, остается транскрипционно активным в клетках многих различных типов. Таким образом, считается, что AAVS1 представляет собой GSH, и этот сайт широко используется для нацеленного трансгеноза в человеческом геноме.

Сайт hROSA26 был идентифицирован на основе аналогии последовательности с GSH мышей (ROSA26 – обратно ориентированный сайт акцептора сплайсинга #26). Хотя ортологичный сайт был идентифицирован у человека, однако, этот сайт не часто используется для встраивания индуцибельного кластера. Авторами настоящего изобретения была разработана система таргетинга, специфичная для сайта hROSA26 и, таким образом, было осуществлено встраивание генетического материала в этот локус. Локус hROSA26 находится на хромосоме 3 (3p25.3) и его можно найти в базе данных Ensembl (GenBank: CR624523). Точными геномными координатами сайта интеграции являются 3:9396280–

9396303:Ensembl. Сайт интеграции находится в открытой рамке считывания (ORF) длинной некодирующей РНК THUMPД3 (обратная цепь). Поскольку сайт hROSA26 имеет эндогенный промотор, встроенный в генетический материал, то этот генетический материал может быть встроен с использованием этого эндогенного промотора, или, альтернативно, он может быть встроен путем функционального присоединения к промотору.

Инtron 2 гена, подобного гену цитрат-лиазы-бета (CLYBL), на длинном плече хромосомы 13, был идентифицирован как подходящий GSH, поскольку он является одной из идентифицированных горячих точек интеграции интегразы phiC31, происходящей от фага. Исследования показали, что индуцибельные кластеры, случайно встроенные в этот локус, являются стабильными и могут экспрессироваться. Было показано, что инсерция индуцибельных кластеров в этот GSH не оказывает негативного влияния на экспрессию локального гена (Cerbibi и др 2015, PLoS One, DOI: 10.1371). Таким образом, CLYBL обеспечивает GSH, который может быть подходящим для его использования в настоящем изобретении.

CCR5, который расположен на хромосоме 3 (в положении 3p21.31), представляет собой ген, который кодирует основной корецептор ВИЧ-1. Интерес к использованию этого сайта в качестве GSH продиктован нулевой мутацией в этом гене, что, очевидно, не будет давать побочных эффектов, но будет вызывать предрасположенность к резистентности к ВИЧ-1-инфекции. Были получены нуклеазы «цинковые пальцы», которые нацелены на третий экзон, что позволяет осуществлять инсерцию генетического материала в этом локусе. Учитывая то, что природная функция CCR5 пока еще не изучена, следует отметить, что этот сайт остается, предположительно, подходящим GSH для его использования в настоящем изобретении.

Ген гипоксантин-гуанинфосфорибозилтрансферазы (HPRT) кодирует фермент трансферазу, который играет центральную роль в продуцировании пуриновых нуклеотидов посредством пути «спасения» пурина. Таким образом, необходимо провести дополнительное исследование для гарантии того, что инсерции в этом сайте не будут нарушать нормальную клеточную функцию. Однако, этот сайт

рассматривается как GSH-сайт. Инсерции в этот сайт могут быть более приемлемыми для зрелых клеток, такие как модификации для генотерапии.

GSH были идентифицированы и в других организмах и ими являются локусы ROSA26, HRPT и Hippl1 (HLL) у мышей. Геномы млекопитающих могут включать GSH-сайты на основе псевдо-attP-сайтов. Для таких сайтов была получена интеграза hiC31, то есть, рекомбиназа, происходящая от фага *Streptomyces*, и эта интеграза служила в качестве невирусного инструмента для встраивания, поскольку она обладает способностью интегрировать плазмиду, содержащую индуцибельный кластер и несущую attB-сайт, в псевдо-attP-сайты.

GSH также присутствует в геномах растений, и модификация растительных клеток может составлять часть настоящего изобретения. GSH были идентифицированы в геномах риса (Cantos et al, Front. Plant Sci., 26 June 2014, Volume 5, Article 302, <http://dx.doi.org/10.3389/fpls.2014.00302>).

В способах согласно изобретению, инсерции присутствуют в другом GSH, и таким образом, для осуществления способа согласно изобретению необходимо по меньшей мере два GSH. Первый GSH модифицируют путем введения белка-регулятора транскрипции. Второй GSH модифицируют путем введения индуцибельного кластера, который содержит генную последовательность, функционально присоединенную к индуцибельному промотору. Другой генетический материал может быть также встроен вместе с любыми этими элементами или с обоими элементами. Генной последовательностью функционально присоединенной к индуцибельному промотору в индуцибельном кластере, предпочтительно является последовательность ДНК. Генная(ые) последовательность(и) индуцибельного кластера, предпочтительно, кодирует(ют) молекулу РНК, и, таким образом, она (они) может (могут) транскрибироваться. Транскрипция регулируется индуцибельным промотором. Молекулой РНК может быть любая последовательность, а предпочтительно, мРНК, кодирующая белок; кшРНК или рРНК.

Первым GSH может быть любой подходящий GSH-сайт. Этот GSH может, но необязательно, присутствовать вместе с эндогенным

промотором, который экспрессируется, что будет приводить к конститутивной экспрессии встроенного белка-регулятора транскрипции. Подходящим GSH является сайт hROSA26 для человеческих клеток. Альтернативно, встроенный белок-регулятор транскрипции функционально присоединяют к промотору, а предпочтительно, к конститутивному промотору. Конститутивный промотор может быть использован в комбинации с инсерцией в сайте hROSA26.

Белок-регулятор транскрипции представляет собой белок, который связывается с ДНК, а предпочтительно, последовательность-специфически связывается с ДНК-сайтом, и либо облегчает связывание фактора транскрипции с промотором, и, таким образом приводит к транскрипции последовательности ДНК (активатор транскрипции), либо блокирует этот процесс (репрессор транскрипции). Такие молекулы также известны как факторы транскрипции.

Последовательность ДНК, которая связывается с белком-регулятором транскрипции, называется сайтом связывания с фактором транскрипции или чувствительным элементом, и они находятся в промоторе или поблизости от промотора регулируемой последовательности ДНК.

Белки-активаторы транскрипции связываются с чувствительным элементом и стимулируют экспрессию генов. Такие белки являются предпочтительными в способах согласно изобретению для регуляции экспрессии индуцибельного кластера.

Белки-репрессоры транскрипции связываются с чувствительным элементом и предотвращают экспрессию гена.

Белки-регуляторы транскрипции могут быть активированы или дезактивированы посредством ряда механизмов, включая связывание вещества, взаимодействие с другими факторами транскрипции (например, гомо- или гетеродимеризация) или корегуляторными белками, фосфорилирование и/или метилирование. Регулятор транскрипции может регулироваться посредством активации или дезактивации.

Если белком-регулятором транскрипции является белок-активатор транскрипции, то, предпочтительно, чтобы белок-

активатор транскрипции активировался. Эта активация может быть осуществлена любыми подходящими способами, но, предпочтительно, чтобы белок-регулятор транскрипции активировался посредством добавления экзогенного вещества в клетку. Подача экзогенного вещества в клетку может регулироваться, и таким образом может регулироваться активация белка-регулятора транскрипции. Альтернативно, может быть введено экзогенное вещество для дезактивации белка-регулятора транскрипции, а затем, оно может быть удалено для активации белка-регулятора транскрипции.

Если белком-регулятором транскрипции является белок-репрессор транскрипции, то, предпочтительным, чтобы белок-репрессор транскрипции был дезактивирован. А поэтому, вводят вещество для предотвращения репрессии транскрипции белком-репрессором транскрипции, и таким образом может стимулироваться транскрипция.

При этом может быть использован любой подходящий белок-регулятор транскрипции, а предпочтительно, чтобы он был активируемым или дезактивируемым. Предпочтительно, чтобы экзогенное вещество могло подаваться для регуляции белка-регулятора транскрипции. Такие белки-регуляторы транскрипции также называются индуцибельными белками-регуляторами транскрипции.

Активация транскрипции, регулируемая тетрациклином, представляет собой метод индуцибельной экспрессии генов, где транскрипция обратимо включается или выключается в присутствии антибиотика тетрациклина или одного из его производных (например, доксициклина, который является более стабильным). В этой системе, белком-активатором транскрипции является белок-активатор транскрипции, реагирующий на тетрациклин (rtTa) или его производное. Белок rtTA способен связываться с ДНК в специфических последовательностях оператора TetO. Несколько повторов таких последовательностей TetO расположены выше минимального промотора (такого как промотор CMV), которые вместе образуют элемент, реагирующий на тетрациклин (TRE). Существуют две формы этой системы, в зависимости от того, приводит ли добавление тетрациклина или его производного к активации (Tet-

On) или дезактивации (Tet-Off) белка rtTA.

В системе Tet-Off, тетрациклин или его производное связываются с rtTA и деактивируют rtTA, что делает его неспособным связываться с последовательностями TRE, предотвращая тем самым транскрипцию TRE-регулируемых генов. Эта система была впервые описана Bujard et al., (1992). Proc. Natl. Acad. Sci. USA 89 (12): 5547-51.

Система Tet-On состоит из двух компонентов; (1) конститутивно экспрессируемого белка-активатора транскрипции, реагирующего на тетрациклин (rtTa) и индуцибельного промотора, чувствительного к rtTa (Tet-восприимчивого элемента, TRE). Эта система может связываться с тетрациклином или с его более стабильными производными, включая доксициклин (dox), что приводит к активации rtTa и позволяет ему связываться с последовательностями TRE и индуцировать экспрессию TRE-регулируемых генов. Применение этой системы может оказаться предпочтительным в способе согласно изобретению. Эта система показана на фигуре 11.

Таким образом, белком-регулятором транскрипции может быть белок-активатор транскрипции, реагирующий на тетрациклин (rtTa), который может активироваться или дезактивироваться антибиотиком тетрациклином или одними из его производных, которые подаются экзогенно. Если белком-регулятором транскрипции является rtTA, то индуцибельный промотор, встроенный во второй GSH-сайт, включает элемент, реагирующий на тетрациклин (TRE). Экзогенно подаваемым веществом является антибиотик тетрациклин или одно из его производных.

Варианты и модифицированные белки rtTa могут быть использованы в способах согласно изобретению, и такими белками являются усиленный трансаактиватор Tet-On (также известный как rtTA2S-M2) и Tet-On 3G (также известный как rtTA-V16, происходящий от rtTA2S-S2).

Элемент, реагирующий на тетрациклин (TRE), обычно состоит из 7 повторов бактериальной 19 п.о.-последовательности TetO, разделенных спейсерными последовательностями, вместе с минимальным промотором. Возможны варианты и модификации

последовательности TRE, поскольку минимальным промотором может быть любой подходящий промотор. Предпочтительно, чтобы минимальный промотор не обнаруживал каких-либо уровней экспрессии или обнаруживал минимальные уровни экспрессии в отсутствии связывания с rtTa. Таким образом, индуцибельный промотор, встроенный во второй GSH, может содержать TRE.

Модифицированной системой на основе регуляции тетрациклином является система T-Rex™ (ThermoFisher Scientific), в которой белком-регулятором транскрипции является белок-репрессор транскрипции, TetR. Компонентами этой системы являются (i) индуцибельный промотор, содержащий сильный предранний промотор человеческого цитомегаловируса (CMV) и два сайта оператора тетрациклина 2 (TetO2) и репрессор Tet (TetR). Последовательности TetO2 состоят из 2 копий последовательности из 19 нуклеотидов, 5'-TCCSTATCAGTGATAGAGA-3', отделенных друг от друга спейсером из 2 пар оснований. В отсутствие тетрациклина, репрессор Tet образует гомодимер, который связывается с очень высокой аффинностью с каждой последовательностью TetO2 в индуцибельном промоторе, и предотвращает транскрипцию с промотора. После добавления, тетрациклин связывается с высокой аффинностью с каждым гомодимером репрессора Tet, что делает его неспособным связываться с оператором Tet. Комплекс «репрессор Tet: тетрациклин» затем диссоциируется из оператора Tet и индуцирует экспрессию. В этом случае, белком-регулятором транскрипции является TetR, а индуцибельный промотор содержит два сайта TetO2. Экзогенно подаваемым веществом является тетрациклин или его производное.

Кроме того, настоящее изобретение относится к TetR, оптимизированному по кодонам (OPTtetR). Он может быть использован в любом описанном здесь способе, либо он может дополнительно использоваться в том случае, если желательно индуцировать стимуляцию. Этот белок может быть получен посредством мультипараметрической оптимизации бактериальной последовательности кДНК TetR. OPTtetR позволяет в десять раз повысить уровень экспрессии tetR по сравнению со стандартной

последовательностью (STDtetR). Экспрессия гомозиготного OPTtetR из tetR была достаточной для утечки кшРНК при сохранении индуцирования нокдауна, как описано в примерах. Последовательность OPTtetR включена в настоящее описание в качестве последовательности, сравниваемой со стандартной последовательностью. Заявленные здесь последовательности по меньшей мере на 75%, 80%, 85% или 90%, а более конкретно на 91, 92, 93, 94, 95, 96, 97 или 99% гомологичны этой последовательности. В этих последовательностях указаны остатки, которые как было показано, были заменены между STDtetR и OPTtetR, и предпочтительно, чтобы эти остатки не были заменены в любом производном OPTtetR, поскольку считается, что они играют важную роль в улучшении свойств. Любое производное будет, но необязательно, сохранять эти модификации в указанных положениях.

Известны и другие системы индуцибельной экспрессии, которые могут быть использованы в способе согласно изобретению. Они включают полную индуцибельную систему регуляции от Agilent Technologies. Она основана на гормоне экдизоне насекомых или его аналоге понастероне A (ponA), которые могут активировать транскрипцию в клетках млекопитающих, трансфицированных геном рецептора экдизона *Drosophila melanogaster* (EcR) и индуцибельным промотором, содержащим сайт связывания с рецептором экдизона. EcR является членом семейства ядерных рецепторов, а именно, ретиноид-Х-рецепторов (RXR). У человека, EcR образует гетеродимер с RXR, который связывается с элементом, реагирующим с экдизоном (EcRE). В отсутствии ponA, транскрипция подавляется гетеродимером.

Таким образом, белком-регулятором транскрипции может быть белок-репрессор, такой как рецептор экдизона или его производное. Примерами последнего являются синтетический рецептор VgEcR от Agilent Technologies, который представляет собой гибрид EcR, ДНК-связывающего домена рецептора глюкокортикоида и домена активации транскрипции вируса простого герпеса VP16. Индуцибельный промотор содержит последовательность EcRE или ее модифицированные варианты вместе с минимальным промотором. Модифицированными вариантами являются

последовательность распознавания E/GRE от Agilent Technologies, в которую были введены мутации. Последовательность распознавания E/GRE содержит инвертированные элементы распознавания половинных сайтов для ретиноид-Х-рецептора (RXR) и GR-связывающие домены. При всех пермутациях, экзогенно вводимым веществом является понастерон А, который отменяет репрессивное действие EcR или его производных на индуцибельный промотор и позволяет осуществлять транскрипцию.

Альтернативно, индуцибельные системы могут быть основаны на синтетическом стероиде мифепристоне, используемом в качестве экзогенно вводимого вещества. В этом случае вводят гибридный белок-регулятор транскрипции, который был получен на основе ДНК-связывающего домена дрожжевого белка GAL4, усеченного лиганд-связывающего домена (LBD) человеческого рецептора прогестерона и домена активации (AD) человеческого NF-κB. Этот гибридный белок-регулятор транскрипции поставляется от Thermofisher Scientific (Gene Switch™). Мифепристон активирует гибридный белок и инициирует транскрипцию из индуцибельного промотора, который содержит GAL4, расположенный выше активирующих последовательностей (UAS) и TATA-бокса аденовируса Elb. Эта система описана Wang, Y. et al (1994) Proc. Natl. Acad. Sci. USA 91, 8180-8184.

Таким образом, белком-регулятором транскрипции может быть любой подходящий белок-регулятор, либо белок-активатор, либо белок-репрессор. Подходящими белками-активаторами транскрипции являются белок-активатор транскрипции, реагирующий на тетрациклин (rtTa), или гибридный белок-регулятор транскрипции Gene Switch. Подходящими белками-репрессорами являются вариант Tet-Off rtTA, TetR или EcR. Белками-регуляторами транскрипции могут быть модифицированные белки или их производные, если это необходимо.

Индукцибельный промотор может содержать элементы, подходящие для связывания или взаимодействия с белком-регулятором транскрипции. Взаимодействие белка-регулятора транскрипции с индуцибельным промотором, предпочтительно регулируют с

использованием экзогенно вводимого вещества.

Экзогенно вводимым веществом может быть любое подходящее вещество, которое связывается или взаимодействует с белком-регулятором транскрипции. Подходящими веществами являются тетрациклин, понастерон А и мифепристон.

Таким образом, инсерция гена, кодирующего белок-регулятор транскрипции, в первый GSH обеспечивает механизм регуляции экспрессии индуцибельного кластера, который функционально присоединен к индуцибельному промотору и встроен во второй другой GSH-сайт.

Ген белка-регулятора транскрипции может быть взят для инсерции вместе с другим генетическим материалом. Такой материал включает гены маркерных или репортерных молекул, таких как гены, которые индуцируют визуально идентифицируемые признаки, включая флуоресцентные и люминесцентные белки. Примерами являются ген, который кодирует белок медузы, флуоресцирующий в зеленом диапазоне спектра (GFP), и который сообщает клеткам экспрессию, и вызывает зеленое свечение под действием синего/УФ-излучения люциферазы, катализирующей реакцию с люциферинем и испускающей свечение, а также белок, флуоресцирующий в красном диапазоне спектра и происходящий от гена DsRed. Такие маркерные или репортерные гены являются подходящими, поскольку присутствие репортерного белка подтверждает экспрессию белка от первого GSH, что указывает на успешную инсерцию. Селективные маркеры могут также включать гены резистентности к антибиотикам или к другим лекарственным средствам. Могут быть также введены и другие маркеры или репортерные генные последовательности, которые позволяют изучать экспрессию эндогенных (или экзогенных) генов. Такими белками являются белки Cas, включая CasL, белки Cas9, которые вырезают представляющие интерес гены, а также гибридные белки Cas, которые опосредуют изменение уровня экспрессии других генов, например, благодаря их действию как энхансеров или репрессоров транскрипции. Кроме того, может оказаться желательной не-индуцибельная экспрессия молекулярных систем, включая оптогенетические системы, гибридные белки ядерных рецепторов, такие как системы ERT, индуцируемые тамоксифеном, и

сконструированные рецепторы, активируемые исключительно специально созданными лекарственными средствами. Кроме того, последовательности, которые кодируют факторы передачи сигнала, изменяющие функцию той же самой клетки или соседних или даже отдаленных клеток в организме, включая аутокринные гормоны или паракринные факторы, могут быть коэкспрессированы из того же самого GSH в виде белка-регулятора транскрипции.

Кроме того, другой генетический материал может включать последовательности, кодирующие некодирующую РНК, как обсуждается в настоящей заявке. Примерами такого генетического материала являются гены миРНК, которые могут функционировать в качестве генетического переключателя.

Предпочтительно, чтобы ген, кодирующий белок-регулятор транскрипции, был функционально присоединен к конститутивному промотору. Альтернативно, первый GSH может быть выбран так, чтобы он уже имел конститутивный промотор, а также мог индуцировать экспрессию гена белка-регулятора транскрипции и любого ассоциированного с ним генетического материала. Конститутивные промоторы обеспечивают стабильную экспрессию гена на высоком уровне. Обычно используемыми конститутивными промоторами являются промоторы человеческого β -актина (ACTB), цитомегаловируса (CMV), фактора элонгации-1 α (EF1 α), фосфоглицераткиназы (PGK) и убихитина С (UBC). Промотором CAG является сильный синтетический промотор, часто используемый для инициации высоких уровней экспрессии генов, и этот промотор был сконструирован из следующих последовательностей: (C) раннего энхансерного элемента цитомегаловируса (CMV), (A) промотора, первого экзона и первого интрона гена куриного бета-актина, и (G) акцептора сплайсинга гена кроличьего бета-глобина.

Кроме того, регулятор транскрипции, а также любой другой генетический материал может быть взят вместе с расщепляемыми последовательностями. Такие последовательности представляют собой последовательности, которые распознаются молекулой, способной специфически разрезать ДНК, и включают рестрикционные сайты, которые являются последовательностями-мишенями для

рестриктирующих ферментов или последовательностями, распознаваемыми другими ДНК-расщепляющими молекулами, такими как нуклеазы, рекомбиназы, рибозимы или искусственные конструкции. Может быть включена по меньшей мере одна расщепляемая последовательность, но предпочтительно, чтобы присутствовали две или более последовательностей. Эти расщепляемые последовательности могут находиться в любом подходящем положении вставки, так, чтобы отобранная часть вставки или все вставки, могли быть селективно удалены из GSH. Таким образом, указанный способ может также быть применен для удаления вставки или ее части из GSH и/или для ее замены. Таким образом, расщепляемые сайты могут фланкировать часть вставки/всю вставку, которую желательно удалить. Регулятор транскрипции и/или другой генетический материал могут быть удалены с применением этого способа.

Часть вставки может представлять собой любую часть до 99% от вставки, то есть, 1-99%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10% или менее, чем 10%.

Может оказаться предпочтительным, чтобы часть вставки, фланкированная расщепляемыми сайтами, включала конститутивный промотор. Альтернативно, конститутивный промотор не входит в часть, фланкированную расщепляемыми последовательностями.

Предпочтительной расщепляемой последовательностью является сайт LoxP для рекомбиназы Cre, поскольку он позволяет осуществлять прямую замену удаленной вставки. Альтернативно или дополнительно, расщепляемой последовательностью является сайт rox для рекомбиназы Dre.

Предпочтительно, чтобы вставка в первом GSH находилась в обоих локусах генома, а поэтому каждый аллель модифицируют путем инсерции. Это приводит к повышению уровня экспрессии из гена, кодирующего регулятор транскрипции и любой ассоциированный с ним генетический материал.

Второй GSH может представлять собой любой подходящий GSH-сайт. Может оказаться предпочтительным, чтобы второй GSH-сайт не ассоциировался с эндогенным промотором, то есть, чтобы экспрессия встроенного индуцибельного кластера находилась под

контролем лишь белка-регулятора транскрипции.

Индукцибельный кластер включает нужную генную последовательность, а предпочтительно, последовательность ДНК, которую переносят в клетку. Введение индуцибельного кластера в геном может изменять фенотип этой клетки, либо в результате добавления генной последовательности, которая обеспечивает экспрессию гена, либо в результате нокдауна/нокаута эндогенной экспрессии. Способы согласно изобретению обеспечивают регулируемую транскрипцию генной(ых) последовательности(ей) в индуцибельном кластере в клетке.

Предпочтительно, чтобы нужная генная последовательность для вставки представляла собой последовательность ДНК, кодирующую молекулу РНК. Молекулой РНК может быть любая последовательность, а предпочтительно, кодирующая или не кодирующая РНК. Кодирующая или матричная РНК кодирует полипептидные последовательности, и транскрипция такой РНК приводит к экспрессии белка в клетке. Не кодирующая РНК может быть функциональной и может включать, но не ограничивается ими, микроРНК, короткую интерферирующую РНК, Рiwi-взаимодействующую РНК, антисмысловую РНК, малую ядерную РНК, малую ядрышковую РНК, малую РНК телец Каджала, Y-РНК, эхансерные РНК, руководящую РНК, рибозимы, короткую шпилечную РНК, малую временную РНК, транс-действующую РНК, короткую интерферирующую РНК и субгеномную матричную РНК. Не кодирующая РНК также известна как функциональная РНК. Некоторые типы РНК являются регуляторными по своей природе, и, например, могут ингибировать экспрессию гена, поскольку они являются комплементарными части мРНК или ДНК гена. МикроРНК (миРНК, 21-22 нуклеотидов) были обнаружены в эукариотах и действуют посредством РНК-интерференции (РНКи), где эффекторный комплекс миРНК и ферментов может расщеплять комплементарную мРНК, блокировать трансляцию мРНК или ускорять ее разложение. РНК другого типа, а именно, короткие интерферирующие РНК (киРНК, 20-25 нуклеотидов) действуют посредством РНК-интерференции по механизму, аналогичному механизму действия миРНК. Некоторые миРНКи и киРНК могут вызвать метилирование генов-мишеней, и тем самым, могут снижать или повышать уровень транскрипции этих

генов. У животных присутствуют Piwi-взаимодействующие РНК (пиРНК, 29–30 нуклеотидов), которые являются активными в клетках зародышевых линий и, вероятно, обеспечивают защиту от транспозонов. Многие прокариоты имеют РНК CRISPR, то есть, регуляторную систему, подобную РНК-интерференции, и такая система включает руководящую РНК (рРНК). Антисмысловые РНК широко распространены, и в большинстве случаев ингибируют ген, и лишь немногие из них являются активаторами транскрипции. Антисмысловая РНК может действовать посредством связывания с мРНК с образованием двухцепочечный РНК, которая подвергается ферментативному расщеплению. Существует множество длинных некодирующих РНК, которые регулируют гены в эукариотах, и одной из таких РНК является Xist, которая покрывает одну X-хромосому у самок млекопитающих и инактивирует ее. Таким образом, существует множество функциональных РНК, которые могут быть использованы в способах согласно изобретению.

Таким образом, индуцибельный кластер может включать генную последовательность, которая представляет собой ген, кодирующий белок. Этот ген, по своей природе, может не присутствовать в клетке, либо он может присутствовать в клетке, но для этого требуется его регулируемая экспрессия. Альтернативно, индуцибельный кластер может представлять собой мутированный, модифицированный или скорректированный вариант гена, присутствующего в клетке, а в частности, он может быть использован в генотерапии или для создания моделей заболеваний. Таким образом, индуцибельный кластер может включать трансен, происходящий от другого организма того же вида (то есть, поврежденный/мутированный вариант человеческого гена, или человеческий ген дикого типа) или от других видов.

В любом аспекте или варианте осуществления изобретения, генная последовательность, содержащаяся в индуцибельном кластере, может представлять собой синтетическую последовательность.

Индуцибельный кластер может включать любую подходящую генную последовательность, которую желательно встроить в геном клетки. Следовательно, генная последовательность может

представлять собой ген, кодирующий белковый продукт или последовательность, транскрибируемую в рибонуклеиновую кислоту (РНК), которая обладает определенной функцией (такую как, малая ядерная РНК (мяРНК), антисмысловая РНК, микроРНК (миРНК), короткая интерферирующая РНК (киРНК), транспортная РНК (тРНК) и другие некодирующие РНК (нкРНК), включая CRISPR-РНК (cr-РНК) и руководящую РНК (рРНК)).

Таким образом, индуцибельный кластер может включать любую генную последовательность, транскрипцию которой желательно регулировать в клетке. Выбор генной последовательности будет зависеть от типа клетки и от цели применения клетки после модификации, как подробно обсуждается ниже.

Так, например, для генотерапии может оказаться желательным получение последовательности гена дикого типа в виде компонента индуцибельного кластера. В этом случае, генной последовательностью может быть любой белок-кодирующий ген человека или животного. Примерами белок-кодирующих генов являются человеческий ген β -глобина, человеческий ген липопротеинлипазы (LPL), белок-шаперон Rab 1 у человека, кодируемый геном *CHM* и многие другие. Альтернативно, индуцибельный кластер может экспрессировать факторы роста, включая BDNF, GDF, NGF, IGF, FGF и/или ферменты, которые могут расщеплять про-пептиды с образованием активных форм. Генотерапия может быть также осуществлена посредством экспрессии индуцибельного кластера, содержащего генную последовательность, кодирующую антисмысловую РНК, миРНК, киРНК или РНК любого типа, которые блокируют экспрессию другого гена в клетке.

Альтернативно, клеткой может быть стволовая клетка, где индуцибельный кластер может включать генную последовательность, кодирующую основной регулятор, специфичный к ключевой линии дифференцировки, и такой регулятор сокращенно называется здесь главным регулятором. Главными регуляторами могут быть один или более из следующих факторов, таких как факторы транскрипции, регуляторы транскрипции, рецепторы цитокинов или сигнал-передающие молекулы и т.п. Главный регулятор представляет собой

экспрессируемый ген, который негативно влияет на дифференцировку клеток, экспрессирующих этот ген. Этот регулятор может представлять собой сеть главных регуляторов, необходимых для определения линий дифференцировки клеток. Используемый здесь ген главного регулятора, который экспрессируется в начале периода развития линии дифференцировки или клеток определенного типа, участвует в распознавании этой линии дифференцировки посредством регуляции множества генов, расположенных непосредственно ниже, или посредством изменения экспрессии каскада генов. При экспрессии главного регулятора, он приобретает способность переопределить статус клеток с образованием других линий дифференцировки. Примерами основных регуляторов являются миогенный фактор транскрипции MyoD и гемопоэтический фактор транскрипции SCL. В частности, главными регуляторами являются, но не ограничиваются ими:

Линии дифференцировки нервных клеток: олигодендроциты: SOX10, OLIG2, NKX2.2., NKX6.2; астроциты: NFIA, NFIB и SOX9; нейроны: Ascl1, нейрогенин, и NeuroD, Pax6, Neurog2, Ascl1, Dlx2, и NeuroD1; гемопоэтические клетки, включая эритроциты и мегакариоциты: GATA1, FLI1 и TAL1.

Мезенхимальные линии дифференцировки: скелетные мышцы: MYOD; кардиомиоциты: Gata4, Mef2c, Baf60c и Tbx5; кость: L-Мус (RXOL) Runx2, Osterix, Oct4; хрящ: с-Мус Klf4, SOX9; и коричневые адипоциты: C/EBP- β и с-Мус

Эндодерма

Типы панкреатических клеток: PDX1 и GATA6.

Стволовые клетки: эпибластные SC: Oct4, Sox2, Klf4 и с-Мус.

Альтернативно или дополнительно, генная последовательность или другой генетический материал могут представлять собой гены, функции которых требует исследования для определения, может ли регулируемая экспрессия влиять на эффект экспрессии в клетке; может ли ген включать гены факторов роста и/или цитокинов в клетках, используемых для трансплантации; и/или может ли ген быть компонентами анализа с использованием репортера.

Кроме того, генная последовательность может кодировать

некодирующую РНК, функция которой заключается в нокдауне экспрессии эндогенного гена или последовательности ДНК, кодирующей некодирующую РНК в клетке. Альтернативно, генная последовательность может кодировать руководящую РНК системы CRISPR-cas9 и осуществлять нокаут эндогенного гена.

Таким образом, способы согласно изобретению охватывают способы нокдауна экспрессии эндогенного гена в клетке. Эти способы являются такими, как они были описаны ранее, и в этих способах, индуцибельный кластер содержит генную последовательность, кодирующую некодирующую РНК, функционально присоединенную к индуцибельному промотору, где некодирующая РНК подавляет экспрессию указанного эндогенного гена. Некодирующая РНК может подавлять экспрессию генов посредством любого подходящего механизма, включая интерференцию РНК и антисмысловую РНК. Таким образом, генная последовательность может кодировать кшРНК, которая может негативно влиять на матричную РНК, кодируемую эндогенным геном.

Снижение уровня экспрессии эндогенного гена может быть частичным или полным, то есть, экспрессия может снижаться на 50, 55, 65, 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 или 100% по сравнению с экспрессией в клетке до индуцирования транскрипции некодирующей РНК.

Способы согласно изобретению также охватывают способы нокаута эндогенных генов в клетке благодаря системе CRISPR-cas9, хотя могут быть использованы и любые другие подходящие системы нокаута гена. В этом случае, предпочтительно, чтобы гены cas9 конститутивно экспрессировались и, таким образом, встраивались в первый GSH с геном регулятора транскрипции. Генные последовательности, кодирующие рРНК, могут быть включены в индуцибельный кластер, который встраивают во второй GSH. рРНК представляет собой короткую синтетическую РНК, состоящую из каркасной последовательности, необходимой для связывания с Cas9, и таргетирующей последовательности приблизительно из 20 нуклеотидов, которая определяет степень модификации геномной мишени. Таким образом, геномная мишень cas9 может быть модифицирована путем простого изменения таргетирующей

последовательности, присутствующей в рРНК. Хотя основное применение такой системы заключается в конструировании рРНК для нацеливания на эндогенный ген с последующим его нокаутом, однако, эта система может быть также модифицирована для селективной активации или репрессии генов-мишеней, очистки специфических областей ДНК и даже визуализации ДНК. При этом рассматриваются все возможные применения.

Индуцибельный кластер включает генную последовательность, функционально присоединенную к индуцибельному промотору. «Промотор» представляет собой нуклеотидную последовательность, которая инициирует и регулирует транскрипцию полинуклеотида. «Индукцибельный промотор» представляет собой нуклеотидную последовательность, где экспрессия генной последовательности, функционально присоединенной к промотору, регулируется аналитом, кофактором, регуляторным белком и т.п. В соответствии с настоящим изобретением, такая регуляция достигается с использованием белка-регулятора транскрипции. Следует отметить, что термин «промотор» или «регуляторный элемент» включает полноразмерные промоторные области и функциональные (например, регулирующие транскрипцию или трансляцию) сегменты этих областей. Термин «функционально присоединенный» относится к расположению элементов, где описанные компоненты имеют такую конфигурацию, которая позволяет им выполнять свою обычную функцию. Таким образом, данный промотор, функционально присоединенный к генной последовательности, способен осуществлять экспрессию этой последовательности, если присутствуют соответствующие ферменты. Промотор, необязательно должен быть смежным с последовательностью, при условии, что его функция будет заключаться в регуляции экспрессии. Так, например, промежуточные нетранслированные, но транскрибируемые последовательности могут присутствовать между промоторной последовательностью и генной последовательностью, и такая промоторная последовательность может рассматриваться как последовательность, «функционально присоединенная» к генной последовательности. Таким образом, термин «функционально присоединенный» охватывает любую пространственную или обычную ориентацию промоторного элемента и

генной последовательности в индуцибельном кластере, что позволяет инициировать транскрипцию индуцибельного кластера после распознавания промоторного элемента транскрипционным комплексом.

Кроме того, другой генетический материал может быть также функционально присоединен к индуцибельному промотору. Другой генетический материал может включать гены, кодирующие последовательности РНК, генетический материал, такой как маркерные или репортерные гены. Такой дополнительный генетический материал обсуждался ранее. В некоторых случаях может быть желательным включение в индуцибельный кластер гена-«самоубийцы», поскольку сама генная последовательность не может представлять собой ген-«самоубийцу», используемый для противораковой генотерапии. Ген-«самоубийца» может использовать один и тот же индуцибельный промотор в индуцибельном кластере, либо он может представлять собой отдельный индуцибельный промотор для обеспечения отдельной регуляции. Такой ген может быть полезен в стратегиях генотерапии, когда желательно, чтобы этот ген разрушал донорные/трансфицированные клетки при выполнении определенных условий. Генами-«самоубийцами» являются гены, которые экспрессируют белок, способствующий клеточному апоптозу, или, альтернативно, эти гены могут потребовать присутствия экзогенного кофактора или фактора, действующего совместно с лекарственным средством. Кофактор или фактор, действующий совместно с лекарственным средством, может превращаться под действием продукта гена-«самоубийцы» в высокой степени цитотоксическую молекулу.

Кроме того, индуцибельный кластер может включать расщепляемые последовательности. Такие последовательности представляют собой последовательности, которые распознаются молекулой, способной специфически разрезать ДНК, и включают рестрикционные сайты, которые являются последовательностями-мишенями для рестриктирующих ферментов или последовательностями, распознаваемыми другими ДНК-расщепляющими молекулами, такими как нуклеазы, рекомбиназы, рибозимы или искусственные конструкции. Может быть включена по меньшей мере одна расщепляемая

последовательность, но предпочтительно, чтобы присутствовали две или более последовательности. Эти расщепляемые последовательности могут находиться в любом подходящем положении вставки, так, чтобы отобранная часть кластера или весь кластер, могли быть селективно удалены из GSH. Таким образом, указанный способ может быть также применен для удаления кластера или его части из GSH и/или для его замены. Таким образом, расщепляемые сайты могут фланкировать часть генной последовательности/всю генную последовательность, которую желательно удалить. Этот способ позволяет удалять индуцибельный кластер и/или другой генетический материал.

Часть кластера может представлять собой любую часть до 99% от кластера, то есть, 1-99%, 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, 10% или менее, чем 10%.

Может оказаться предпочтительным, чтобы часть вставки, фланкированная расщепляемыми сайтами, включала промотор, функционально присоединенный к генной последовательности. Альтернативно, промотор, функционально присоединенный к генной последовательности, не входят в часть, фланкированную расщепляемыми последовательностями.

Предпочтительной расщепляемой последовательностью является сайт LoxP для рекомбиназы Cre, поскольку он позволяет осуществлять прямую замену удаленной вставки. Альтернативно или дополнительно, расщепляемой последовательностью может быть сайт rox для рекомбиназы Dre.

Белок-регулятор транскрипции и индуцибельный кластер вместе с любым ассоциированным с ними генетическим материалом, встраивают в другой GSH в геноме клетки.

Вставки в GSH, предпочтительно специфически осуществляют в последовательность GSH, как описано ранее. Может быть применен любой подходящий способ введения полинуклеотида в специфическую последовательность, и некоторые из них описаны в литературе. Подходящие способы включают любой способ, который вводит разрыв в нужное положение и позволяет осуществлять рекомбинацию вектора в пропуске. Таким образом, первой важной стадией нацеленной сайт-специфической геномной модификации является создание

двухцепочечного ДНК-разрыва (DSB) в модифицируемом геномном локусе. Различные механизмы репарации клеток могут быть применены для репарации DSB и введения нужной последовательности, и такими механизмами являются репарация посредством присоединения негомологичных концов (NHEJ), которые более предрасположены к ошибкам, и репарация с гомологичной рекомбинацией (HR), опосредуемая донорной ДНК-матрицей, и такие механизмы могут быть использованы для встраивания индуцибельных кластеров.

Существует несколько методов специально разработанной сайт-специфической генерации DSB в геноме. Многие из них включают использование специализированных эндонуклеаз, таких как нуклеазы «цинковые пальцы» (ZFN), эффекторные нуклеазы, подобные активатору транскрипции (TALEN), или система кластеризованных регулярно повторяющихся коротких палиндромных повторов/белка, ассоциированного с CRISPR (CRISPR/Cas9) (Gaj, T, et al «ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering» Trends Biotechnol, 31:397-405, July 2013).

Нуклеазы «цинковые пальцы» представляют собой искусственные ферменты, которые получают путем присоединения домена, связывающегося с ДНК «цинковые пальцы», к нуклеазному домену рестриктирующего фермента FokI. Последний имеет неспецифический домен расщепления, который должен димеризоваться и расщеплять ДНК. Это означает, что два мономера ZFN необходимы для димеризации доменов FokI и расщепления ДНК. ДНК-связывающий домен может быть сконструирован так, чтобы он мог быть нацелен на любую представляющую интерес геномную последовательность, представляет собой tandemный массив «цинковых пальцев» Cys₂His₂, каждый из которых распознает три смежных нуклеотида в последовательности-мишени. Два связывающих сайта отделены друг от друга 5-7 п.о., что позволяет осуществлять оптимальную димеризацию доменов FokI. Таким образом, этот фермент способен расщеплять ДНК в специфическом сайте, а его специфичность к мишени повышается благодаря тому, что два события связывания с проксимальной ДНК должны обеспечивать образование двухцепочечного разрыва.

Эффекторные нуклеазы, подобные активатору транскрипции, или TALEN, представляют собой комплекс «димерный фактор транскрипции/нуклеаза». Эти нуклеазы получают путем присоединения ДНК-связывающего домена эффектора TAL к ДНК-расщепляющему домену (нуклеазы). Эффекторные нуклеазы, подобные активатору транскрипции (TALE), могут быть сконструированы так, чтобы они связывались практически с любой нужной последовательностью ДНК, и так, чтобы при их объединении с нуклеазой, ДНК могла разрезаться в специфических сайтах. Эффекторы TAL представляют собой белки, которые секретируются бактериями *Xanthomonas*, ДНК-связывающий домен которых содержит повторяющуюся высококонсервативную аминокислотную последовательность из 33–34 аминокислот с модификациями 12-й и 13-й аминокислоты. Эти два положения являются в высокой степени переменными и в значительной степени коррелируют со специфическим распознаванием нуклеотида. Такая прямая взаимосвязь между аминокислотной последовательностью и распознаванием ДНК позволяет сконструировать специфические ДНК-связывающие домены посредством отбора комбинации повторяющихся сегментов, содержащих соответствующие остатки в двух различных положениях. Таким образом, TALEN конструируют из массивов модулей, состоящих из 33–35 аминокислот, каждый из которых нацелен на один нуклеотид. Посредством отбора массива модулей можно осуществлять нацеливание почти на любую последовательность. И в этом случае, используемой нуклеазой могут быть FokI или его производное.

Были идентифицированы три типа механизмов CRISPR, из которых тип II является наиболее изученным. Система CRISPR/Cas9 (тип II) использует нуклеазу Cas9 для создания двухцепочечного разрыва в ДНК в сайте, распознаваемом короткой руководящей РНК. Система CRISPR/Cas является прокариотической иммунной системой, сообщающей резистентность к чужеродным генетическим элементам. CRISPR представляют собой сегменты прокариотической ДНК, содержащей короткие повторы последовательностей оснований. За каждым повтором расположены короткие сегменты «протоспейсерной» ДНК, образованные в результате предыдущего экспонирования

чужеродными генетическими элементами. Спейсеры CRISPR распознают и разрезают экзогенные генетические элементы посредством РНК-интерференции. Иммунный ответ CRISPR происходит в две стадии: биогенеза CRISPR-РНК (cr-РНК) и интерференции посредством руководящей cr-РНК. Молекулы cr-РНК состоят из вариабельной последовательности, транскрибируемой из протоспейсерной ДНК и CRISPR-повтора. Каждая молекула cr-РНК затем гибридизуется со второй РНК, известной как транс-активирующая CRISPR-РНК (tracr-РНК), и вместе они, в конечном счете, образуют комплекс с нуклеазой Cas9. Протоспейсерная ДНК кодирует часть cr-РНК, регулирующую Cas9, что приводит к расщеплению комплементарных последовательностей ДНК-мишеней, если они расположены поблизости от коротких последовательностей, известных как мотивы, смежные с протоспейсером (PAM). Эта природная система была разработана и использована для введения DSB-разрывов в специфические сайты геномной ДНК, хотя она может быть применена и во многих других целях. В частности, может быть использована система CRISPR типа II от *Streptococcus pyogenes*. В самом простом варианте, система CRISPR/Cas9 содержит два компонента, которые доставляются в клетку для редактирования генома: сама нуклеаза Cas9 и небольшая руководящая РНК (gРНК). gРНК представляет собой гибрид специально созданной сайт-специфической cr-РНК (нацеленной на последовательность-мишень) и стандартизированной tracr-РНК.

После получения DSB вводят донорную матрицу, гомологичную локусу-мишени; и репарация DSB может быть осуществлена по пути гомологично ориентированной репарации (HDR), позволяющему точно вводить вставки.

Могут быть также использованы и производные этой системы. Мутантные формы Cas9, такие как Cas9D10A, являются доступными только с никазной активностью. Это означает, что они расщепляют только одну цепь ДНК и не активируют NHEJ. Вместо этого, при введении матрицы гомологичной репарации, репарацию ДНК проводят только по пути HDR с высокой степенью достоверности. Cas9D10A (Cong L., et al. (2013) Science, 339, 819-823) может быть использован в парных комплексах Cas9, сконструированных для создания смежных одноцепочечных ДНК-разрывов в комбинации с

двумя оцрРНК, комплементарных смежной области на противоположных цепях сайта-мишени, которые могут быть особенно предпочтительными.

Элементы для создания двухцепочечного ДНК-разрыва могут быть введены в один или более векторов, таких как плазмиды для экспрессии в клетке.

Таким образом, в способе согласно изобретению может быть применен любой метод создания специфических двухцепочечных разрывов-мишеней в геноме для встраивания гена/индуцибельного кластера. Может оказаться предпочтительным, чтобы в способе встраивания гена/индуцибельного кластера использовались любые один или более ZFN, TALEN и/или систем CRISPR/Cas9 или любые их производные.

После получения DSB любым подходящим методом, ген/индуцибельный кластер для инсерции может быть введен любым подходящим способом, как описано ниже. Ген/индуцибельный кластер и ассоциированный с ним генетический материал образуют донорную ДНК для репарации ДНК в DSB, и их встраивают по стандартным механизмам/путям репарации клеток. Способ инициации разрыва с последующей модификацией пути репарации повреждения описан выше.

Белок-регулятор транскрипции и индуцибельный кластер могут быть введены способом согласно изобретению в отдельные векторы. «Вектор» представляет собой молекулу нуклеиновой кислоты, такую как молекула ДНК, которая используется в качестве носителя для искусственного введения генетического материала в клетку. Вектор обычно представляет собой последовательность нуклеиновой кислоты, которая состоит из вставки (такой как индуцибельный кластер или ген белка-регулятора транскрипции) и более крупной последовательности, которая служит в качестве «остова» вектора. Вектор может быть получен в любой подходящей форме, включая плазмиды, небольшую кольцевую или линейную ДНК. Вектор содержит по меньшей мере ген регулятора транскрипции или индуцибельного кластера, функционально присоединенного к индуцибельному промотору, вместе с минимальными последовательностями, позволяющими осуществлять встраивание генов в релевантный GSH. Векторы также, но необязательно, имеют точку начала репликации

(ori), который позволяет амплифицировать вектор, например, в бактериях. Дополнительно или альтернативно, вектор включает селективные маркеры, такие как гены резистентности к антибиотикам, гены цветковых маркеров и гены-«самоубийцы».

Примеры векторов, используемых в Примерах, описаны на фигурах 20-33.

Клетками, используемыми в способе согласно изобретению, могут быть любые клетки человека или животных. Предпочтительными являются клетки млекопитающих, например, клетки грызунов, таких как мыши и крысы; сумчатых, таких как кенгуру и коалы; приматов, не являющихся человеком, таких как шимпанзе бонобо, шимпанзе, лемуры, гиббоны и человекообразные обезьяны; животных семейства верблюжьих, таких как верблюды и ламы; крупного рогатого скота, такого как лошади, свиньи, крупный рогатый скот, буйволы, бизоны, козы, овцы, олени, северные олени, ослы, бантенги, яки, куры, утки и индюки; домашних животных, таких как кошки, собаки, кролики и морские свинки. Предпочтительной клеткой является человеческая клетка. В некоторых аспектах изобретения, клеткой предпочтительно является клетка крупного рогатого скота.

Тип клеток, используемых в способе согласно изобретению, будет зависеть от цели применения клеток после завершения встраивания генетического материала в GSH-сайты.

Если целью является продуцирование зрелых клеток из клеток-предшественников, то модифицированной клеткой является стволовая клетка, а предпочтительно плюрипотентная стволовая клетка. Плюрипотентные стволовые клетки могут дифференцироваться почти во все клетки организма. Существует несколько источников плюрипотентных стволовых клеток. Эмбриональные стволовые клетки (ES-клетки) являются плюрипотентными стволовыми клетками, происходящими от внутренней клеточной массы бластоцистов эмбриона на ранней стадии перед имплантацией. Индуцированные плюрипотентные стволовые клетки (iPSC) представляют собой зрелые клетки, которые были генетически перепрограммированы в состояние, подобное эмбриональным стволовым клеткам, путем принудительной экспрессии генов и факторов, играющих важную роль в сохранении определенных свойств эмбриональных стволовых

клеток. В 2006 году было показано, что введение четырех специфических генов, кодирующих факторы транскрипции, может превращать зрелые клетки в плюрипотентные стволовые клетки (Takahashi, K; Yamanaka, S (2006), Cell 126 (4): 663–76), но в последующем, ряд необходимых генов был уменьшен/модифицирован. Oct-3/4 и некоторые члены семейства генов Sox были идентифицированы как потенциально важные регуляторы транскрипции, участвующие в процессе индуцирования. Дополнительные гены, включая некоторые члены семейства Klf, семейства Мус, Nanog и LIN28, могут повышать эффективность индуцирования. Примерами генов, которые могут содержаться в факторах перепрограммирования, являются Oct3/4, Sox2, Sox1, Sox3, Sox15, Sox17, Klf4, Klf2, с-Мус, N-Мус, L-Мус, Nanog, Lin28, Fbx15, ERas, ECAT15-2, Tcl1, бета-катенин, Lin28b, Sall1, Sall4, Esrrb, Nr5a2, Tbx3 и Glis1, и эти факторы перепрограммирования могут быть использованы по отдельности или в комбинации двух или более видов.

Если целью последующих исследований является продуцирование стволовых клеток с нокдауном или нокаутом гена, например, в исследованиях развития или в исследованиях генных функций, то модифицированной клеткой может быть стволовая клетка, а предпочтительно, плюрипотентная стволовая клетка или зрелая клетка. Источники плюрипотентных стволовых клеток обсуждаются выше.

Если клетки, модифицированные путем инсерции индуцибельного кластера, используются для введения человеку, то может оказаться предпочтительным, чтобы клетка представляла собой iPSC, взятая у этого индивидуума. Такое применение аутологичных клеток позволяет избежать необходимости поиска соответствующих клеток для реципиента. Альтернативно, могут быть использованы коммерчески доступные iPSC, например, клетки, поставляемые WiCell® (WiCell Research Institute, Inc, Wisconsin, US). Альтернативно, клетками могут быть тканеспецифические стволовые клетки, которые могут быть также аутологичными или взятыми у донора. Подходящими клетками являются эпибластные стволовые

клетки, индуцированные нервные стволовые клетки и другие тканеспецифические стволовые клетки.

В некоторых вариантах осуществления изобретения может оказаться предпочтительным, чтобы используемой клеткой была эмбриональная стволовая клетка или стволовая клеточная линия. В настоящее время доступно множество эмбриональных стволовых клеточных линий, например, WA01 (HI) и WA09 (H9) могут быть получены от WiCell, а KhES-1, KhES-2 и KhES-3 могут быть получены из Института Медицинских наук, Университета Киото (Киото, Япония).

Может оказаться предпочтительным, чтобы эмбриональная стволовая клетка была получена без разрушения эмбриона, а в частности, в том случае, если клетки представляют собой человеческие клетки, поскольку такие методы являются легко доступными (Chung, Young et al., Cell Stem Cell, Volume 2, Issue 2, 113-117). Также доступными являются стволовые клеточные линии, которые были получены без разрушения эмбриона. В одном из своих аспектов, настоящее изобретение не относится к любым способам, которые включают разрушение человеческого эмбриона.

Предпочтительный аспект настоящего изобретения представляет собой прямое перепрограммирование плюрипотентных стволовых клеток в зрелые клетки. Таким образом, способ согласно изобретению может быть применен для получения зрелых клеток из плюрипотентных стволовых клеток. В этом аспекте изобретения, индуцибельным кластером для встраивания во второй GSH предпочтительно является один или более главных регуляторов, обсуждаемых ранее. Эти индуцибельные кластеры могут способствовать превращению клеток в конкретную линию дифференцировки, и различные индуцибельные кластеры могут быть использованы для прямой дифференцировки в зрелые клетки. При этом рассматриваются зрелые клетки любого типа, включая, но не ограничиваясь ими, нервные клетки, миоциты, остециты, хондроциты, эпителиальные клетки, секреторные клетки и/или клетки крови.

Авторами настоящей заявки был разработали быстрый, эффективный и масштабируемый способ получения зрелых клеток

фактически любого типа. Такой простой и дешевый метод будет иметь особую ценность для регенеративной медицины. В предыдущих методах прямого перепрограммирования клеток используются системы Tet-On, но были предприняты попытки включить весь этот материал в один вектор/сайт («Tet-On все-в-одном») или встроить индуцибельный кластер в один аллель AAVS1, а регуляторную систему в другой аллель AAVS1 (DeKelver *et al*, 2010, *Genome Res.*, 20, 1133-43 и Qian *et al*, 2014, *Stem Cells*, 32, 1230-8). Удивительно то, что этот разработанный и описанный здесь способ нацеливания на два GSH имеет много непредвиденных преимуществ. Во-первых, отсутствует какая-либо возможная интерференция промотора между геном, встроенным в первый GSH, и генной последовательностью индуцибельного кластера, встроенного во второй GSH. Во-вторых, это позволяет осуществлять встраивание больших «грузов» из векторов, поскольку в этом случае, требуется меньше материала для встраивания в каждый сайт. В-третьих, этот способ максимизирует число безопасно встроенных копий. В-четвертых, это обеспечивает большую гибкость конструкции. И наконец, это позволяет встраивать дополнительный генетический материал, включая репортерные гены и переключатели миРНК. Способ согласно изобретению был продемонстрирован как надежный и эффективный способ получения зрелых клеток из плюрипотентных клеток.

После встраивания гена в первый GSH и встраивания индуцибельного кластера, содержащего трансген, во второй GSH, плюрипотентные стволовые клетки могут быть культивированы для их прямого перепрограммирования. Эти условия культивирования могут быть специфическими для используемых плюрипотентных стволовых клеток любого типа, либо они могут зависеть от зрелых клеток конечного типа. Независимо от используемых условий культивирования, экзогенное вещество будет регулировать экспрессию генной последовательности в индуцибельном кластере; и может подаваться непрерывно, а затем оно может быть удалено для индуцирования транскрипции или подано по мере потребности в транскрипции, в зависимости от способа его действия, как обсуждалось ранее.

Если целью является перепрограммирование стволовых клеток, то может оказаться предпочтительным, чтобы клетка имела внеклеточные стимулы для облегчения дифференцировки в комбинации с подачей индуцибельных кластеров, кодирующих главные регуляторы. Стратегии перепрограммирования клеток могут быть усовершенствованы путем объединения сверхэкспрессии основного регулятора или фактора транскрипции с внеклеточными сигнал-передающими молекулами. Таким образом, может быть осуществлен систематический скрининг на факторы про-дифференцировки путем модуляции основных каскадов передачи сигнала, участвующих в развитии зрелых клеток этого конкретного типа. Такой скрининг описан, например, в Примере 3.

В одном из своих аспектов, настоящее изобретение относится к способу продуцирования миоцитов из плюрипотентных стволовых клеток, где указанный способ включает стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

б) нацеленного встраивания гена MYOD1, функционально присоединенного к индуцибельному промотору, во второй генетически безопасный якорный сайт, где указанный индуцибельный промотор регулируется белком-регулятором транскрипции, и где указанные первый и второй генетически безопасные якорные сайты являются различными; и

культивирования указанных клеток в присутствии ретиноевой кислоты.

Ген MYOD1 представляет собой ген, кодирующий миогенный белок дифференцировки 1. Предпочтительной ретиноевой кислотой (RA) являются все транс-RA.

В другом своем аспекте, настоящее изобретение относится к способу продуцирования миоцитов из плюрипотентных стволовых клеток, экспрессирующих MYOD1, где указанный способ включает культивирование указанных клеток в присутствии ретиноевой кислоты.

Предпочтительно, RA представляет собой все транс-RA. Предпочтительно, клетки сверхэкспрессируют MYOD1.

В другом своем аспекте, настоящее изобретение относится к способу продуцирования олигодендроцитов из плюрипотентных стволовых клеток, где указанный способ включает стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

б) нацеленного встраивания гена SOX 10, функционально присоединенного к индуцибельному промотору, во второй генетически безопасный якорный сайт, где указанный индуцибельный промотор регулируется белком-регулятором транскрипции; и где указанные первый и второй генетически безопасные якорные сайты являются различными, и

культивирования указанных клеток в присутствии ретиноевой кислоты.

Клетками, используемыми для этой цели, могут быть клетки животного или человека. Если клетками являются клетки животного, то предпочтительно, чтобы такое животное принадлежало к крупному рогатому скоту, как было определено ранее.

Ген SOX-10 кодирует фактор транскрипции SOX-10. Предпочтительной ретиноевой кислотой (RA) являются все транс-RA.

Если клеткой, используемой в способах согласно изобретению, является плюрипотентная клетка, то полученная клетка может представлять стволовую клетку, специфичную к линии дифференцировки, клетку-предшественника или зрелую клетку с нужными свойствами, сообщаемыми экспрессией главного регулятора. Эти стволовые клетки, специфичные к линии дифференцировки, клетки-предшественники или зрелые клетки могут быть использованы в любом подходящем способе. Так, например, зрелые клетки могут быть использованы непосредственно для трансплантации в организм человека или животного, в зависимости от типа клетки. Альтернативно, клетки могут представлять собой тест-материал для исследований, включая влияние лекарственных средств на экспрессию гена и взаимодействие лекарственных средств с конкретным геном. Клетки для исследования могут включать индуцибельный кластер с генной последовательностью, имеющей неизвестную функцию, для исследования регулируемой экспрессии

этой генной последовательности. Кроме того, эти клетки могут быть использованы для получения больших количеств желаемых веществ, таких как факторы роста или цитокины.

В другом аспекте, клетки могут быть использованы для конструирования ткани. Для конструирования ткани требуется продуцирование ткани, которая может быть использована для замены тканей или даже целых органов человека или животного. Методы конструирования ткани известны специалистам в данной области и включают использование каркаса (внеклеточного матрикса), на который наносят клетки для получения тканей/органов. Эти методы могут быть применены для создания «искусственных» органов, таких как трахея, мочевого пузыря, печень, поджелудочная железа, желудок, тонкий кишечник, кровеносные сосуды, ткань сердца, кости, костный мозг, ткань слизистой, нервы, мышцы, кожа, почки или любые другие ткани или органы. Методы получения тканей могут включать дополнительную процедуру, также известную как трехмерную (3D) печать, которая может включать прямую печать клеток с получением тканей. Таким образом, настоящее изобретение относится к способу получения тканей с использованием клеток, полученных как описано в любом аспекте настоящего изобретения.

Ткани, генерированные с использованием клеток, полученных способами согласно изобретению, могут быть использованы для их трансплантации в организм человека или животного. Альтернативно, если клетки происходят от животного, то ткани могут быть использованы для *in vitro*/выращивание мяса. Первичными клетками для культивирования мяса являются миоциты. Однако, такая ткань может включать комбинацию клеток, полученных способами согласно изобретению. Такими клетками могут быть миоциты (мышечные клетки), клетки кровеносных сосудов, клетки крови и адипоциты (жировые клетки). Если целью конструирования ткани является выращивание мяса, то такая клетка может быть взята у крупного рогатого скота.

Способы согласно изобретению могут быть также осуществлены на клетках, которые не являются плюрипотентными стволовыми клетками, по разным причинам, в том числе для исследований, для генотерапии с использованием генетических вакцин, для получения

моделей заболеваний *in vitro* и создания не-человеческих моделей *in vivo*.

Клетки, используемые в способе согласно изобретению, могут представлять зрелые стволовые клетки любого типа, и такими клетками являются неспециализированные клетки, которые могут развиваться в клетки многих, но не всех типов. Зрелые стволовые клетки представляют собой недифференцированные клетки, присутствующие во всем организме, которые делятся, для пополнения погибших клеток и регенерации поврежденных тканей. Эти клетки также известны как соматические стволовые клетки, то есть, они не являются плюрипотентными. Зрелые стволовые клетки были выявлены во многих органах и тканях, включая головной мозг, костный мозг, периферическую кровь, кровеносные сосуды, скелетную мышцу, кожу, зубы, сердце, кишечник, печень, эпителий яичника и яички. Для того, чтобы установить, что клетка является соматической стволовой клеткой, специалист должен продемонстрировать, что одна зрелая стволовая клетка может генерировать линию генетически идентичных клеток, которые затем дают начало развитию всех соответствующих дифференцированных клеток ткани. Для того, чтобы экспериментально подтвердить, что предполагаемые зрелые стволовые клетки, действительно являются стволовыми клетками, эти клетки должны продуцировать эти генетически идентичные клеткам в культуре, или очищенная популяция этих клеток должна повторно заселять ткань после трансплантации животному. Подходящими типами клеток являются, но не ограничиваются ими, нервные, мезенхимальные и эндодермальные стволовые клетки и клетки-предшественники.

Альтернативно, используемыми клетками могут быть зрелые клетки. Такие клетки являются дифференцированными и специализированными и не способны развиваться в клетки другого типа. Зрелыми клетками являются, но не ограничиваются ими, нервные клетки, миоциты, остециты, хондроциты, эпителиальные клетки, секреторные клетки и/или клетки крови. Зрелые клетки могут представлять собой любые клетки, выделенные из организма человека или животного.

Соматические стволовые клетки и зрелые клетки могут быть

модифицированы в соответствии с настоящим изобретением, а затем использованы для таких целей, как генотерапия или генетическая вакцинация. Генотерапия может быть определена как нацеленное встраивание чужеродной ДНК в ядро клетки в терапевтических целях. Такое определение включает введение гена или генов в клетку для получения варианта дефектного гена дикого типа, добавление генов, кодирующих молекулы РНК, которые препятствуют экспрессии гена-мишени (которые могут быть дефектными), введение генов-«самоубиц» (таких как ферменты тимидинкиназа вируса простого герпеса (HSV-tk) и цитозиндезаминаза (CD), которые превращают безопасное пролекарство ганцикловир (GCV) в цитотоксическое лекарственное средство), приготовление ДНК-вакцин для иммунизации или противораковой терапии (включая клеточную адаптивную иммунотерапию) и введение любых других генов в клетку в терапевтических целях.

Обычно, способ согласно изобретению может быть применен для встраивания нужной генной последовательности для транскрипции в клетке, предпочтительно для экспрессии, а в частности, для приготовления ДНК-вакцин. ДНК-вакцины обычно кодируют модифицированную форму ДНК инфицированного организма. ДНК-вакцины вводят индивидууму, после чего они экспрессируют нужный белок инфицированного организма, инициируя иммунный ответ против этого белка, который обычно является протективным. ДНК-вакцины могут также кодировать опухолевый антиген при его применения в противораковой иммунотерапии.

ДНК-вакцина может включать последовательность нуклеиновой кислоты, кодирующую антиген, и используется для лечения или профилактики ряда состояний, включая, но не ограничиваясь ими, рак, аллергию, токсичность и инфицирование патогеном, таким как, но не ограничиваясь ими, грибки, вирусы, включая вирусы человеческой папилломы (HPV), ВИЧ-инфекции, HSV2/HSV1, вирус гриппа (типы А, В и С), поливирус, вирус RSV, риновирусы, ротавирусы, вирус гепатита А, вирус кори, вирус парагриппа, вирус паротита, вирус ветряной оспы, цитомегаловирус, вирус Эпштейна-Барра, аденовирусы, вирус коревой краснухи, вирус человеческой Т-клеточной лимфомы типа I (HTLV-I), вирус гепатита

В (HBV), вирус гепатита С (HCV), вирус гепатита D, вирус оспы, вирус Зика, Марбург и Эбола; бактерии, включая менингококки, *Haemophilus influenza* (типа В); и паразитарные патогены. ДНК-вакцины могут включать последовательность нуклеиновой кислоты, кодирующую антиген любого подходящего патогена. Антиген может происходить от патогена, ответственного за развитие заболевания у человека или животного, а в частности, он может происходить от вирусного патогена.

ДНК-вакцины, введенные в GSH, могут также включать последовательность нуклеиновой кислоты, кодирующую опухолевые антигены. Примерами опухолеассоциированных антигенов являются, но не ограничиваются ими, раковые антигены, такие как члены семейства MAGE (MAGE 1, 2, 3 и т.п.), NY-ESO-1 и SSX-2, антигены дифференцировки, такие как антигены тирозиназы, gp100, PSA, Her-2 и CEA, мутированные аутоантигены и вирусные опухолевые антигены, такие как Е6 и/или Е7, происходящие от HPV онкогенных типов. Другими примерами конкретных опухолевых антигенов являются MART-1, Melan-A, p97, beta-HCG, GaINAc, MAGE-1, MAGE-2, MAGE-4, MAGE-12, MUC1, MUC2, MUC3, MUC4, MUC18, CEA, DDC, PlA, EpCam, антиген меланомы gp75, Hker 8, высокомолекулярный антиген меланомы, KI 9, Tyr1, Tyr2, члены семейства генов pMel 17, c-Met, PSM (муциновый антиген предстательной железы), PSMA (мембранный антиген, специфичный к предстательной железе), секреторный белок предстательной железы, альфа-фетопроtein, антигены CA 125, CA 19.9, TAG-72, BRCA-1 и BRCA-2.

Встроенная генная последовательность может продуцировать терапевтические молекулы ДНК других типов. Так, например, такие молекулы ДНК могут быть использованы для экспрессии функционального гена у индивидуума, страдающего генетическим заболеванием, вызванным дисфункциональным вариантом этого гена. Примерами таких заболеваний являются мышечная дистрофия Дюшенна, кистозный фиброз, болезнь Гоше и дефицит аденозиндезаминазы (ADA). Другими заболеваниями, при которых может быть показана генотерапия, являются воспалительные, аутоиммунные, хронические и инфекционные заболевания, включая такие заболевания, как СПИД, рак, нервные заболевания, сердечно-сосудистые заболевания,

гиперхолестемия, различные заболевания крови, включая различные анемии, талассемию и гемофилию, и эмфизема. Для лечения солидных опухолей может оказаться желательной экспрессия генов, кодирующих токсические пептиды (то есть, химиотерапевтические агенты, такие как рицин, дифтерийный токсин и фактор яда кобры); генов-супрессоров опухоли, таких как p53; генов, кодирующих последовательности мРНК, которые являются антисмысловыми по отношению к трансформирующим онкогенам; противоопухолевых пептидов, таких как фактор некроза опухоли (TNF) и другие цитокины, или трансдоминантно негативные мутанты трансформирующих онкогенов.

Также рассматриваются и другие типы терапевтических молекул ДНК. Так, например, могут быть встроены молекулы ДНК, которые транскрибируются в активную некодирующую форму РНК, например, короткую интерферирующую РНК (киРНК). Таким образом, способы согласно изобретению охватывают способы нокдауна экспрессии эндогенного гена или нокаута эндогенных генов с использованием некодирующих РНК в индуцибельном кластере.

Таким образом, способ согласно изобретению может быть применен для специфического и стабильного встраивания генной последовательности в индуцибельный кластер, где эта последовательность может регулируемо транскрибироваться. Это имеет множество преимуществ для соматических стволовых клеток и зрелых клеток. Этот способ позволяет проводить тщательно регулируемую генотерапию и будет гарантировать, что важные гены не будут подвергаться разрушению, а экспрессия индуцибельного кластера будет отключаться в случае возникновения каких-либо побочных эффектов. Он также позволяет проводить тщательно регулируемый нокдаун или нокаут эндогенного гена для выявления функции и экспрессии гена.

Настоящее изобретение относится к клеткам, полученным способом согласно изобретению. Клетки могут быть определены как модифицированные в первом генетически безопасном якорном сайте, включающем белок-регулятор транскрипции, и во втором генетически безопасном якорном сайте, включающем генную последовательность, функционально присоединенную к индуцибельному промотору, который

регулируется белком-регулятором транскрипции. Два GSH отличаются друг от друга и имеют характерные признаки. Предпочтительно, чтобы клетки были гомозиготными в обоих сайтах инсерции. Все элементы являются такими, как они описаны выше.

Клетки, полученные любыми способами согласно изобретению, находят свое применение в диагностических и терапевтических методах. Клетки могут быть использованы *in vitro* для изучения развития клеток, разработки систем тестирования новых лекарственных средств, способов скрининга, тщательной разработки схем терапии, разработки диагностических тестов и т.п. Такое применение является частью настоящего изобретения. Альтернативно, клетки могут быть трансплантированы человеку или животному в диагностических или терапевтических целях. Применение клеток в терапии также входит в объем настоящего изобретения. Клетки могут быть аллогенными (то есть, зрелые клетки могут быть удалены, модифицированы и возвращены тому же индивидууму) или могут быть взяты у донора (включая стволовую клеточную линию).

Все цитируемые здесь документы вводятся в настоящее описание посредством ссылки.

Последовательности

AAVS1 - NCBI GenBank S51329.1

SEQ ID NO:1: последовательность TetO2 19n

SEQ ID NO:2: геномная последовательность сайта инсерции hROSA

SEQ ID NO:3: STDtetR-nls (нуклеотиды) и SEQ ID NO:4 - STDtetR-nls (аминокислоты)

SEQ ID NO:5: OPTtetR-nls (нуклеотиды) и SEQ ID No6 - OPTtetR-nls (аминокислоты)

SEQ ID NO:7-80: Праймеры, представленные в таблице 3.

SEQ ID NO:81: Фигура 18B, AAVS1 FWD; SEQ ID NO:82: Фигура 18B, AAVS1 REV

SEQ ID NO:83: Фигура 18B, маркер; SEQ ID NO:84: Фигура 18B, маркер REV

SEQ ID NO:85: Фигура 19E, HI POL3 FWD; SEQ ID NO:82: Фигура 19E, HI POL3 REV.

Эта последовательность представляет собой геномную последовательность сайта инсерции hROSA26; включает гомологичную 5'-цепь, сайт разрезания (показан жирным), и гомологичную 3'-цепь: (SEQ ID NO:2).

GCTCGAAACCGGACGGAGCCATTGCTCTCGCAGAGGGAGGAGCGCTTCCGGCTAGCCTC
 TTGTCGCCGATTGGCCGTTTCTCCTCCCGCCGTGTGTGAAAACACAAATGGCGTATTCTGGTTG
 GAGTAAAGCTCCTGTCAGTTACGCCGTCGGGAGTACGCAGCCGCTTAGCGACTCTCGCGTTGCC
 CCCTGGGTGGGGCGGGTAGGTAGGTGGGGTGTAGAGATGCTGGGTGTGCGGGCGCGGCCGGCCCT
 CCTGCGGCGGGAGGGGAGGGTCAGTGAAATCGGCTCTGGCGCGGGCGTCTCCACCCCTCCCCT
 TCCTTCGGGGGAGTCGGTTTACCCGCCGCTGCTTGTCTTCGACACCTGATTGGCTGTGGAAGC
 TGTGGGACCGGGCCCTTGCTACTGGCTCGAGTCTCACATGAGCGAAACCACTGCGCGGGGCGCG
 GGGGTGGCGGGGAGGCGGGCGTTGGTACGGTCCTCCCCGAGGCCGAGCGCCGCAGTGTCTGGCC
 CCGCGCCCCTGCGCAACGTGGCAGGAAGCGCGCGCTGGAGGCGGGGGCGGGCTGCCGGCCGAGA
 CTTCTGGATGGCGGCGGCCGCGGCTCCGCCCCGGGTTCACCGCCTGAAGGGCGAGACAAGCC
 CGACCTGCTACAGGCACTCGTGGGGGTGGGGGAGGAGCGGGGGTCGGTCCGGCTGGTTTGTGGG
 TGGGAGGCGCTTGTTCTCCAAAACCGGCGCGAGCTGCAATCCTGAGGGAGCTGCGGTGGAGGA
 GGTGGAGAGAAGGCCGCACCCTTCTGGGCAGGGGGAGGGGAGTGCCGCAATACCTTTATGGGAG
 TTCTCTGCTGCCTCCCGTCTTGTAAGGACCGCCCTGGGCCTGGAAGAAGCCCTCCCTCCTTTCC
 TCCT**CGCGTGATCTCGTCATCGCCTCCATGTGCGAGTCGCTTCTCGATTAT**GGGCGGGATTCTTT
 TGCCTAGGCTTAAGGGGCTAACTTGGTCCCTGGGCGTTGCCCTGCAGGGGAGTGAGCAGCTGTA
 AGATTTGAGGGGCGACTCCGATTAGTTTATCTTCCCACGGACTAGAGTTGGTGTGAGGTTATT
 GTAATAAGGGTGGGGTAGGGAAATGGAGCTTAGTCATTACCTGGGGCTGATTTTATGCAACGA
 GACTGCGGATTATCACTACTTATCATTTTTTGGAGCATTTTTCTAGAGACAGACATAAAGCATGA
 TCACCTGAGTTTTATACCATTTGAGACCCTTGCTGCACCACCAAAGTGTAGCATCAGGTAAAT
 CTTAATAGAAAAATTTTAGCTTTTGCTTGAGAAACCAGTGCTTCCCTCCCTCACCTCTCTCCC
 CAGGCTCTCTACCCCTTTGCATCCCTACCAGGCATCTTAGCAACTCTCACTCATACTTGATCCC
 ATTTTCCATTTGTTGTACTTGCTCCTCTAGTATTCAGACATAGCACTAGCTTTCTCCCTCTCTT
 GATCTTGGGTAGCCTGGTGTCTCGCGAAACCAGACAGATTGGTTCCACCACAAATTAAGGCTTG
 AGCTGGGGCTTGACTCTTACCCAGCAGTGCTTTTATTCCTCCCTAGTTCACGTTCTTAAATGTT
 TATCTTGATTTTTCATTTTATCCTTTTTCCTTAGCTGGGATTCTGTCCCTGACCGTCTTCACAGT
 CCAGGTGATCTTGACTACTGCTTTACAGAGAATTGGATCTGAGGTAGGCAACATCTCCCTTTT
 TCTTCCTCTAAATACCTCTCATTTCTGTTCTTACCAGTTAGTAACTGATCTCAGATGCCTGTGT
 GATAGCTTCC

STDtetR-nls: (SEQ ID NO:3 и 4)

Нуклеотидные и аминокислотные последовательности белка-

репрессора, чувствительного к тетрациклину (tetR) и содержащего N-концевой сигнал локализации в ядре SV40 (NLS, выделено серым). Последовательности представлены до или после их оптимизации по кодам (STDtetR и OPTtetR, соответственно). Точками показаны синонимичные мутации, введенные в OPTtetR.

```

ATGCCAAAAAAGAAGAGGAAGCTATCTAGATTAGATAAAAGTAAAGTGATTAACAGCGCATTAGAGC
M P K K K R K V S R L D K S K V I N S A L E L
TGCTTAATGAGGTGGAATCGAAGGTTTAAACAACCCGTAAGCTCGCCAGAAAGCTAGGTGTAGAGCA
L N E V G I E G L T T R K L A Q K L G V E Q
GCCTACATTGTATTGGCATGTAAAAATAAGCGGGCTTTGCTCGACGCCTTAGCCATTGAGATGTTA
P T L Y W H V K N K R A L L D A L A I E M L
GATAGGCACCATACTCACTTTTGGCCCTTTAGAAGGGGAAAGCTGGCAAGATTTTACGTAATAACG
D R H H T H F C P L E G E S W Q D F L R N N A
CTAAAAGTTTGTAGATGTGCTTTACTAAGTCATCGCGATGGAGCAAAAGTACATTTAGGTACACGGCC
K S F R C A L L S H R D G A K V H L G T R P
TACAGAAAAACAGTATGAAACTCTCGAAAATCAATTAGCCTTTTATGCCAACAAGGTTTTCACATA
T E K Q Y E T L E N Q L A F L C Q Q G F S L
GAGAATGCATTATATGCACTCAGCGCTGTGGGGCATTTTACTTTAGGTTGCGTATTGGAAGATCAAG
E N A L Y A L S A V G H F T L G C V L E D Q E
AGCATCAAGTCGCTAAAGAAGAAAGGGAAACACCTACTACTGATAGTATGCCGCCATTATTACGACA
H Q V A K E E R E T P T T D S M P P L L R Q
AGCTATCGAATTATTTGATCACCAAGGTGCAGAGCCAGCCTTCTTATTCGGCCTTGAATTGATCATA
A I E L F D H Q G A E P A F L F G L E L I I
TGCGGATTAGAAAAACAACCTTAAATGTGAAAGTGGGTCTCCGCGGTAA
C G L E K Q L K C E S G S P R *

```

Последовательность для оптимизированного tetR: OPTtetR-nls
(SEQ ID NO:5 и 6):

```

ATGCCCAAGAAAAAGCGGAAGGTGTCCCGGCTGGAACAAGACCAAGTGATCAACAGCGCCCTGGAAC
M P K K K R K V S R L D K S K V I N S A L E L
TGCTGAACGAAGTGGGCATCGAGGGCCTGACCACCCGGAAGCTGGCCAGAAACTGGGCGTGGAACA
L N E V G I E G L T T R K L A Q K L G V E Q
GCCCACCCCTGTACTGGCACTGGAAGAACAAGCGGGCCTGCTGGACGCCCTGGCCATCGAGATGCTG
P T L Y W H V K N K R A L L D A L A I E M L
GACCGGCACCAACACACACTTTTGGCCCTTGAAGGCGAAAGCTGGCAGGACTTCTGCGGAACAACG
D R H H T H F C P L E G E S W Q D F L R N N A
CCAAGAGCTTCAGATGCGCCCTGCTGAGCCACCGGGAACGGCGCCAAAGTGCACTGGGCACCAAGACC
K S F R C A L L S H R D G A K V H L G T R P
CACCGAGAAGCAGTACGAGACACTGGAAAACAGCTGGCCTTCTCTGTGCCAGCAGGGCTTCAGCCTG
T E K Q Y E T L E N Q L A F L C Q Q G F S L
GAAAACGCCCTGTACGCCCTGAGCGCCGTGGGCCACTTTACCTTGGGCTGCGTGCTGGAAGATCAGG
E N A L Y A L S A V G H F T L G C V L E D Q E
AACACCAGGTGCGCAAGAGGAAAGAGAGACACCCACCAAGACAGCATGCCCCCTGCTGAGACA
H Q V A K E E R E T P T T D S M P P L L R Q
GGCCATCGAGCTGTTCGATCATCAAGGCGCCGAGCCCGCCTTCTGTTCGGCCTGGAAGTATCATC
A I E L F D H Q G A E P A F L F G L E L I I
TGCGGCTCGAGAAGCAGCTGAAGTGGAGAGCGGCTCCCCAGATGA
C G L E K Q L K C E S G S P R *

```

Настоящее изобретение описано ниже со ссылками на следующие неограничивающие примеры:

Примеры

Материалы и методы, используемые в Примерах:

Поддерживающая культура hPSC и дифференцировка зародышевого слоя

Было проведено культивирование фидерных и несывороточных hESC (линии H9; WiCell) и hiPSC (Cheung et al, Nat. Biotechnol. 30, 165-173 (2012)). Вкратце, клетки высевали на культуральные чашки, покрытые средой с желатином/MEF [MEF-средой, состоящей из улучшенной среды DMEM/F12 (90%, Gibco), фетальной бычьей сыворотки (10%, Gibco), L-глутамина (1 mM, Gibco), 2-меркаптоэтанола (0,1 mM, Sigma-Aldrich) и пенициллина/стрептомицина (1%, Gibco)], и культивировали в среде с определенным химическим составом [CDM, состоящей из IMDM (50%, Gibco), F12 (50%, Gibco), концентрированных липидов (100x, Gibco), монотиоглицерина (450 мкМ, Sigma-Aldrich), инсулина (7 мкг/мл, Roche), трансферина (15 мкг/мл, Roche), фракции альбумина бычьей сыворотки V (5 мг/мл), и пенициллина/стрептомицина (1%), в которую было добавлено 10 нг/мл активина-A и 12 нг/мл FGF2. Клетки пассировали в виде небольших скоплений с использованием коллагеназы через каждые 5-6 дней.

Дифференцировка hPSC в зародышевые слои была индуцирована в адгезивных культурах hESC в соответствии с ранее опубликованными протоколами прямой дифференцировки в энтодерму, боковую пластинчатую мезодерму и нейроэктодерму (Touboul, T. et al. Hepatology 51, 1754-1765 (2010), Cheung et al, (2012) и Douvaras, P. et al. Stem Cell Reports 3, 250-259 (2014)). Вкратце, энтодерму в конечной стадии получали путем культивирования hPSC в течение 3 дней в среде CDM-PVA (без инсулина), в которую добавляли FGF2 (20 нг/мл), активин-A (100 нг/мл), BMP4 (10 нг/мл, Marko Hyvonen, Dept. of Biochemistry, University of Cambridge) и LY-294002 (10 мкМ, Promega) 3. Для дериватизации нейроэктодермы, hPSC культивировали в течение 6 дней в среде CDM-BSA, в которую добавляли SB-431542 (10 мкМ, Tocris), LDN-193189 (0,1 мкМ, Tocris) и RA (0,1 мкМ, Sigma) 4.

Боковую пластинчатую мезодерму получали путем культивирования hPSC в течение 36 часов в среде CDMFVA, в которую добавляли FGF2 (20 нг/мл), 10 нг/мл BMP4 (R&D) и LY294002 (10 мкМ), и в течение последующих 3,5 дней в среде CDM-PVA, в которую добавляли FGF2 (20 нг/мл) и BMP4 (50 нг/мл).

Дифференцировка hESC. Дифференцировку инициировали в адгезивных культурах hESC через 48 часов после пассажей. Среду обычно заменяли ежедневно, и объемы корректировали для достижения нужной плотности клеток. Зрелые клетки получали методами, описанными ранее в литературе. Полученными зрелыми клетками являются нервные клетки, остециты, хондроциты, клетки гладких мышц, сердечные фибробласты, кардиомиоциты, клетки кишечника, клетки поджелудочной железы, гепатоциты, холангиоциты или клетки легких.

Конструкции для таргетирования генов и молекулярное клонирование

В данном разделе описаны разработка и конструирование рРНК hROSA26 и экспрессионных плазмид Cas9n: Стратегия на основе CRISPR/Cas9n была разработана для специфического нацеливания на локус hROSA26 и для встраивания индуцибельных кластеров посредством гомологичной рекомбинации. Для индуцирования геномного DSB в правильном сайте интеграции была сконструирована система никазы CRISPR/Cas9. В отличие от широко применяемой нуклеазы Cas9 дикого типа, которая вводится одной рРНК в геномный сайт-мишень, D10A-мутантная никаза Cas9 (Cas9n) направляется посредством пары соответствующим образом сконструированных рРНК, что приводит к одновременному введению одноцепочечных разрывов в обе цепи ДНК-мишени. Эта стратегия позволяет эффективно удваивать число оснований, необходимых для редактирования генома и, тем самым, повышения специфичности. Компьютерное Web-приложение «CRISPR Design Tool» было использовано для определения потенциальных сайтов-мишеней для cr-РНК-регулируемых нуклеаз, которые находятся поблизости от сайта интеграции. Во фрагменте последовательности из 250 п.о., расположенном вокруг сайта-мишени (125 п.о. на каждом участке фактического сайта интеграции), самым лучшим выходом является

пара рРНК, которые в совокупности дают оценку «высокого качества» 97 баллов без предсказанных побочных эффектов. рРНК [рРНК-5'-GTCGAGTCGCTTCTCGATTA-(TGG)-3' и рРНК-В, 5'-GGCGATGACGAGATCACGCG-(AGG)-3' (PAM-сайты, указанные в скобках), были синтезированы *de novo* и лигированы в экспрессионные векторы. Конечные плазмиды кодируют любую из двух рРНК, соответственно, и D10A-мутант Cas9n (фигуры 20 и 21).

Была сконструирована донорная плазида, которая служила в качестве матричной ДНК для облегчения гомологично ориентированной репарации Cas9n-индуцированного DSB. Два гомологичных плеча hROSA26 были получены путем ПЦР-амплификации с высокой степенью достоверности. Геномная ДНК, которая была выделена из H9 hESC, служила в качестве матрицы. Гомологичные 5'- и 3'-плечи имели длину 904 п.о. и 869 п.о., соответственно. Обе эти последовательности затем встраивали в сайт множественного клонирования вектора pUC19. Для нацеливания на локус hROSA26, клетки трансфецировали плазмидой, двумя конструкциями рРНК/Cas9n и донорной плазмидой EGFP (фигура 22).

Таргетирующий вектор pR26_CAG-rtTA (фигура 23) конструировали путем клонирования кодирующей последовательности третьей генерации rtTA (ПЦР-амплифицированной из pLVX-Tet3G) в BamHI/MluI-сайты pR26_CAG-EGFP и, тем самым, замены последовательности EGFP. Экспрессионные плазмиды AAVS1 ZFN были любезно предоставлены Dr. Kosuke Yusa (Wellcome-Trust Sanger Institute). Индуцибельный таргетирующий вектор EGFP AAVS1 был сконструирован специалистами компании Gibson Assembly (New England Biolabs), где три вставки лигировали в EcoRI/HindIII-сайты сайта множественного клонирования вектора pUC19 (Thermo Fisher Scientific): Первая вставка содержала вышерасположенное гомологичное плечо AAVS1, акцептор сплайсинга, T2A-сайт и кластер резистентности к пуromину (ПЦР-амплифицированный из pTRE-EGFP; дополнительный ген 22074, депонированный Rudolf Jaenisch). Вторая вставка содержала индуцибельный промотор TRE3G (ПЦР-амплифицированной из pLVX-TRE3G). Третья вставка содержала экспрессионный кластер EGFP и нижерасположенное гомологичное

плечо AAVS1 (ПЦР-амплифицированное из pTRE-EGFP; дополнительный ген 22074, депонированный Rudolf Jaenisch). Полученную плазмиду обозначали pAAV_TRE-EGFP (фигура 32). Таргетирующие векторы pAAV_TRE-NGN2 и pAAV_TRE-MYOD1 (фигура 33) были сконструированы путем клонирования кодирующей последовательности NGN2 и MYOD1, соответственно (NGN2: ПЦР-амплифицированный из pLVX-TRE-NGN2 и любезно предоставлен Oliver Brüstle; MYOD1: ПЦР-амплифицированный из коммерчески доступной кДНК-плазмиды, Open Biosystems MHS6278-202832821, рег. No.: BC064493, Клон ID: 5022419) в SpeI/EcoRI-сайты pAAV_TRE-EGFP и замены последовательности EGFP.

Другие плазмиды также получали аналогичными методами, и все используемые плазмиды представлены на фигурах 20-33. Эти плазмиды либо конструировали, либо получали в качестве подарка. Плазмидами, используемыми в Примерах, являются (по порядку на фигурах 20-33): pSpCas9n(BB),_R26-R, pSpCas9n(BB) (комбинация этих двух плазмид, как было предсказано, индуцирует специфический двухцепочечный разрыв в интроне между экзонами 1 и 2 THUMPDS3-AS1 на хромосоме 3 (локус ROSA26)),_R26-L pR26_CAG_EGFP, pR26_CAG_rtTA, pZFN-AAVS1-L-ELD (нуклеаза «цинковые пальцы», слева), pZFN-AAVS1-R-KKR (нуклеаза «цинковые пальцы», справа), pAAV_CAG_EGFP (донор), pR26-Neo_CAG-OPtTetR (нацеливание оптимизированного по кодонам tetR на hROSA26), pAAV-Puro_iKD (нацеливание индуцибельной кшРНК на AAVS1), pAAV-Neo_CAG-Cas9 (нацеливание Cas9 на AAVS1), pAAV-Puro_siKO (нацеливание индуцибельной pРНК на AAVS1), pAAV-Puro_siKO-2TO (нацеливание индуцибельной pРНК на AAVS1, вариант с 2 оперонами tet в промоторе), pAAV_TRE-EGFP (индуцибельная сверхэкспрессия EGFP, сцепленная с геном) и pAAV_TRE-MYOD1 (индуцибельная сверхэкспрессия MYOD1 в мышцах).

Таргетинг гена

Таргетинг локуса hROSA26 и AAVS1 для нокдауна и нокаута гена осуществляли путем нуклеофекции. Человеческие плюрипотентные стволовые клетки (PSC) диссоциировали на одиночные клетки с использованием TrypLE Select (Gibco), и 2×10^6

клеток подвергали нуклеофекции (в реакционном объеме 100 мкл; всего 12 мкг ДНК, которые были в равной степени разделены между двумя плазмидами рРНК/Cas9n и таргетирующим вектором) с использованием набора для нуклеофекции Lonza P3 Primary Cell 4D-Nucleofector X и термоячейки CA-137 системы Lonza 4D-Nucleofector. hPSC, подвергнутые нуклеофекции, высевали на облученные мышинные эмбриональные фибробласты, резистентные ко множеству лекарственных средств (DR4), и культивировали в среде KSR [состоящей из улучшенной среды DMEM/F12 (80%), заменителя сыворотки для нокаута (20%, Gibco), L-глутамина (1 мМ), 2-меркаптоэтанола (0,1 мМ) и пенициллина/стрептомицина (1%)], в которую был добавлен FGF2 (4 нг/мл, Department of Biochemistry, University of Cambridge). Y-27632 (5 мкМ, Tocris) добавляли за 24 часа до и после нуклеофекции для стимуляции выживания клеток. Через 3-6 дней, резистентные к неомицину hPSC отбирали путем добавления G418 (50 мкг/мл, Sigma-Aldrich) в течение 7-10 дней. Затем, отдельные клоны собирали, размножали в условиях без фидерных клеток и, наконец, анализировали путем генотипирования.

Таргетинг локуса AAVS1 также осуществляли путем липофекции. Человеческие PSC высевали в 6-луночные планшеты без фидерных клеток и трансфецировали через 48 часов после пересева. Трансфекцию осуществляли в среде Opti-MEM (Gibco), в которую добавляли липофектамин 2000 (10 мкл/лунку, Thermo Fisher Scientific) и всего 4 мкг ДНК (в равной степени разделенные между двумя плазмидами AAVS1 ZFN и таргетирующим вектором) в течение 24 часов. Через 3-5 дней, резистентные hPSC отбирали путем добавления пурамицина (1 мкг/мл, Sigma-Aldrich) в течение 5-8 дней. Затем, отдельные клоны собирали, размножали и анализировали путем генотипирования. Резистентность к антибиотикам может быть использована для отбора клональных линий.

Клоны hPSC, резистентные к лекарственному средству и полученные в экспериментах по таргетированию, скринировали с помощью геномной ПЦР для подтверждения сайт-специфической интеграции индуцибельного кластера, для определения числа аллелей-мишеней и для исключения интеграций, не являющихся

мишенями. ПЦР осуществляли с использованием ДНК-полимеразы LongAmp Taq (New England Biolabs). В Таблице 2 представлены комбинации праймеров, используемых для различных таргетирующих векторов. Результаты всех экспериментов по таргетированию систематизированы в Таблице 1. Анализ кариотипа осуществляли стандартными методами G-бэндинга (Medical Genetics Service, Cambridge University Hospitals). Для получения человеческих PSC-мишеней в целях проведения анализа хромосом, клетки инкубировали в свежей культуральной среде, в которую были добавлены Y-27632 (5 мкМ, Tocris) и кольцемид KaryoMAX (100 нг/мл, Gibco) в течение 4 часов при 37°C. Затем, клетки собирали в виде одиночных клеток, промывали и осаждали. Набухание ядра и распределение хромосом достигалось путем обработки гипотоническим 0,055 М раствором KCl в течение 5-10 минут. И наконец, клетки фиксировали метанолом и ледяной уксусной кислотой (в отношении 3:1).

Для OPTiKD, таргетинг AAVS1 targeting осуществляли путем липофекции, как описано ранее. Вкратце, hPSC высевали в 6-луночные планшеты без фидерных клеток и трансфецировали через 48 часов после пересева клеток 4 микрограммами ДНК (в равной степени разделенными между двумя плазмидами AAVS1 ZFN и таргетирующим вектором) с использованием 10 мкл/лунку липофектамина 2000 в среде Opti-MEM (Gibco) в течение 24 часов в соответствии с инструкциями производителей. Через 4 дня, 1 мкг/мл пурамицина добавляли в культуральную среду и отдельные клоны собирали и размножали после отбора в течение 7-10 дней.

Для каждого сайта OPTiKO, таргетинг AAVS1 осуществляли путем нуклеофекции. hESC, предварительно обработанные в течение 16 часов 10 мкМ Y-27632 (Tocris), диссоциировали с образованием скоплений из 2-8 клеток с использованием аккутазы (Gibco), и 2×10^6 клеток подвергали нуклеофекции в объеме 100 мкл вместе с 12 мкг ДНК (4 мкг каждой для двух плазмид ZFN и 2 мкг каждой для двух таргетирующих векторов) с использованием набора Lonza P3 Primary Cell 4D-Nucleofector X и термоячейки на системе CA-137 Lonza 4D-Nucleofector, в соответствии с инструкциями

производителей. hESC, подвергнутые нуклеофекции, высевали на облученные мышинные эмбриональные фибробласты DR4 (резистентные к пурамицину и неомицину) и культивировали в среде KSR, в которую было добавлено 4 нг/мл FGF2 и 10 мкМ Y-27632 (последнюю процедуру проводили только в первые 24 часа). Через 4 дня, колонии hPSC, несущие ген резистентности к пурамицину и неомицину, отбирали в течение 7-10 дней с 25 мкг/мл генетицина (сульфата G418, Gibco) и 0,5 мкг/мл пурамицина. Затем, отдельные клоны собирали и размножали в условиях без фидеров.

AAVS1-EGFP, ROSA26-EGFP, ROSA26-STDtetR, ROSA26-OPTtetR и ROSA26-EGFPd2 hESC получали путем липофекции (локуса AAVS1) или нуклеофекции (локуса ROSA26) таргетирующих векторов с использованием AAVS1 ZFN или пар ROSA26 CRISPR/Cas9n (как описано выше). 2 мкг/мл бластицидина S-HCl (Gibco) использовали для плазмиды pR26-Bsd_CAG-EGFPd2. Получение hESC, сверхэкспрессирующих индуцибельный EGFP и несущих трансген ROSA26-rtTA и AAVS1-TRE-EGFP, осуществляли как описано в настоящей заявке. Вкратце, клетки последовательно таргетировали геном сначала путем нуклеофекции плазмид pR26-Neo_CAG-rtTA с ROSA26 CRISPR/Cas9n, а затем путем липофекции плазмид pAAVPuro_TRE-EGFP с AAVS1 ZFN.

Клональные линии hPSC с геном-мишенью скринировали с помощью геномной ПЦР для подтверждения сайт-специфического таргетинга, для определения числа аллелей-мишеней и для исключения интеграций таргетирующей плазмиды, не являющихся мишенями (см. фиг. 16A).

Сверхэкспрессия индуцибельного кластера

Сверхэкспрессию индуцибельных кластеров (EGFP, NGN2, MYOD1 и OLIG2-SOX10, соответственно) индуцировали путем добавления гиклата доксициклина (Sigma-Aldrich) в культуральную среду. Если это не оговорено особо, то доксициклин использовали в конечной концентрации 1 мкг/мл. Среду, содержащую доксициклин, хранили в защищенном от света месте и заменяли через каждые 24 часа. Клетки, экспрессирующие EGFP, обозначали здесь как OPTi-EGFP, клетки, экспрессирующие NGN2, обозначали как OPTi-NGN2, клетки, экспрессирующие MYOD1, обозначали как OPTi-MYOD1, а клетки,

экспрессирующие OLIG2-SOX10, обозначали как OPTi-OLIG2-SOX10.

Нокаут и нокдаун индуцибельного гена

Если это не указано особо на чертежах или в примерах, то гидрохлорид тетрациклина (sigma-Aldrich) использовали в концентрации 1 мкг/мл для индуцирования нокдауна или нокаута гена.

Индукция нейронов

Плюрипотентные клетки OPTi-NGN2 диссоциировали на одиночные клетки с использованием TrypLE и высевали в чашки, покрытые Матригелем (35 мкг/см², Scientific Laboratory Supplies) при плотности 75000 клеток на лунку 12-луночного планшета. Через 24-48 часов после разделения инициировали прямое перепрограммирование. Если это не оговорено особо, то индуцирование осуществляли в среде DMEM/F12 (Gibco), в которую добавляли глутамакс (100×, Gibco), заменимые аминокислоты (100×, Gibco), 2-меркаптоэтанол (50 мкМ), пенициллин/стрептомицин (1%), и доксициклин (1 мкг/мл). После индуцирования в течение 2 дней, среду заменяли на нейробазальную среду, в которую добавляли глутамакс (100×), B27 (50×, Gibco), BDNF (10 нг/мл, Peprotech), NT3 (10 нг/мл, R&D Systems), пенициллин/стрептомицин (1%) и доксициклин (1 мкг/мл).

Индукция миоцитов скелетной мышцы

Плюрипотентные клетки OPTi-MYOD1 диссоциировали на одиночные клетки с использованием TrypLE и высевали в чашки, покрытые желатином/средой MEF при плотности 100000 клеток на лунку 12-луночного планшета. Через 24-48 часов после разделения инициировали прямое перепрограммирование. Если это не оговорено особо, то индуцирование осуществляли в среде DMEM (Sigma-Aldrich), в которую добавляли L-глутамин (2 мМ), 2-меркаптоэтанол (50 мкМ), пенициллин/стрептомицин (1%), инсулин (7 мкг/мл), все транс-ретиновые кислоты (1 мкМ, Sigma-Aldrich) и доксициклин (1 мкг/мл). После индуцирования в течение 5 дней, в среду добавляли CHIR99021 (3 мкМ, Tocris) и термоинактивированную лошадиную сыворотку (2%, Gibco) для ускорения созревания.

Индукцирование олигодендроцитов

Плюрипотентные клетки hPSC OLIG2-2A-SOX10 OPTi-OX культивировали в виде колоний в чашках для культивирования, покрытых желатином/MEF. Перед началом индуцирования, их обрабатывали SB и LDN в течение ночи. На следующий день, индуцирование инициировали в среде CDM, в которую добавляли доксициклин (1 мкг/мл) и RA (0,1 мкМ). На следующий день после индуцирования, клетки разделяли в среде CDM, в которую добавляли RA (0,1 мкМ), PM (1 мкМ) и Y-27632 (5 мкМ), PDGF $\alpha\alpha$ (20 нг/мл, Peprotech), FGF2 (5 нг/мл) в чашках для культивирования, покрытых PDL/ламинином (100000 клеток на лунку 12-луночного планшета). На следующий день, клетки переносили в среду для олигодендроцитов, состоящую из DMEM/F12, в которую добавляли глутамакс (100 $\times$), заменимые аминокислоты (100 $\times$), 2-меркаптоэтанол (1000 мкМ), пенициллин/стрептомицин (100 $\times$), добавку N2 (100 $\times$), добавку B27 (50 $\times$), инсулин, 7 мкг/мл (Marko Hyvonen), T3, 60 нг/мл (Sigma), биотин, 100 нг/мл (Sigma), db-cAMP, 1 мкМ (Sigma). В среду для олигодендроцитов добавляли dox (1 мкг/мл), PDGF $\alpha\alpha$ (20 нг/мл), FGF2 (5 нг/мл), RA (0,1 мкМ) и PM (1 мкМ). Через семь дней после индуцирования, RA и PM удаляли. Для поддержания индуцированных клеток в пролиферативном состоянии, клетки пересеивали через каждые 4 дня (75000 клеток на лунку 24-луночного планшета) при постоянном присутствии митогенов PDGF $\alpha\alpha$ и FGF2. Для дифференцировки пролиферирующих предшественников олигодендроцитов, PDGF $\alpha\alpha$ и FGF2 не использовали. Затем добавляли человеческий рекомбинантный NT3 (5 нг/мкл, R&D Systems) для повышения выживаемости клеток.

Количественная ПЦР в реальном времени (кол.ПЦР)

РНК экстрагировали с использованием набора для экстракции общей РНК млекопитающих GenElute Mammalian Total РНК Miniprep и серии ферментов для расщепления ДНКазы I на колонке (Sigma-Aldrich). Синтез кДНК осуществляли с использованием набора для синтеза первой цепи кДНК Maxima (Thermo Fisher Scientific). Для кол.ПЦР использовали основную смесь для ПЦР с использованием белка SYBR, флуоресцирующего в зеленом диапазоне спектра,

Applied Biosystems. Образцы загружали на устройство для быстрой ПЦР Applied Biosystems 7500. Все образцы анализировали в рабочих дубликатах и нормализовали на ген домашнего хозяйства, то есть, ген порфобиллиноген-деаминазы 1 (PBGD). Результаты анализировали методом $\Delta\Delta C_t$. См. Таблицу 3 для праймерных последовательностей.

Проточная цитометрия

Для анализа на экспрессию EGFP, клетки собирали с использованием TrypLE Select (Gibco) в течение 5-10 минут при 37°C с получением моноклеточной суспензии. После промывки PBS, клетки ресуспендировали в охлажденном льдом PBS, в который добавляли DAPI (10 мкг/мл), и инкубировали в течение 5 минут на льду. Клетки анализировали на проточном цитометре Cyan ADP для определения уровней экспрессии EGFP жизнеспособными клетками (DAPI-негативными). Для окрашивания и анализа на экспрессию тяжелых цепей миозина, клетки собирали с использованием TrypLE Select (как и в анализе на экспрессию EGFP), один раз промывали PBS, фиксировали и делали проницаемыми с использованием раствора Cytofix/Cytoperm (BD Biosciences). Затем, клетки промывали и блокировали в буфере для сообщения проницаемости/промывочном буфере (BD Biosciences), в который был добавлен 3% альбумин бычьей сыворотки (BSA), при 4°C в течение ночи. Окрашивание ФЭ-конъюгированным анти-MYH антителом (Таблица 4) осуществляли в буфере для сообщения проницаемости/промывочном буфере в течение 1 часа при 4°C в темноте. После трех промывок буфером для сообщения проницаемости/промывочным буфером, клетки анализировали на проточном цитометре Cyan ADP для определения уровней экспрессии МНС. Анализ данных проводили с использованием программ FlowJo (v10) и Graphpad Prism (v6).

Вестерн-блот-анализ

Белок целых клеток экстрагировали средой CellLytic M (Sigma-Aldrich), в которую был добавлен полный ингибитор протеазы (Roche), а затем количественно оценивали с использованием набора для быстрой количественной оценки белка (Sigma-Aldrich). Электрофорез белка осуществляли с использованием буфера для

образцов NuPAGE LDS и 4-12% предварительно сформованного бис-трис-геля NuPAGE Precast (Invitrogen). После переноса белка на ПВДФ, мембраны блокировали PBS, в который был добавлен 0,05% Твина-20 (PBST) в 4% молоке в течение 1 часа при комнатной температуре, и инкубировали с первыми антителами в течение ночи в PBST в 4% молоке. Мембраны промывали PBST, инкубировали с ПХ-конъюгированными вторыми антителами (Sigma-Aldrich) в PBST в 4% молоке, инкубировали с субстратом для вестерн-блоттинга Pierce ECL2 (Thermo Fisher Scientific), и экспонировали с рентгеновской пленкой Super RX (Fujifilm).

Иммуноцитохимический анализ

Клетки фиксировали в 4% параформальдегиде (разведенном в PBS) в течение 20 минут при комнатной температуре, а затем три раза промывали PBS. После этого, клетки блокировали 10% ослиной сывороткой (Sigma-Aldrich) и делали проницаемыми с использованием 0,3% Тритона X-100 (разведенного в PBS) в течение 20 минут при комнатной температуре. Затем, клетки инкубировали с соответствующим образом разведенными первыми антителами (дополнительные экспериментальные процедуры) в 2% ослиной сыворотке и 0,1% Тритоне X-100 (разведенном в PBS) при 4°C в течение ночи. Тритон-X не использовали ни в одной стадии, если проводили окрашивание поверхностного антигена PDGFRA, A2B5, и O4. После трех промывок PBS, клетки инкубировали в течение 1 часа при комнатной температуре с соответствующими вторыми ослиными антителами, конъюгированными с флуорофором (Alexa Fluor 488, 555, 568, и/или 647) в PBS, в который была добавлена 1% ослиная сыворотка. Ядра визуализировали с использованием 4',6-диамидино-2-фенилиндола (DAPI, Thermo Fisher Scientific). Экспрессию EGFP и иммунологическое окрашивание визуализировали на конфокальном микроскопе Zeiss LSM 700 (Leica). Процент клеток, позитивных по β III-тубулину, вычисляли путем определения уровней экспрессии β III-тубулина по меньшей мере в 50 произвольно выбранных DAPI-позитивных клетках в 3 визуальных полях для 3 биологических дубликатов на инвертированном флуоресцентном микроскопе Olympus IX71.

Статистический анализ проводили с использованием программы GraphPad Prism (v6). Число дубликатов, используемый статистический анализ и результаты анализа описаны в пояснениях к чертежам. Если это не оговорено особо, то данные представлены как среднее $\pm$ ср. кв. ош.

Пример 1: Двойной таргетинг EGFP

Для разработки стратегии индуцибельной сверхэкспрессии в hPSC, авторами было осуществлено последовательное нацеливание двух компонентов системы Tet-ON на два различных GSH. Конститутивно экспрессируемый rtTA третьей генерации нацеливали на человеческий локус ROSA26 (hROSA26) с использованием стратегии нацеливания на основе CRISPR/Cas9n, а индуцибельный EGFP-кластер встраивали в AAVS1 (фиг. 1A; Фигуры 4a-c). Нацеливание на оба hROSA26 и AAVS1 было в высокой степени эффективным (Фигуры 4d-e, таблица 1), и не влияло на геномную стабильность, самообновление и дифференцировку hPSC (данные не приводятся), что является аргументом против rtTA-зависимой клеточной токсичности.

Затем отбирали клоны, нацеленные на два GSH и несущие одну или две копии каждого из двух индуцибельных кластеров (фигура 5a). Гомозиготный таргетинг на rtTA давал приблизительно в два раза более высокие уровни белка rtTA (фигура 5b), а также значительно более высокие уровни EGFP после индуцирования, по сравнению с гетерозиготной экспрессией rtTA (фигуры 5c-5e). Кроме того, клоны с гомозиготным таргетингом индуцибельного кластера EGFP давали более высокие и более однородные уровни EGFP по сравнению с линиями с гетерозиготным таргетингом (фигуры 5c-5e). Важно отметить, что все правильно нацеленные линии показали надежную индуцибельную экспрессию EGFP, которая была по меньшей мере в двадцать раз выше по сравнению с экспрессией сильного конститутивного промотора CAG (фиг. 1b, фигуры 5c-e). В целом, эти результаты подтвердили первоначальную гипотезу авторов о том, что нацеливание на две копии обоих элементов системы Tet-ON приведет к максимальной экспрессии после индуцирования. Пик уровней EGFP был достигнут приблизительно

через четыре дня после индуцирования, и экспрессия быстро прекращалась после удаления доксициклина (рис. 1с). Кроме того, экспрессия EGFP может быть оттитрована путем коррекции дозы доксициклина (фиг. 1d). Важно отметить, что индуцибельная экспрессия EGFP была не только в высокой эффективной в hPSC, но также и во время дифференцировки в зародышевые слои (цветные фотографии не показаны, данные на фигурах 6a-6d). И, наконец, в соответствии с известной жесткой регуляцией транскрипции систем Tet-ON третьего поколения не было обнаружено какой-либо детектируемой фоновой экспрессии мРНК или белка EGFP в отсутствии доксициклина, как было определено с помощью проточной цитометрии и кол.ПЦР, соответственно (фиг. 1b, фигура 6d). В целом, эти результаты показали, что нацеливание системы Tet-ON на два GSH является мощной стратегией для оптимальной экспрессии индуцибельных кластеров в hPSCs и их производных.

Пример 2: Получение возбуждающих нейронов коры головного мозга из hESC и hiPSC

Предыдущие исследования показали, что эти клетки могут быть легко получены посредством сверхэкспрессии лентивирусами любого из про-нейронных bHLH-факторов (ASCL1, Ngn2 или NEUROD1) в hPSC. Таким образом, авторами были получены Opti-NGN2 hPSC (фиг. 2a, Таблица 1). Индуцирование NGN2 приводило к быстрому ингибированию факторов плюрипотентности (фигура 7) и к инициации процесса транскрипции в нейронах (фиг. 2b). В индуцированных клетках наблюдались нейронные процессы уже через три дня после индуцирования (данные не приводятся). Через одну неделю, все клетки обнаруживали нейронную морфологию и экспрессировали все маркерные белки нейронов, такие как β III-тубулин и MAP2 (фигура 2c). Количественный ОТ-ПЦР-анализ показал сильное индуцирование типичных маркеров переднего мозга, таких как *BRN2* и *FOXG1*, и глутаматергических нейронов, включая *GRIA4* и *VGLUT2* (фиг. 2b), что свидетельствует об идентичности возбуждающих нейронов коры головного мозга. В целом, эти результаты показали резкое повышение скорости и эффективности генерирования нейронов по сравнению с традиционными протоколами дифференцировки hPSC, и

значительному повышению эффективности и чистоты по сравнению с протоколами трансдифференцировки и прямого перепрограммирования на основе лентивирусов. Аналогичные результаты были получены для OPTi-NGN2 hPSC, что подтверждает надежность этого метода. И наконец, авторами настоящего изобретения не было обнаружено какого-либо снижения эффективности индуцирования нейронов в течение всего длительного периода культивирования Opti-NGN2 hPSC (>25 пассажей, фиг. 2с). В целом, полученные авторами результаты показали, что OPTi-NGN2 hPSC могут быть использованы в качестве неисчерпаемого источника для неограниченного в высокой степени масштабируемого, быстрого, одностадийного, безвирусного и почти детерминированного генерирования нейронов.

Пример 3: Получение миоцитов скелетной мышцы

Известно, что фактор транскрипции MYOD1 индуцирует миогенную трансдифференцировку, если он сверхэкспрессируется в соматических клетках различных типов, однако, в настоящее время ведутся дебаты по поводу способности hPSC подвергаться прямому миогенному перепрограммированию, индуцированному MYOD1. Авторами были получены OPTi-MYOD1 hPSC (Таблица 1), однако, они заметили, что индуцирование экспрессии MYOD1 после обработки доксициклином приводило почти к полной гибели клеток в течение 3-5 дней в широком диапазоне условий культивирования, которые, как было предположительно установлено ранее, облегчают превращение hPSC в скелетные миоциты. Поскольку уже установлено, что стратегии клеточного перепрограммирования могут быть улучшены путем объединения сверхэкспрессии фактора транскрипции с внеклеточными молекулами передачи сигнала, то авторами был проведен систематический скрининг на про-миогенные факторы посредством модуляции основных каскадов передачи сигнала, которые участвуют в образовании примитивных штриховых полос, в сомитогенезе и в миогенезе. Авторами было обнаружено, что добавление всех транс-ретиновых кислот (RA) в комбинации со сверхэкспрессией MYOD1 приводит к быстрому и почти полному превращению в миоциты, позитивные по миогенину и тяжелым цепям миозина (MHC), на день 5 после индуцирования. Эффект RA зависел от концентрации и

опосредовался изоформами RA-рецепторов $RAR\alpha$ и $RAR\beta$, что соответствовало паттерну экспрессии рецепторов RA во время миогенеза на стадии развития (фигура 8). Считается, что этот эффект не зависит от механизма сверхэкспрессии MYOD1. Индуцированные скелетные миоциты обнаруживали типичную удлинённую морфологию, подобную морфологии веретенообразных клеток, подвергались интенсивному слиянию и обнаруживали сверхэкспрессию миогенных маркеров на уровнях мРНК и белка (фиг. 3b, фигуры 9a-9c). Добавление ацетилхолина (ACh) или селективного агониста ACh-рецептора карбахола в наномолярных концентрациях приводило к полному сокращению мышечных волокон, что указывало на функциональность индуцированных миоцитов. Аналогичные результаты были получены с использованием Opti-MYOD1 hiPSC (данные не приводятся). Важно отметить, что эффективность индуцирования миогенов не снижалась в течение длительных периодов культивирования (>50 пассажей, фиг. 3D), что продемонстрировало надёжность и воспроизводимость этого метода. И наконец, авторы отметили, что уровни MYOD1-индуцируемого кластера позитивно коррелируют с эффективностью превращения, что еще раз подчеркивает важность надёжной доставки генов и преимущество этого метода по сравнению с методами перепрограммирования, опосредуемого лентивирусами (фигура 10). В целом, стратегия прямого перепрограммирования OPTi-MYOD1 действует приблизительно в семь раз быстрее и в пять раз эффективнее, чем большинство уже известных протоколов дифференцировки hPSC в скелетные миоциты. По сравнению с предыдущими протоколами прямого перепрограммирования (Tanaka, A. *et al. PLoS One* **8**, e61540 (2013) and Abujarour, R. *et al. Stem Cells Transl. Med.* **3**, 149-60 (2014)), этот метод является более эффективным (>95% по сравнению с 30-80%), не предусматривает использования произвольно встроенных индуцибельных кластеров, проводится с использованием улучшенного химического состава, является полностью воспроизводимым и более масштабируемым.

Эти результаты показали, что такой способ регуляции экспрессии индуцибельного кластера в hPSC может быть применен в

качестве неисчерпаемого источника для высокоэффективного и крупномасштабного производства гомогенных клеточных популяций. В настоящее время, скорость индуцирования и чистота желаемых клеток-мишеней никогда еще не достигалась с применением других современных методов.

Пример 4: Получение олигодендроцитов и их предшественников:

OPTi-OX hPSC, несущие индуцибельный SOX10, либо по отдельности, либо в комбинации с OLIG2, образуют бицистронный экспрессионный кластер. Хотя клетки, индуцированные только одним SOX10, надежно экспрессировали маркер O4 предшественников олигодендроцитов (OPC), после индуцирования в течение 10 дней, однако, эти клетки больше не дифференцировались в миелин-экспрессирующие клетки, и постепенно погибали. В противоположность этому, клетки, экспрессирующие два OLIG2-SOX10, легко превращались из O4-позитивной стадии клеток-предшественников в зрелые клетки CNP/MBP-позитивного фенотипа через 20 дней после индуцирования. Кроме того, дополнительный анализ на экспрессию маркерного белка подтвердил, что OPTi-OLIG2-SOX10 hPSC индуцируются в среде для олигодендроцитов (Douvaras et al., 2014), в которую были добавлены митогены PDGF $\alpha\alpha$ и FGF2, и сначала проходили стадию, подобную OPC-стадии, в которой они в высокой степени пролиферировались и коэкспрессировали PDGFRA, A2B5 и O4. Эти клетки были в высокой степени пролиферативными и могли сохраняться по меньшей мере в трех пассажах (фигура 12b), благодаря их культивированию в присутствии митогенов. Поэтому, авторы называли эти клетки iOPC, то есть, индуцированные OPC. Примечательно то, что после удаления митогенов и в непрерывном присутствии доксициклина, i-OPC легко дифференцировались приблизительно за одну неделю в зрелые олигодендроциты, экспрессирующие главные миелиновые белки CNP, PLP, MAG, MOG и MBP (фиг. 12c-12d), которые были способны образовывать миелиновые «ножницы» (данные не приводятся). В целом, эти результаты показали, что настоящее изобретение позволяет разработать новый, надежный и быстрый протокол прямого перепрограммирования hPSC с образованием олигодендроцитов и их предшественников.

Таблица 1: Систематизация результатов генотипирования:

Локус	Клеточная линия	Трансген	# собранных клонов	# клонов без интеграции мишеней (a)	# клонов НЕТ+нежелательные мишени (b)	# клонов НОМ+нежелательные мишени (b)	# клонов НЕТ (c)	# клонов НОМ (c)	Эффективность без нежелательных мишеней [%] (d)	Общая эффективность [%] (e)
ROSA26	H9	rtTA	23/27/60*	2/3/1*	7/13/36*	5/3/6*	8/8/14*	1/0/3*	39/30/28*	91/89/98*
ROSA26	iPSC	rtTA	48	8	11	2	25	2	56	83
AAVS1	H9	EGFP	12/12/24*	2/1/2*	0/0/0*	4/5/11*	0/1/4*	6/5/7*	50/50/46*	83/92/92*
AAVS1	H9	NGN2	6	0	0	0	0	6	100	100
AAVS1	iPSC	NGN2	3	0	0	2	1	0	33	100
AAVS1	H9	MYOD1	12	2	0	3	0	7	58	75
AAVS1	iPSC	MYOD1	3	0	1	1	0	1	33	100

(a) Некорректный таргетинг: Отсутствие какого-либо явного таргетинга (отсутствие полос в 5'- и 3'-интегрирующей ПЦР и присутствие полосы дикого типа в ПЦР локуса) или наличие таргетинга, но наличие неправильного размера продукта 5'- и 3'-интегрирующей ПЦР.

(b) Коррекция интеграции мишени с дополнительной случайной интеграцией плазмиды (полосы в ПЦР 3'-остова).

(c) Коррекция интеграции мишени (НЕТ, гетерозиготный; НОМ, гомозиготный).

(d) Процент клонов с правильной интеграцией мишени (без дополнительной нежелательной интеграции)

(e) Процент клонов с правильной интеграцией мишени (с дополнительной интеграцией или без дополнительной нежелательной интеграции).

* Три цифры представляют собой числа для трех других экспериментов по таргетингу в hESC.

Таблица 2: Список праймеров, используемых для ПЦР-генотипирования

Локус	Типы ПЦР	Сайт связывания с праймером	Последовательность праймера
hROSA2 6	ПЦР локуса	Геном (5')	GAGAAGAGGCTGTGCTTCGG
		Геном (3')	ACAGTACAAGCCAGTAATGGAG
	5'-INT-ПЦР	Геном (5')	GAGAAGAGGCTGTGCTTCGG
		Акцептор сплайсинга	AAGACCGCGAAGAGTTTGTCC
	3'-INT-ПЦР	rtTA	GAAACTCGCTCAAAAGCTGGG
		Геном (3')	ACAGTACAAGCCAGTAATGGAG
	3'-BB- ПЦР	rtTA	GAAACTCGCTCAAAAGCTGGG
		Остов вектора (3')	TGACCATGATTACGCCAAGC
AAVS1	ПЦР локуса	Геном (5')	CTGTTTCCCCTTCCCAGGCAGGTCC
		Геном (3')	TGCAGGGGAACGGGGCTCAGTCTGA
	5'-INT-ПЦР	Геном (5')	CTGTTTCCCCTTCCCAGGCAGGTCC
		Пуромицин	TCGTCGCGGGTGGCGAGGCGCACCG
	3'-INT-ПЦР	Индукцибельный кластер	Специфическая последовательность индуцибельного кластера
		Геном (3')	TGCAGGGGAACGGGGCTCAGTCTGA
	3'-BB-ПЦР	Индукцибельный кластер	Специфическая последовательность индуцибельного кластера
		Остов вектора (3')	ATGCTTCCGGCTCGTATGTT

Таблица 3: Список праймеров, используемых для количественной ПЦР

Ген	Ориентация	Последовательность праймера
CNP	Прямая	TCCTCATCATGAACAGAGGCTT
	Обратная	AAACTGCAGCTCAGGCTTGT
DES	Прямая	CCAACAAGAACAACGACGCC
	Обратная	ATCAGGGAATCGTTAGTGCCC
DMD	Прямая	TGGTGGGAAGAAGTAGAGGACT

	Обратная	TGCTGCTTCCCAAACCTTAGA
EGFP	Прямая	CCCGACAACCACTACCTGAG
	Обратная	GTCCATGCCGAGAGTGATCC
FOXG1	Прямая	TGCCAAGTTTTACGACGGGA
	Обратная	GGGTGGAAGAAGACCCCTG
GRIA4	Прямая	GGCCAGGGAATTGACATGGA
	Обратная	AACCAACCTTTCTAGGTCCTGTG
HMBS (PBGD)	Прямая	ATTACCCCGGGAGACTGAAC
	Обратная	GGCTGTTGCTTGGACTTCTC
MAG	Прямая	CAGAAGACGTCCCCAACTCA
	Обратная	CCTCGGGAGGCTGAAATCATAA
MAP2	Прямая	AGACTGCAGCTCTGCCTTTAG
	Обратная	AGGCTGTAAGTAAATCTTCCTCC
MBP	Прямая	TGGTGATGGAGATGTCAAGCAGGT
	Обратная	GCTGTGGTTTGGAACGAGGTTGT
MOG	Прямая	AGAGATAGAGAATCTCCACCGGA
	Обратная	TGATCAAGGCAACCAAGGGTC
MYH1	Прямая	CACACTAGTTTCACAGCTCTCG
	Обратная	CAGGGCACTCTTGGCCTTTA
MYH2	Прямая	GGAAGCTCTGGTGTCTCAGTT
	Обратная	CAGGGCGTTCTTGGCTTTTAT
MYH3	Прямая	GCTGCATACCCAGAACACCA
	Обратная	CCCTGCTGGCATCTTCTACC
MYH4	Прямая	TCGCATTTGTCAGCCAAGGG
	Обратная	TGAAACCCAGGATGTCCACAG
MYH7	Прямая	GAGACTGTCGTGGGCTTGTA
	Обратная	GCCCTTCTCAATAGGCGCATC
MYH8	Прямая	TGAAGCAGATAGCAGCGCGA
	Обратная	CGTACGAAGTGAGGGTGTGT
MYOD1 (эндо)	Прямая	GCCGCTTTCCTTAACCACAA
	Обратная	CTGAATGCCACCCACTGTC
MYOD1	Прямая	CGACGGCATGATGGACTACA
	Обратная	TAGTAGGCGCCTTCGTAGCA
NANOG	Прямая	AGCAGATGCAAGAACTCTCCAA

	Обратная	TGAGGCCTTCTGCGTCACAC
NEUROG2 (NGN2)	Прямая	TGTTTCGTCAAATCCGAGACCT
	Обратная	CGATCCGAGCAGCACTAACA
PAX6	Прямая	CGAGATTTTCAGAGCCCCATA
	Обратная	AAGACACCACCGAGCTGATT
PLP	Прямая	AACAGCTGAGTTCCAAATGACC
	Обратная	ACGGCAAAGTTGTAAGTGGC
POU3F2 (BRN2)	Прямая	ACCCGCTTTATCGAAGGCAA
	Обратная	CCTCCATAACCTCCCCCAGA
POU5F1 (OCT4)	Прямая	GTGGAGGAAGCTGACAACAA
	Обратная	ATTCTCCAGGTTGCCTCTCA
RYR1	Прямая	CAATCGCCAGAACGGAGAGA
	Обратная	GTCGTGTTCCCTGTCTGTGT
SLC17A6 (VGLUT2)	Прямая	GTAGACTGGCAACCACCTCC
	Обратная	CCATTCCAAAGCTTCCGTAGAC
SYP	Прямая	ACCTCGGGACTCAACACCTCGG
	Обратная	GAACCACAGGTTGCCGACCCAG
SYN1	Прямая	CCCTGGGTGTTTGCCCAGAT
	Обратная	ACCACGGGGTACGTTGTACT
TUBB3	Прямая	CAACCAGATCGGGGCCAAGTT
	Обратная	CCGAGTCGCCCACGTAGTT

Таблица 4: Список антител

Антиген	Виды	Изотип	Клональность	Компания	Но. по каталогу.	Разведение
A2B5	мышь	IgM	моноклональное	Millipore	MAB312	1:300
ACTN2 (α -актинин)	мышь	IgG1	моноклональное	Sigma	A7811	1:200
BrdU	мышь	IgG1	моноклональное	BD Bio	347580	1:100
CNP	мышь	IgG1	моноклональное	Abcam	ab6319	1:500
DES (десмин)	кролик	IgG	моноклональное	Abcam	ab32362	1:500
EOMES	кролик	IgG	поликлональное	Abcam	ab23345	1:200
MAG	мышь	IgG1	моноклональное	Abcam	ab89780	1:400
MAP2	мышь	IgG1	моноклональное	Sigma	M4403	1:200

MBP	крыса	IgG2a	моноклональное	Millipore	MAB386	1:200
MYOD1	кролик	IgG	моноклональное	Abcam	ab133627	1:250
MYOG (миогенин)	мышь	IgG1	моноклональное	DSHB	F5D	1:100
MYOG (миогенин)	кролик	IgG	моноклональное	Abcam	ab124800	1:500
MYH (тяжелые цепи миозина)	мышь	IgG2b	моноклональное	DSHB	MF20	1:100
MYH-PE	мышь	IgG2b	моноклональное	BD Biosc.	564408	1:20 (поток)
NANOG	коза	IgG	поликлональное	R&D	AF1997	1:200
NCAM	мышь	IgG1	моноклональное	DSHB	5.1H11	1:100
O4	мышь	IgM	моноклональное	R&D	MAB1326	1:1000
NKX2.5	кролик	IgG	поликлональное	Santa Cruz	sc14033	1:200
OCT4	мышь	IgG2b	моноклональное	Santa Cruz	sc5279	1:200
PAX3	мышь	IgG2a	моноклональное	DSHB	Pax3	1:100
PAX6	мышь	IgG1	моноклональное	DSHB	PAX6	1:100
PAX7	мышь	IgG1	моноклональное	DSHB	PAX7	1:100
PLP	кролик	IgG	моноклональное	Abcam	Ab183493	1:2000
TNNT2 (тропонин T)	мышь	IgG2a	моноклональное	DSHB	CT3	1:100
TTN (титин)	мышь	IgM	моноклональное	DSHB	9D10	1:100
TetR (репрессор tet)	мышь	IgG1	моноклональное	Clontech	631131	1:1000 (WB)
TUBA4A (α 4- тубулин)	мышь	IgG1	моноклональное	Sigma	T6199	1:10000 (WB)
TUBB3 (β III- тубулин)	мышь	IgG1	моноклональное	Millipore	MAB1637	1:1000
VGLUT1	коза	IgG	поликлональное	Abcam	ab104899	1:500

Пример 5: Система индуцибельного нокдауна TET-ON

Разработка стратегии оптимизированного индуцибельного нокдауна в hPSC.

Авторами были получены hESC-линии, в которых трансен EGFP может быть подвергнут сайленсингу по индуцибельному механизму (фиг. 14B). Для этого, авторами было проведено: (1) встраивание экспрессионного кластера CAG-tetR в локус ROSA26; и (2) встраивание трансгена CAG-EGFP плюс индуцибельного кластера кшРНК EGFP в локус AAVS1 (фиг. 14A, B). Для экспрессии более высоких уровней белка tetR необходимо более сильное подавление экспрессии кшРНК в отсутствие тетрациклина. Для этого, авторами были получены мультипараметрическая РНК и оптимизированная по кодонам бактериальная кДНК tetR, и полученную оптимизированную по кодомам tetR (OPTtetR) использовали для получения новых hESC-линий с индуцибельным нокдауном EGFP (фиг. 14B). Эта модификация давала в десять раз больше увеличение уровня экспрессии tetR по сравнению со стандартной последовательностью (STDtetR; фиг. 14D). Кроме того, гомозиготная экспрессия OPTtetR была достаточной для полного предотвращения утечки кшРНК, но при этом, полностью сохраняла эффективное индуцирование нокдауна (фиг. 14C). Следует отметить, что индуцируемый нокдаун был быстрым, обратимым и дозозависимым (фиг. 14E, F). И наконец, индуцибельные hESC имели нормальный кариотип (данные не приводятся), что свидетельствует о том, что геномная инженерия, необходимая для создания этих линий, не влияет на их генетическую стабильность.

На основе этих обнадеживающих результатов, авторами была дополнительно подтверждена эффективность этого способа в отношении эндогенных генов благодаря получению hESC, несущих индуцибельные кшРНК против POU5F1/OCT4 или B2M (данные не приводятся). Примечательно то, что все проанализированные сублинии (6 для каждого гена) показали надежный индуцибельный нокдаун без какой-либо значительной утечки кшРНК. Титрование тетрациклина позволило идентифицировать концентрации, оптимальные для частичного или полного нокдауна OCT4. Как и ожидалось, значительное снижение уровня OCT4 приводило, в частности, к потере плюрипотентности и индуцированию

нейроэктодермы и маркеров энтодермы на конечной стадии. Аналогичные результаты были получены с использованием 20 дополнительных сублиний hESC с индуцибельным нокдауном OCT4, что подтверждает надежность и воспроизводимость этого способа. Важно отметить, что продуцирование hESC с сильным и жестко регулируемым нокдауном было настолько эффективным, что анализы на фенотип можно было проводить сразу после отбора на антибиотики на смешанных популяциях клеток, и таким образом, полностью избежать необходимости сбора отдельных колоний для выделения клонов. В целом, эти результаты показали, что нацеливание на два GSH вместе с оптимизированной системой индуцибельного нокдауна представляет собой мощный способ регуляции экспрессии генов в hPSC. Этот способ далее будет сокращенно обозначаться OPTiKD, например, оптимизированный индуцибельный нокдаун (фиг. 14A).

Пример 6

Способность к нокдауну генов в различных дифференцированных клетках будет иметь значительное преимущество по сравнению с уже известными системами индуцибельного нокдауна генов. Для тщательной проверки такой возможности, авторами была проанализирована эффективность стратегии OPTiKD в отношении нокдауна трансгена EGFP в hPSC, дифференцированных в три зародышевых слоя, а также в панель из тринадцати полностью дифференцированных клеток (фиг. 15A). Для обоих способов, кол.ПЦР-анализы продемонстрировали сильный и индуцибельный нокдаун транскриптов EGFP во всех протестированных клеточных линиях дифференцировки (фиг. 17). Наблюдение под микроскопом подтвердило надежное снижение уровня экспрессии белка EGFP, а проточная цитометрия показала снижение степени флуоресценции EGFP более, чем на 70% для большинства линий дифференцировки (данные не приводятся).

Пример 7

Разработка оптимизированной индуцибельной стратегии нокаута CRISPR/Cas9 в hPSC.

Авторы сосредоточили свое внимание на разработке стратегии индуцибельного нокаута. Современные методы индуцибельного

нокаута CRISPR/Cas9 основаны на сверхэкспрессии Cas9 в присутствии конститутивно экспрессируемой рРНК в зависимости от условий. В этом случае, регуляция сверхэкспрессии Cas9 достигалась методом TET-ON, в котором, после обработки доксициклином, обратный трансаактиватор, регулируемый тетрациклином (rtTA), активирует промотор Pol II-зависимого элемента, реагирующего на тетрациклин (TRE) (гибрид между множеством оперонов TET и минимальным промотором CMV). Хотя эта стратегия TET-ON была успешно применена для человеческих клеток определенного типа, однако, авторами было обнаружено, что такая индуцибельная система подвергается сайленсингу в процессе дифференцировки hESC во множество клеточных линий дифференцировки (включая кардиомиоциты, гепатоциты и клетки гладких мышц), даже после нацеливания на AAVS1 GSH (данные не приводятся). Авторы исследовали возможность разработки альтернативного и усовершенствованного способа путем объединения конститутивно экспрессируемого Cas9, запускаемого промотором CAG, с индуцибельным кластером рРНК на основе метода, разработанного для индуцибельной экспрессии кшРНК (фиг. 18А,В). Следовательно, авторами были получены hESC-линии, в которых ген флуоресцентного репортера может быть подвергнут нокауту по индуцибельному механизму (фиг. 18С). Для этого, авторами был проведен таргетинг репортерного гена ROSA26-EGFPd2 в hESC, содержащих индуцибельную рРНК EGFP и конститутивный Cas9 в локусе AAVS1, где каждый трансген был интегрирован в один из двух аллелей. Этот метод двойного нацеливания был быстрым (<2 недель) и эффективным (>90% линий, содержащих оба трансгена). Примечательно то, что при культивировании отдельных клональных сублиний в присутствии тетрациклина авторы наблюдали снижение уровня экспрессии EGFPd2 во всех линиях-мишенях, а клетки, гомозиготные по EGFPd2, показали почти гомогенную потерю по меньшей мере одной копии репортерного гена по меньшей мере через пять дней после индуцирования тетрациклином (как было продемонстрировано по 50% снижению интенсивности флуоресценции EGFPd2). Длительная обработка тетрациклином приводила к прогрессирующей полной потере флуоресценции EGFPd2 до 75%

клеток, гомозиготных по EGFPd2 (данные не приводятся). Интересно отметить, что коэкспрессия двух или трех копий одного и того же кластера рНК EGFP из одного и того же локуса AAVS1 была достаточной для значительного увеличения скорости и повышения эффективности индуцибельного нокаута EGFPd2 во всех проанализированных клональных сублиниях. Так, например, одновременное индуцирование трех копий одной и той же рНК приводило к значительной 95% эффективности нокаута после обработки тетрациклином. Важно отметить, что в клетках hESC с индуцибельным нокаутом EGFPd2 не наблюдалось какого-либо значимого снижения ни числа EGFPd2-позитивных клеток, ни флуоресценции после длительного культивирования в отсутствие тетрациклина, даже при использовании нескольких копий рНК. Это свидетельствует о том, что индуцибельная экспрессия рНК была жестко регулируемой. И наконец, тестирование дополнительных рНК против EGFPd2 показало, что скорость и эффективность индуцибельного нокаута в значительной степени зависит от рНК. Действительно, оптимальная последовательность давала 90% нокаут уже после 2 дней индуцирования. Следует отметить, что наиболее эффективная рНК также давала нерегулируемый нокаут EGFPd2, но это ограничение может быть легко устранено просто путем добавления второго оперона ТЕТ к индуцибельному промотору HI, для гарантии еще более строгой регуляции транскрипции. В целом, эти результаты показали, что система нокадауна может быть легко переориентирована на сохранение индуцибельной экспрессии рНК и достижение жестко регулируемой активности CRISPR/Cas9 в широком диапазоне эффективностей рНК. Насколько известно авторам, этот способ представляет собой первый способ применения CRISPR/Cas9, основанный на индуцибельной экспрессии рНК в зависимости от выбранных условий.

СПИСОК ПОСЛЕДОВАТЕЛЬНОСТЕЙ

<110> КЕМБРИДЖ ЭНТЕРПРАЙЗ ЛИМИТЕД

<120> Регулируемая транскрипция

<130> P8072WO1

<150> GB1619876.4

<151> 2016-11-24

<160> 86

<170> PatentIn version 3.5

<210> 1

<211> 19

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Нуклеотидная последовательность TetO2 19

<400> 1

tccctatcag tgatagaga

19

<210> 2

<211> 1797

<212> ДНК

<213> Homo sapiens

<400> 2

gctcgaaacc ggacggagcc attgctctcg cagagggagg agcgcttccg gctagcctct 60

tgtcgccgat tggccgtttc tcctcccgcc gtgtgtgaaa acacaaatgg cgtattctgg 120

ttggagtaaa gctcctgtca gttacgccgt cgggagtagc cagccgctta gcgactctcg 180

cgttgcccc tgggtggggc gggtaggtag gtgggggtgta gagatgctgg gtgtgcgggc 240

gcggccggcc tcctgcggcg ggaggggagg gtcagtgaaa tcggctctgg cgcgggcgtc 300

ctccaccct ccccttcctt cgggggagtc ggtttaccgc ccgcctgctt gtcttcgaca 360

cctgattggc tgtcgaagct gtgggaccgc gcccttgcta ctggctcgag tctcacatga 420

gcgaaaccac tgcgcggggc gcgggggtgg cggggaggcg ggcgttggtc cggctctccc 480

cgaggccgag cgccgcagtg tctggccccg cggccctgcg caacgtggca ggaagcgcgc 540

gctggaggcg ggggcgggct gccggccgag acttctggat ggcggcggcc gcggctccgc 600

cccgggttcc caccgcctga agggcgagac aagcccgacc tgctacaggc actcgtgggg 660

gtgggggagg agcgggggtc ggtccggctg gtttgtgggt gggaggcgct tgttctccaa 720

aaaccggcgc gagctgcaat cctgaggag ctgcgggtgga ggaggtggag agaaggccgc 780

acccttctgg gcagggggag gggagtgcgc caataccttt atgggagttc tctgctgcct 840

cccgtcttgt aaggaccgcc ctgggcctgg aagaagccct ccctcctttc ctccctcgct 900

gatctcgtca tcgcctccat gtcgagtcgc ttctcgatta tgggcgggat tcttttgcct 960

aggcttaagg ggctaacttg gtccctgggc gttgccctgc aggggagtga gcagctgtaa	1020
gatttgaggg gcgactccga ttagttttatc ttcccacgga ctagagttgg tgtcgaggtt	1080
attgtaataa ggggtggggta gggaaatgga gcttagtcat tcacctgggg ctgattttat	1140
gcaacgagac tgcggattat cactacttat catttttggg gcatttttct agagacagac	1200
ataaagcatg atcacctgag ttttatacca tttgagaccc ttgctgcacc accaaagtgt	1260
agcatcaggt taaatcttaa tagaaaaatt ttagcttttg cttgagaaac cagtgccttc	1320
ctccctcacc ctctctcccc aggctctcta cccctttgca tccctaccag gcactttagc	1380
aactctcact cataacttgat cccatttttc atttggtgta cttgctcctc tagtattcag	1440
acatagcact agctttctcc ctctcttgat cttgggtagc ctgggtgtctc gcgaaaccag	1500
acagattggg tccaccacaa attaaggctt gagctggggc ttgactctta cccagcagtg	1560
cttttattcc tccctagttc acgttcttaa atgtttatct tgattttcat tttatccttt	1620
ttccttagct gggattctgt ccctgaccgt cttcacagtc caggtgatct tgactactgc	1680
tttacagaga attggatctg aggttaggca acatctccct ttttcttcct ctaaatacct	1740
ctcatttctg ttcttaccag ttagtaactg atctcagatg cctgtgtgat agcttcc	1797

<210> 3
 <211> 651
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> STDtetR-nls

<400> 3	
atgccaaaaa agaagaggaa ggtatctaga ttagataaaa gtaaagtgat taacagcgca	60
ttagagctgc ttaatgaggt cggaatcgaa ggtttaacaa cccgtaaact cgcccagaag	120
ctaggtgtag agcagcctac attgtattgg catgtaaaaa ataagcgggc tttgctcgac	180
gccttagcca ttgagatggt agataggcac catactcact tttgcccttt agaaggggaa	240
agctggcaag attttttacg taataacgct aaaagtttta gatgtgcttt actaagtcac	300
cgcgatggag caaaagtaca tttaggtaca cggcctacag aaaaacagta tgaaactctc	360
gaaaatcaat tagccttttt atgccaacaa ggtttttcac tagagaatgc attatatgca	420
ctcagcgctg tggggcattt tacttttaggt tgcgtattgg aagatcaaga gcacaaagtc	480
gctaaagaag aaagggaaac acctactact gatagtatgc cgccattatt acgacaagct	540
atcgaattat ttgatcacca aggtgcagag ccagccttct tattcggcct tgaattgatc	600
atatgcggat tagaaaaaca acttaaagtgt gaaagtgggt ctccgcggta a	651

<210> 4
 <211> 216

<212> Белок
<213> Искусственная последовательность

<220>
<223> STDtetR-nls

<400> 4

Met Pro Lys Lys Lys Arg Lys Val Ser Arg Leu Asp Lys Ser Lys Val
1 5 10 15

Ile Asn Ser Ala Leu Glu Leu Leu Asn Glu Val Gly Ile Glu Gly Leu
20 25 30

Thr Thr Arg Lys Leu Ala Gln Lys Leu Gly Val Glu Gln Pro Thr Leu
35 40 45

Tyr Trp His Val Lys Asn Lys Arg Ala Leu Leu Asp Ala Leu Ala Ile
50 55 60

Glu Met Leu Asp Arg His His Thr His Phe Cys Pro Leu Glu Gly Glu
65 70 75 80

Ser Trp Gln Asp Phe Leu Arg Asn Asn Ala Lys Ser Phe Arg Cys Ala
85 90 95

Leu Leu Ser His Arg Asp Gly Ala Lys Val His Leu Gly Thr Arg Pro
100 105 110

Thr Glu Lys Gln Tyr Glu Thr Leu Glu Asn Gln Leu Ala Phe Leu Cys
115 120 125

Gln Gln Gly Phe Ser Leu Glu Asn Ala Leu Tyr Ala Leu Ser Ala Val
130 135 140

Gly His Phe Thr Leu Gly Cys Val Leu Glu Asp Gln Glu His Gln Val
145 150 155 160

Ala Lys Glu Glu Arg Glu Thr Pro Thr Thr Asp Ser Met Pro Pro Leu
165 170 175

Leu Arg Gln Ala Ile Glu Leu Phe Asp His Gln Gly Ala Glu Pro Ala
180 185 190

Phe Leu Phe Gly Leu Glu Leu Ile Ile Cys Gly Leu Glu Lys Gln Leu
195 200 205

Lys Cys Glu Ser Gly Ser Pro Arg
210 215

<210> 5
<211> 651
<212> ДНК
<213> Искусственная последовательность

<220>
<223> OPTtetR

<400> 5
atgcccaaga aaaagcggaa ggtgtcccgg ctggacaaga gcaaagtgat caacagcgcc 60
ctggaactgc tgaacgaagt gggcatcgag ggcttgacca cccggaagct ggcccagaaa 120
ctgggctgtg aacagcccac cctgtactgg cacgtgaaga acaagcgggc cctgctggac 180
gccctggcca tcgagatgct ggaccggcac cacacacact ttgccccct ggaaggcgaa 240
agctggcagg acttcctgcg gaacaacgcc aagagcttca gatgcgccct gctgagccac 300
cgggacggcg ccaaagtgca cctgggcacc agaccaccg agaagcagta cgagacactg 360
gaaaaccagc tggccttcct gtgccagcag ggcttcagcc tggaaaacgc cctgtacgcc 420
ctgagcgccg tgggccactt taccctgggc tgcgtgctgg aagatcagga acaccaggtc 480
gccaaagagg aaagagagac acccaccacc gacagcatgc cccccctgct gagacaggcc 540
atcgagctgt tcgatcatca aggcgccgag cccgccttcc tgttcggcct ggaactgatc 600
atctgcggcc tcgagaagca gctgaagtgc gagagcgggt cccccagatg a 651

<210> 6
<211> 216
<212> Белок
<213> Искусственная последовательность

<220>
<223> OPTtetR

<400> 6
Met Pro Lys Lys Lys Arg Lys Val Ser Arg Leu Asp Lys Ser Lys Val
1 5 10 15
Ile Asn Ser Ala Leu Glu Leu Leu Asn Glu Val Gly Ile Glu Gly Leu
20 25 30
Thr Thr Arg Lys Leu Ala Gln Lys Leu Gly Val Glu Gln Pro Thr Leu
35 40 45
Tyr Trp His Val Lys Asn Lys Arg Ala Leu Leu Asp Ala Leu Ala Ile
50 55 60
Glu Met Leu Asp Arg His His Thr His Phe Cys Pro Leu Glu Gly Glu
65 70 75 80
Ser Trp Gln Asp Phe Leu Arg Asn Asn Ala Lys Ser Phe Arg Cys Ala
85 90 95

Leu Leu Ser His Arg Asp Gly Ala Lys Val His Leu Gly Thr Arg Pro
100 105 110

Thr Glu Lys Gln Tyr Glu Thr Leu Glu Asn Gln Leu Ala Phe Leu Cys
115 120 125

Gln Gln Gly Phe Ser Leu Glu Asn Ala Leu Tyr Ala Leu Ser Ala Val
130 135 140

Gly His Phe Thr Leu Gly Cys Val Leu Glu Asp Gln Glu His Gln Val
145 150 155 160

Ala Lys Glu Glu Arg Glu Thr Pro Thr Thr Asp Ser Met Pro Pro Leu
165 170 175

Leu Arg Gln Ala Ile Glu Leu Phe Asp His Gln Gly Ala Glu Pro Ala
180 185 190

Phe Leu Phe Gly Leu Glu Leu Ile Ile Cys Gly Leu Glu Lys Gln Leu
195 200 205

Lys Cys Glu Ser Gly Ser Pro Arg
210 215

<210> 7
<211> 20
<212> ДНК
<213> Искусственная последовательность

<220>
<223> ПЦР локуса hROSA26

<400> 7
gagaagaggc tgtgcttcgg 20

<210> 8
<211> 22
<212> ДНК
<213> Искусственная последовательность

<220>
<223> ПЦР локуса hROSA26

<400> 8
acagtacaag ccagtaatgg ag 22

<210> 9
<211> 20
<212> ДНК
<213> Искусственная последовательность

<220>

<223>	hROSA26 5'-INT PCR	
<400>	9	
	gagaagaggc tgtgcttcgg	20
<210>	10	
<211>	21	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	hROSA26 5'-INT PCR	
<400>	10	
	aagaccgcga agagtttgtc c	21
<210>	11	
<211>	21	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	hROSA26 3'-INT PCR	
<400>	11	
	gaaactcgct caaaagctgg g	21
<210>	12	
<211>	22	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	hROSA26 3'-INT PCR	
<400>	12	
	acagtacaag ccagtaatgg ag	22
<210>	13	
<211>	21	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	hROSA26 3'-BB PCR	
<400>	13	
	gaaactcgct caaaagctgg g	21
<210>	14	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	hROSA26 3'-BB PCR	
<400>	14	
	tgaccatgat tacgccaagc	20

<210> 15
 <211> 25
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> ПЦР локуса AAVS1

 <400> 15
 ctgtttcccc ttcccaggca ggtcc 25

<210> 16
 <211> 25
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> ПЦР локуса AAVS1

 <400> 16
 tgcaggggaa cggggctcag tctga 25

<210> 17
 <211> 25
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> AAVS1 5'-INT PCR

 <400> 17
 ctgtttcccc ttcccaggca ggtcc 25

<210> 18
 <211> 25
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> AAVS1 5'-INT PCR

 <400> 18
 tcgtcgcggg tggcgaggcg caccg 25

<210> 19
 <211> 25
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> AAVS1 3'-INT PCR

 <400> 19
 tgcaggggaa cggggctcag tctga 25

<210> 20
 <211> 20

<212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> AAVS1 3'-BB PCR

 <400> 20
 atgcttccgg ctcgtatggt 20

 <210> 21
 <211> 22
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> CNP Fw

 <400> 21
 tcctcatcat gaacagaggc tt 22

 <210> 22
 <211> 20
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> CNP Rev

 <400> 22
 aaactgcagc tcaggcttgt 20

 <210> 23
 <211> 20
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> DES Fw

 <400> 23
 ссаасаагаа саасгасгсс 20

 <210> 24
 <211> 21
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> DES Rev

 <400> 24
 atcagggaat cgtagtgcc c 21

 <210> 25
 <211> 22
 <212> ДНК
 <213> Искусственная последовательность

 <220>

<223>	DMD Fw	
<400>	25	
	tggtgggaag aagtagagga ct	22
<210>	26	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	DMD Rev	
<400>	26	
	tgctgcttcc caaacttaga	20
<210>	27	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	EGFP Fw	
<400>	27	
	cccgacaacc actacctgag	20
<210>	28	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	EGFP Rev	
<400>	28	
	gtccatgccg agagtgatcc	20
<210>	29	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	FOXG1 Fw	
<400>	29	
	tgccaagttt tacgacggga	20
<210>	30	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	FOXG1 Rev	
<400>	30	
	gggttggaag aagaccctg	20

<210>	31	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	GRIA4 Fw	
<400>	31	
	ggccagggaa ttgacatgga	20
<210>	32	
<211>	23	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	GRIA4 Rev	
<400>	32	
	aaccaacctt tctaggtcct gtg	23
<210>	33	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	HMBS (PBGD) Fw	
<400>	33	
	attaccccgaggagactgaac	20
<210>	34	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	HMBS (PBGD) Rev	
<400>	34	
	ggctgttgct tggacttctc	20
<210>	35	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	MAG Fw	
<400>	35	
	cagaagacgt ccccaactca	20
<210>	36	
<211>	22	

<212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> MAG Rev

 <400> 36
 cctcgggagg ctgaaatcat aa 22

 <210> 37
 <211> 21
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> MAP2 Fw

 <400> 37
 agactgcagc tctgccttta g 21

 <210> 38
 <211> 23
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> MAP2 Rev

 <400> 38
 aggctgtaag taaatcttcc tcc 23

 <210> 39
 <211> 24
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> MBP Fw

 <400> 39
 tggatgatgga gatgtcaagc aggt 24

 <210> 40
 <211> 24
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> MBP Rev

 <400> 40
 gctgtgggttt ggaaacgagg ttgt 24

 <210> 41
 <211> 23
 <212> ДНК
 <213> Искусственная последовательность

 <220>

<223>	MOG Fw	
<400>	41	
	agagatagag aatctccacc gga	23
<210>	42	
<211>	21	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	MOG Rev	
<400>	42	
	tgatcaaggc aaccaagggt c	21
<210>	43	
<211>	22	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	МҮН1 Fw	
<400>	43	
	cacactagtt tcacagctct cg	22
<210>	44	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	МҮН1 Rev	
<400>	44	
	cagggcactc ttggccttta	20
<210>	45	
<211>	21	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	МҮН2 Fw	
<400>	45	
	ggaagctctg gtgtctcagt t	21
<210>	46	
<211>	21	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	МҮН2 Rev	
<400>	46	
	cagggcggttc ttggctttta t	21

<210>	47	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	МҮН3 Fw	
<400>	47	
	gctgcatacc сагаасасса	20
<210>	48	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	МҮН3 Rev	
<400>	48	
	ccctgctggc атттттасс	20
<210>	49	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	МҮН4 Fw	
<400>	49	
	tcgcatttgt сагссааggg	20
<210>	50	
<211>	21	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	МҮН4 Rev	
<400>	50	
	tgaaaccсag gatgtccaca g	21
<210>	51	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	МҮН7 Fw	
<400>	51	
	gagactgtcg tgggcttgta	20
<210>	52	
<211>	21	

<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	МҮН7 Rev	
<400>	52	
	gcccttctca ataggcgcat c	21
<210>	53	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	МҮН8 Fw	
<400>	53	
	tgaagcagat agcagcgcgga	20
<210>	54	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	МҮН8 Rev	
<400>	54	
	cgtacgaagt gaggggtgtgt	20
<210>	55	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	MYOD1 (эндо) Fw	
<400>	55	
	gccgctttcc ttaaccasaa	20
<210>	56	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	MYOD1 (эндо) Rev	
<400>	56	
	ctgaatgcc acccactgtc	20
<210>	57	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		

<223>	MYOD1 Fw	
<400>	57	
	cgacggcatg atggactaca	20
<210>	58	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	MYOD1 Rev	
<400>	58	
	tagtaggcgc cttcgtagca	20
<210>	59	
<211>	22	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	NANOG Fw	
<400>	59	
	agcagatgca agaactctcc aa	22
<210>	60	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	NANOG Rev	
<400>	60	
	tgaggccttc tgcgtcacac	20
<210>	61	
<211>	21	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	NEUROG2 (NGN2) Fw	
<400>	61	
	tgttcgtcaa atccgagacc t	21
<210>	62	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	NEUROG2 (NGN2) Rev	
<400>	62	
	cgatccgagc agcactaaca	20

<210>	63	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	PAX6 Fw	
<400>	63	
	cgagatttca gagccccata	20
<210>	64	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	PAX6 Rev	
<400>	64	
	aagacaccac cgagctgatt	20
<210>	65	
<211>	22	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	PLP Fw	
<400>	65	
	aacagctgag ttccaaatga cc	22
<210>	66	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	PLP Rev	
<400>	66	
	acggcaaagt tgtaagtggc	20
<210>	67	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	POU3F2 (BRN2) Fw	
<400>	67	
	accgcgttta tcgaaggcaa	20
<210>	68	
<211>	20	

<212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> POU3F2 (BRN2) Rev

 <400> 68
 cctccataac ctcccccaga 20

 <210> 69
 <211> 20
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> POU5F1 (OCT4) Fw

 <400> 69
 gtggaggaag ctgacaaca 20

 <210> 70
 <211> 20
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> POU5F1 (OCT4) Rev

 <400> 70
 attctccagg ttgcctctca 20

 <210> 71
 <211> 20
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> RYR1 Fw

 <400> 71
 caatcgccag aacggagaga 20

 <210> 72
 <211> 20
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> RYR1 Rev

 <400> 72
 gtcgtgttcc ctgtctgtgt 20

 <210> 73
 <211> 20
 <212> ДНК
 <213> Искусственная последовательность

 <220>

<223>	SLC17A6 (VGLUT2) Fw	
<400>	73	
	gtagactggc aaccacctcc	20
<210>	74	
<211>	22	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	SLC17A6 (VGLUT2) Rev	
<400>	74	
	ccattccaaa gcttccgtag ac	22
<210>	75	
<211>	22	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	SYP Fw	
<400>	75	
	acctcgggac tcaacacctc gg	22
<210>	76	
<211>	22	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	SYP Rev	
<400>	76	
	gaaccacagg ttgccgaccc ag	22
<210>	77	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	SYN1 Fw	
<400>	77	
	ccctgggtgt ttgccagat	20
<210>	78	
<211>	20	
<212>	ДНК	
<213>	Искусственная последовательность	
<220>		
<223>	SYN1 Rev	
<400>	78	
	accacgggggt acgttg tact	20

<210> 79
 <211> 21
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> TUBB3 Fw

 <400> 79
 caaccagatc ggggccaagt t 21

<210> 80
 <211> 19
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> TUBB3 Rev

 <400> 80
 ccgagtcgcc cacgtagtt 19

<210> 81
 <211> 52
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> Таргетирующий вектор AAVS1 Fw

 <400> 81
 gtgatataga tccctgggtgc aggtggactg actcacctgc acctgtttta ga 52

<210> 82
 <211> 52
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> Таргетирующий вектор AAVSI Rev

 <400> 82
 actaaaacag gtgcaggtga gtcagtccac ctgcaccagg gatctctatc ac 52

<210> 83
 <211> 24
 <212> ДНК
 <213> Искусственная последовательность

 <220>
 <223> 20 п.о-метка Fw

 <220>
 <221> misc_feature
 <222> (5)..(24)
 <223> n = а или g или с или t, неизвестное или другое

<400> 83
 tcccnnnnnnn nnnnnnnnnnn nnnn 24

<210> 84
 <211> 24
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> 20 п.о.-метка Rev

<220>
 <221> misc_feature
 <222> (5)..(24)
 <223> n = а или g или с или t, неизвестное или другое

<400> 84
 aaacnnnnnnn nnnnnnnnnnn nnnn 24

<210> 85
 <211> 268
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Промотор H1 POLIII H1-TO

<400> 85
 cgaacgctga cgtcatcaac ccgctccaag gaatcgcggg ccagtggtca ctaggcggga 60
 acaccagcg cgcgtgcgcc ctggcaggaa gatggctgtg agggacaggg gagtggcgcc 120
 ctgcaatatt tgcattgtcg tatgtgttct gggaaatcac cataaacgtg aaatgtcttt 180
 ggatttggga atcttataag ttccctatca gtgatagaga tccctgggtgc aggtggactg 240
 actcacctgc acctgtttta gagctaga 268

<210> 86
 <211> 268
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Промотор H1 POLIII H1-2TO

<400> 86
 cgaacgctga cgtcatcaac ccgctccaag gaatcgcggg ccagtggtca ctaggcggga 60
 acaccagcg cgcgtgcgcc ctggcaggaa gatggctgtg agggacaggg gagtggcgcc 120
 ctgcaatatt tgcattgtcg tatgtgttct gggaaatcac cataaacgtg aaatccctat 180
 cagtgataga gacttataag ttccctatca gtgatagaga tccctgggtgc aggtggactg 240
 actcacctgc acctgtttta gagctaga 268

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ регуляции транскрипции генной последовательности в клетке, где указанный способ включает:

а) нацеленное встраивание гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

б) нацеленное встраивание индуцибельного кластера во второй генетически безопасный якорный сайт, где указанный индуцибельный кластер содержит указанную генную последовательность, функционально присоединенную к индуцибельному промотору, и где указанный промотор регулируется белком-регулятором транскрипции, где указанные первый и второй генетически безопасные якорные сайты являются различными.

2. Способ по п.1, где генная последовательность представляет собой трансен.

3. Способ по п.1, где генная последовательность кодирует некодирующую РНК.

4. Способ по любому из пп.1-3, где активность указанного белка-регулятора транскрипции регулируется экзогенно вводимым веществом.

5. Способ по любому из пп.1-4, где указанный белок-регулятор транскрипции является конститутивно экспрессируемым.

6. Способ по любому из предшествующих пунктов, где указанный белок-регулятор транскрипции выбран из: белка-активатора транскрипции, реагирующего на тетрациклин (rtTa), репрессора тетрациклина (TetR), синтетического рецептора VgEsR или гибрида белка-регулятора транскрипции и гена-переключателя и любых их производных.

7. Способ по любому из предшествующих пунктов, где указанным белком-регулятором транскрипции является rtTA или любое его производное.

8. Способ по п.5, где активность tTA регулируется тетрациклином или его производным, необязательно доксициклином.

9. Способ по п.7 или п.8, где индуцибельный промотор включает элемент, реагирующий на Tet (TRE).

10. Способ по любому из предшествующих пунктов, где указанные первый и второй геномно безопасные якорные сайты

выбраны из любых двух локусов hROSA26 и AAVS1, гена CLYBL, гена CCR5 или гена HPRT.

11. Способ по любому из предшествующих пунктов, где введение указанного гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт, осуществляют на обеих хромосомах клетки, и/или введение указанного индуцибельного кластера во второй генетически безопасный якорный сайт осуществляют на обеих хромосомах клетки.

12. Способ по любому из предшествующих пунктов, где дополнительный генетический материал встраивают в первый и/или второй геномно безопасные якорные сайты, содержащие, но необязательно, один или более из:

- a) гена-«самоубийцы»;
- b) селективного маркера;
- c) гена-репортера; и/или
- d) гена некодирующей РНК.

13. Способ по любому из предшествующих пунктов, где указанный способ осуществляют *ex vivo*.

14. Способ по любому из предшествующих пунктов, где указанная клетка выбрана из плюрипотентных стволовых клеток, соматической стволовой клетки или зрелой клетки.

15. Способ по любому из п.1 или 2, где указанный способ представляет собой перепрограммирование плюрипотентных стволовых клеток в определенные зрелые клетки.

16. Способ по любому из п.1 или 2, где указанной клеткой является плюрипотентная стволовая клетка, а указанной генной последовательностью является трансген для одного или более главных регуляторов, необязательно, фактора транскрипции.

17. Способ по п.16, где транскрипция указанной генной последовательности приводит к прямому перепрограммированию клетки в зрелую клетку определенного типа.

18. Способ по п.17, где зрелые клетки указанного типа выбраны из любых клеток следующих типов: нервных клеток, миоцитов, остеоцитов, хондроцитов, эпителиальных клеток, секреторных клеток и/или клеток крови.

19. Способ по любому из пп.14-18, где указанная

плюрипотентная стволовая клетка выбрана из индуцированных плюрипотентных стволовых клеток (iPSC) или эмбриональных стволовых клеток (ESC).

20. Способ по любому из пп.1-19, где указанная клетка представляет собой человеческую клетку.

21. Способ по п.1, где указанный способ включает встраивание генной последовательности в целях генотерапии, и где генная последовательность кодирует, но необязательно, белок, включая белок дикого типа, модифицированный белок, антиген, фермент, селективный маркер или некодирующую молекулу РНК.

22. Клетка, полученная способом по п.1.

23. Клетка с модифицированным геномом, который включает ген, кодирующий белок-регулятор транскрипции и встроенный в первый генетически безопасный якорный сайт; и который включает индуцибельный кластер, содержащий генную последовательность, функционально присоединенную к индуцибельному промотору, и встроенный во второй генетически безопасный якорный сайт, где указанный индуцибельный промотор регулируется белком-регулятором транскрипции, а указанные первый и второй сайты являются различными.

24. Применение клетки по любому из пп.22 или 23 в терапии.

25. Применение клетки по любому из п.22 или 23 для *in vitro* диагностики.

26. Применение клетки по любому из п.22 или 23 для конструирования тканей и, необязательно, выращивания мяса.

27. Способ продуцирования миоцитов из плюрипотентных стволовых клеток, где указанный способ включает стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

б) нацеленного встраивания гена MYOD1, функционально присоединенного к индуцибельному промотору, во второй генетически безопасный якорный сайт, где указанный индуцибельный промотор регулируется белком-регулятором транскрипции, и где указанные первый и второй генетически безопасные якорные сайты являются различными; и

культивирования указанных клеток в присутствии ретиноевой кислоты.

28. Способ продуцирования олигодендроцитов из плюрипотентных стволовых клеток, где указанный способ включает стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

б) нацеленного встраивания гена SOX-10, функционально присоединенного к индуцибельному промотору, во второй генетически безопасный якорный сайт, где указанный индуцибельный промотор регулируется белком-регулятором транскрипции, и где указанные первый и второй генетически безопасные якорные сайты являются различными; и

культивирования указанных клеток в присутствии ретиноевой кислоты.

29. Способ снижения уровня транскрипции и/или трансляции эндогенного гена в клетке, где указанный способ включает следующие стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, в первый генетически безопасный якорный сайт; и

б) нацеленного встраивания индуцибельного кластера во второй генетически безопасный якорный сайт; где указанный индуцибельный кластер содержит ДНК, кодирующую некодирующую последовательность РНК, функционально присоединенную к индуцибельному промотору, и где указанный промотор регулируется белком-регулятором транскрипции, а указанная некодирующая последовательность РНК подавляет транскрипцию или трансляцию эндогенного гена;

где указанные первый и второй генетически безопасные якорные сайты являются различными.

30. Способ нокаута эндогенного гена в клетке, включающий следующие стадии:

а) нацеленного встраивания гена, кодирующего белок-регулятор транскрипции, и гена, кодирующего Cas9 в первый

генетически безопасный якорный сайт; и

b) нацеленного встраивания индуцибельного кластера во второй генетически безопасный якорный сайт; где указанный индуцибельный кластер содержит руководящую РНК, функционально присоединенную к индуцибельному промотору, и где указанный промотор регулируется белком-регулятором транскрипции, а указанная последовательность рРНК нацелена на эндогенный ген;

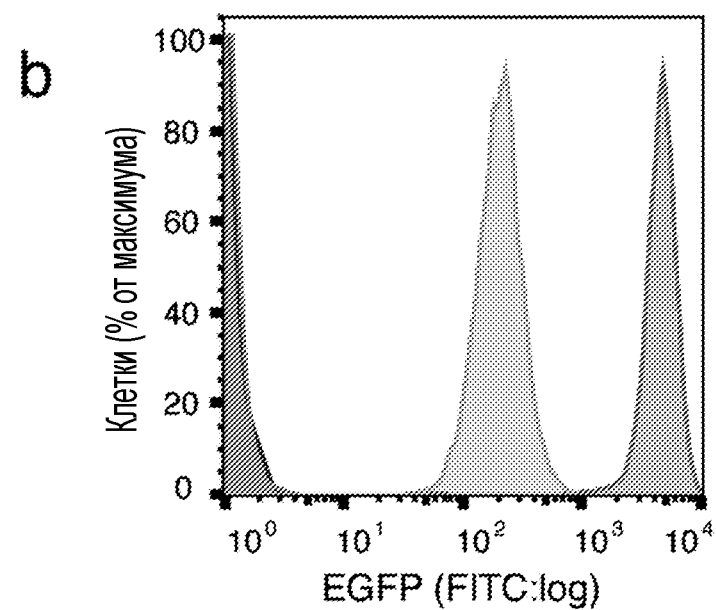
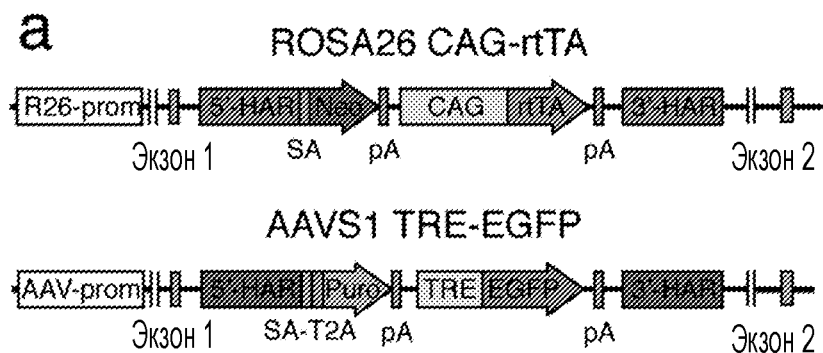
где указанные первый и второй генетически безопасные якорные сайты являются различными.

31. Способ по любому из пп.1-21 или 27-29, где другой индуцибельный кластер или трансген встраивают в другой GSH, отличающийся от первого и второго GSH.

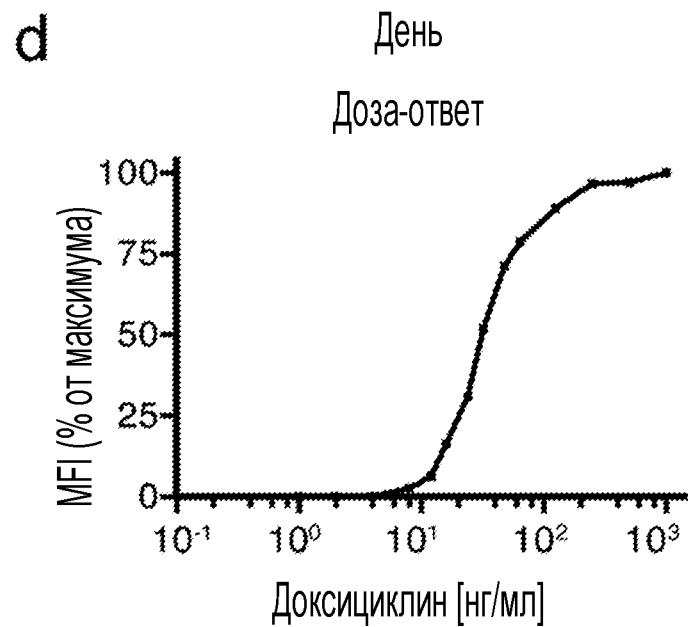
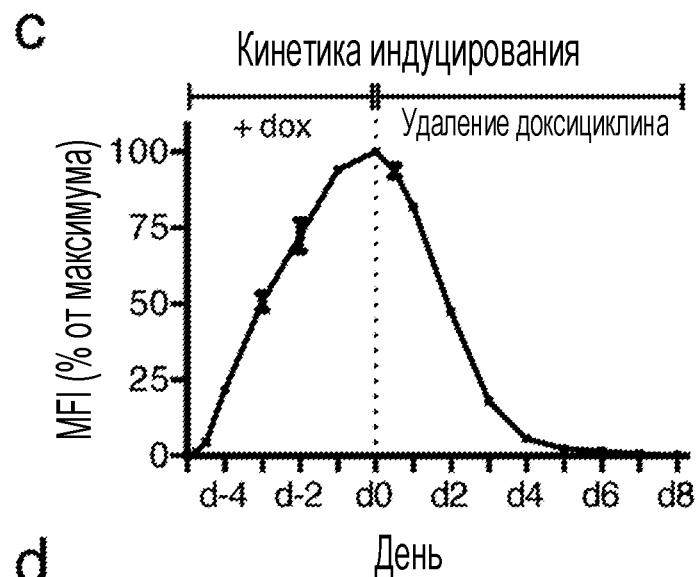
32. Оптимизированная последовательность tetR, включающая последовательность, которая на 80%, 85%, 90%, 95% или 99% гомологична последовательности SEQ ID NO: 6.

33. Способ продуцирования миоцитов из плюрипотентных стволовых клеток, экспрессирующих MYOD1, где указанный способ включает культивирование указанных клеток в присутствии ретиноевой кислоты.

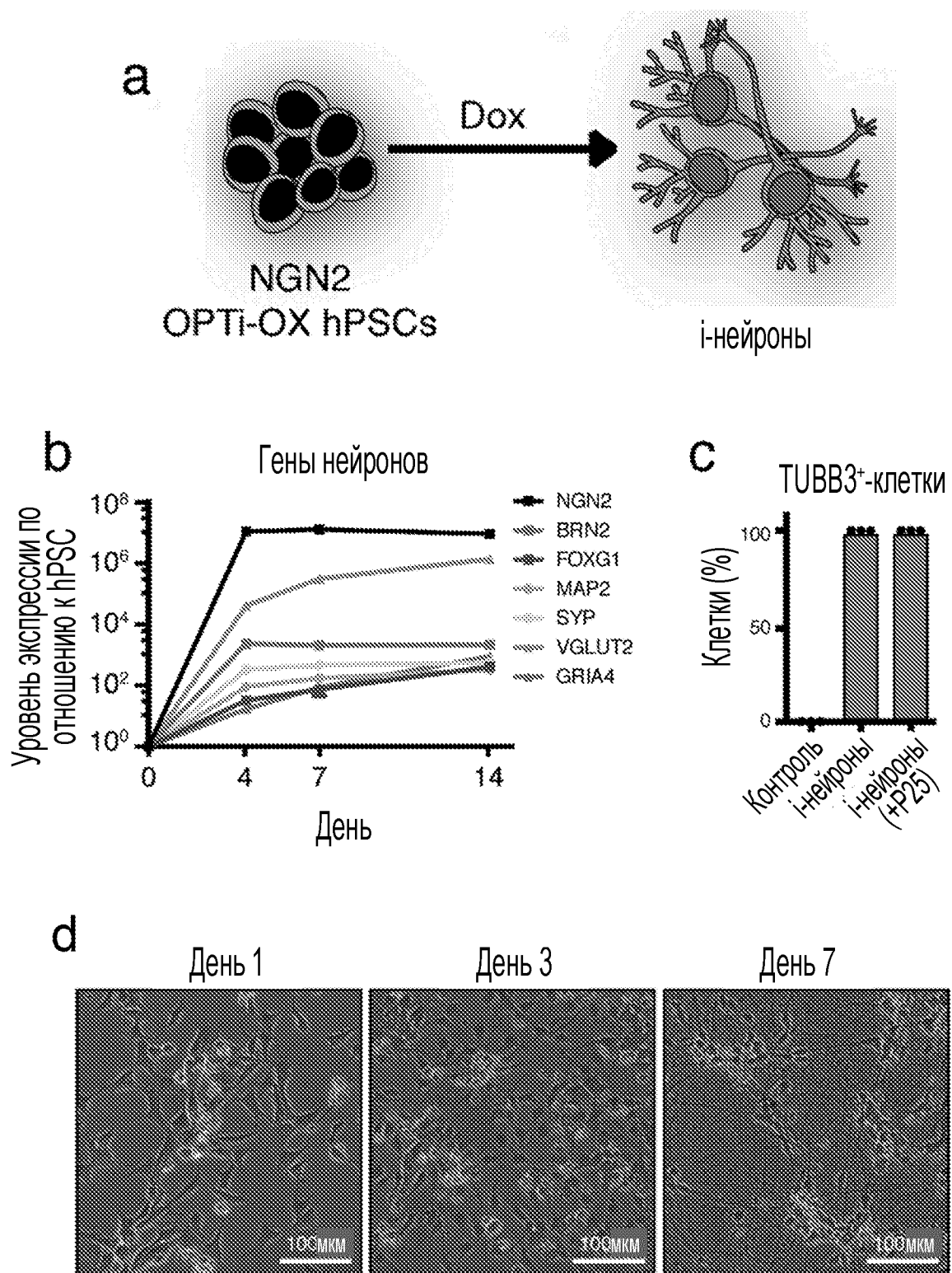
По доверенности



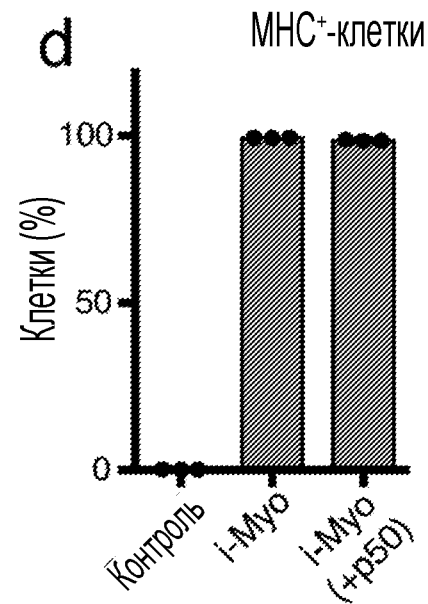
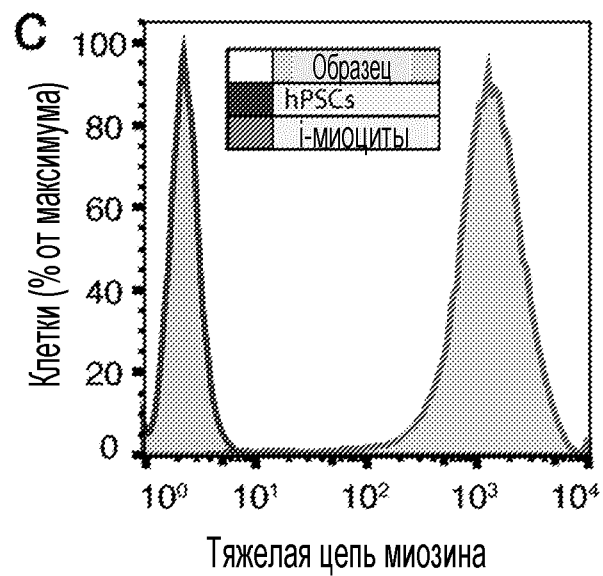
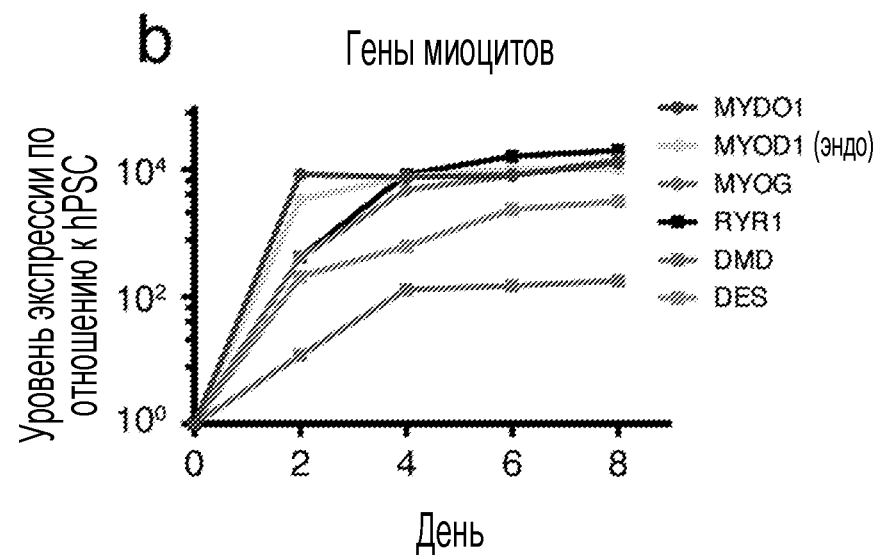
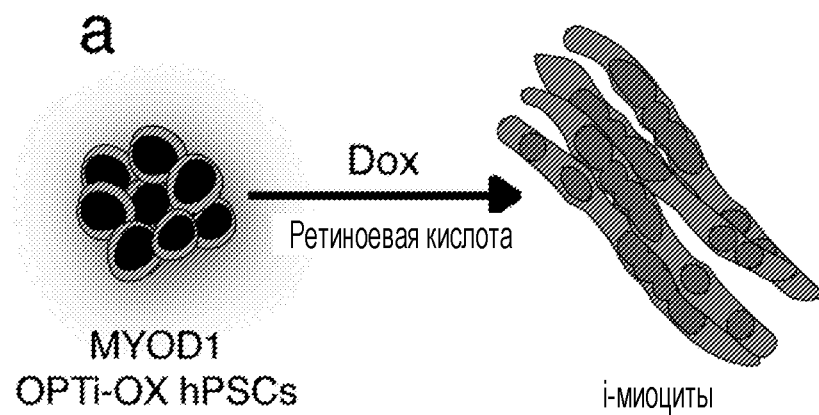
Образец	MFI	EGFP ⁺ (%)
Дикого типа	1	0
CAG-EGFP	204	99.5
i-EGFP CTR	1	0
i-EGFP DOX	4542	99.9



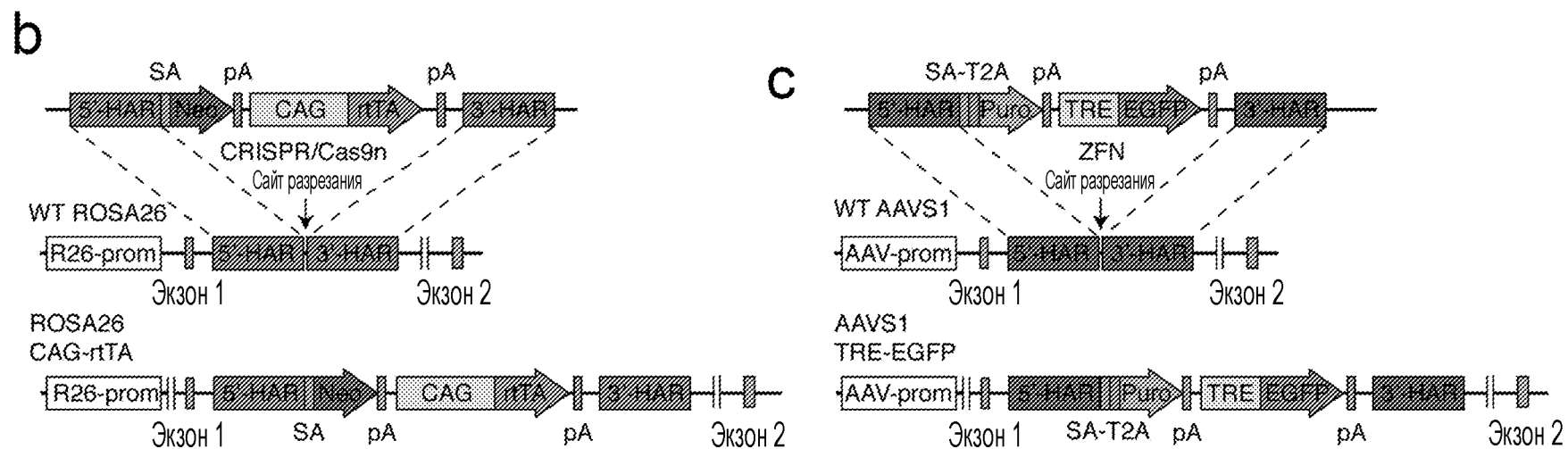
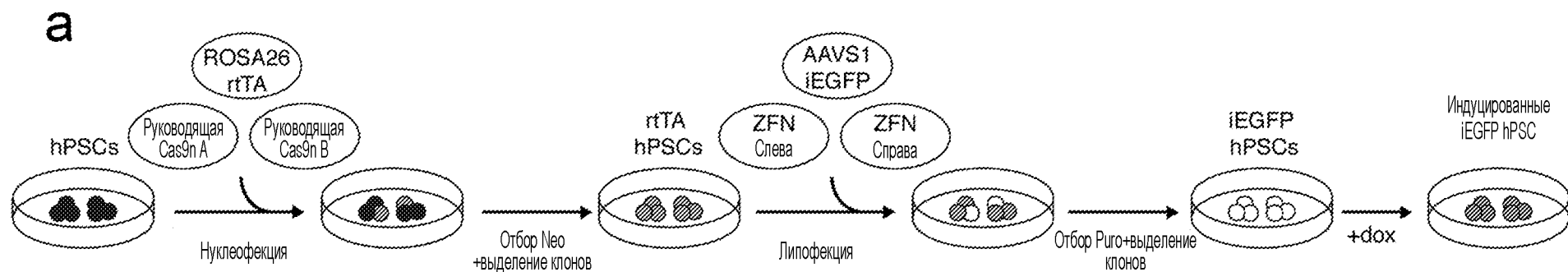
ФИГ. 1



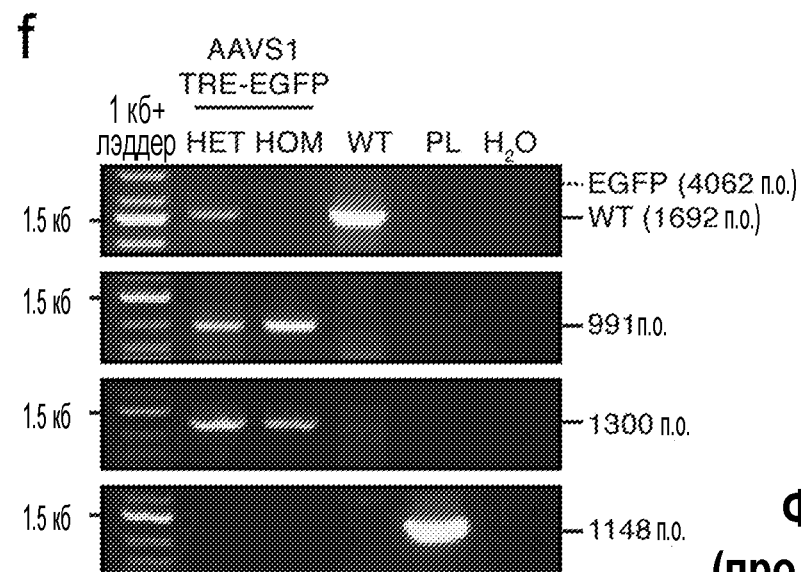
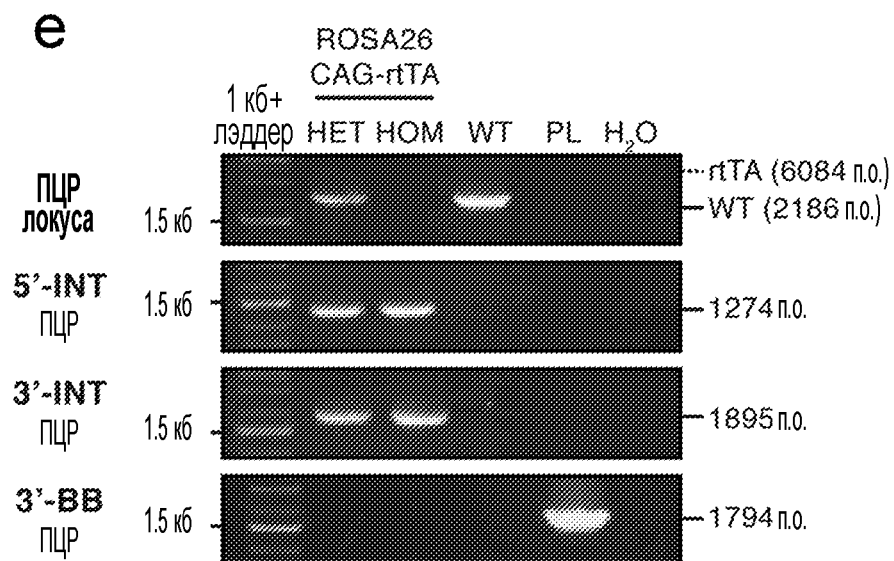
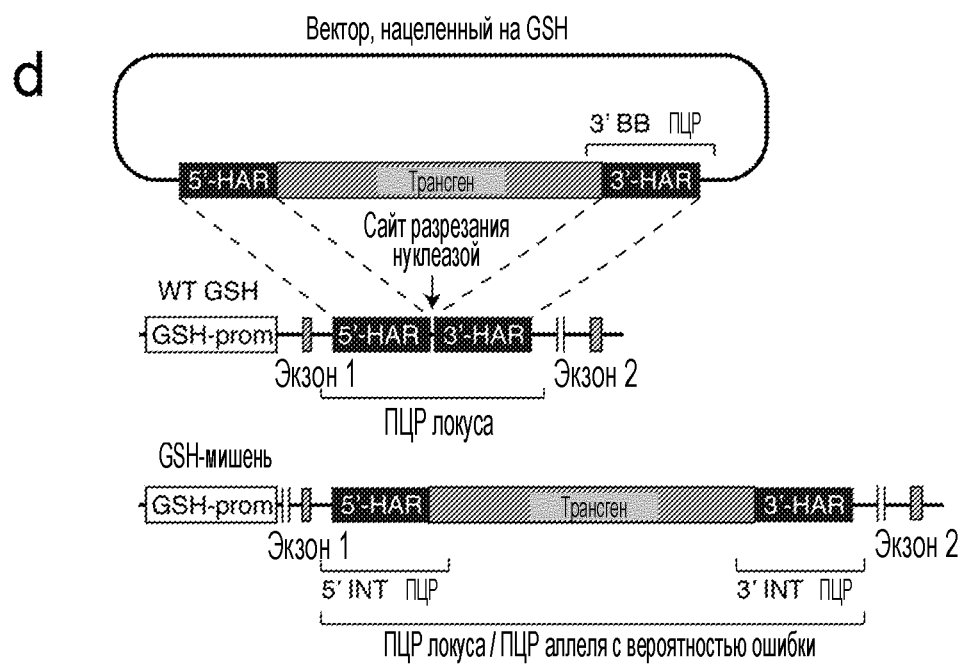
ФИГ. 2



ФИГ. 3

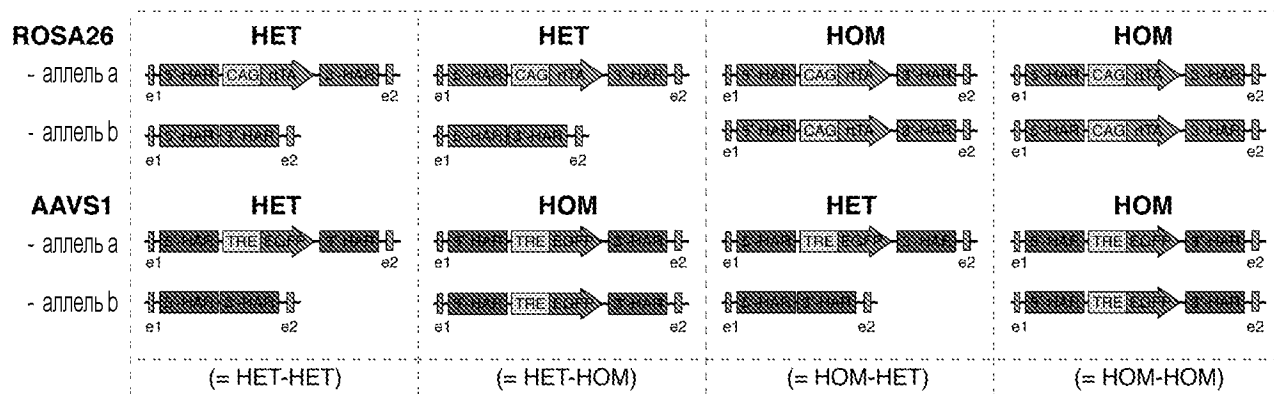


ФИГ. 4

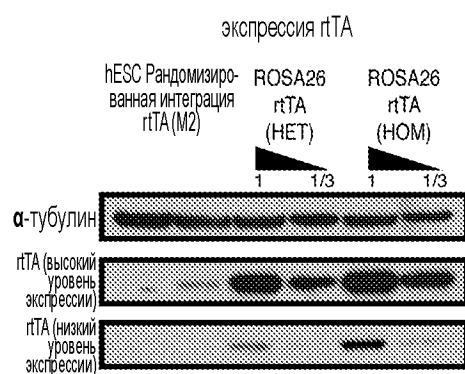


ФИГ. 4
(продолжение)

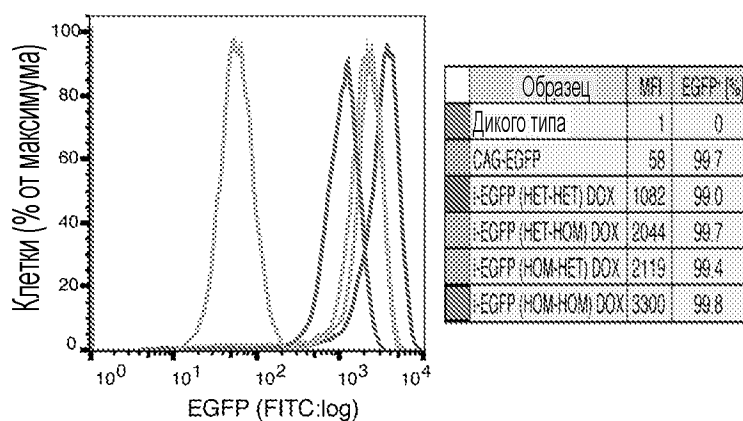
a



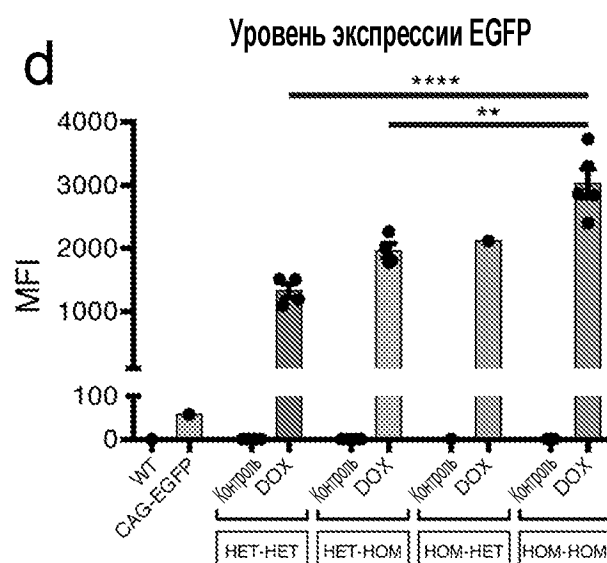
b



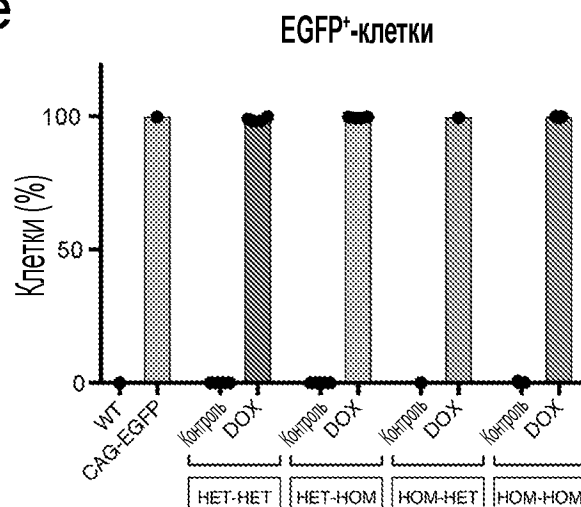
c



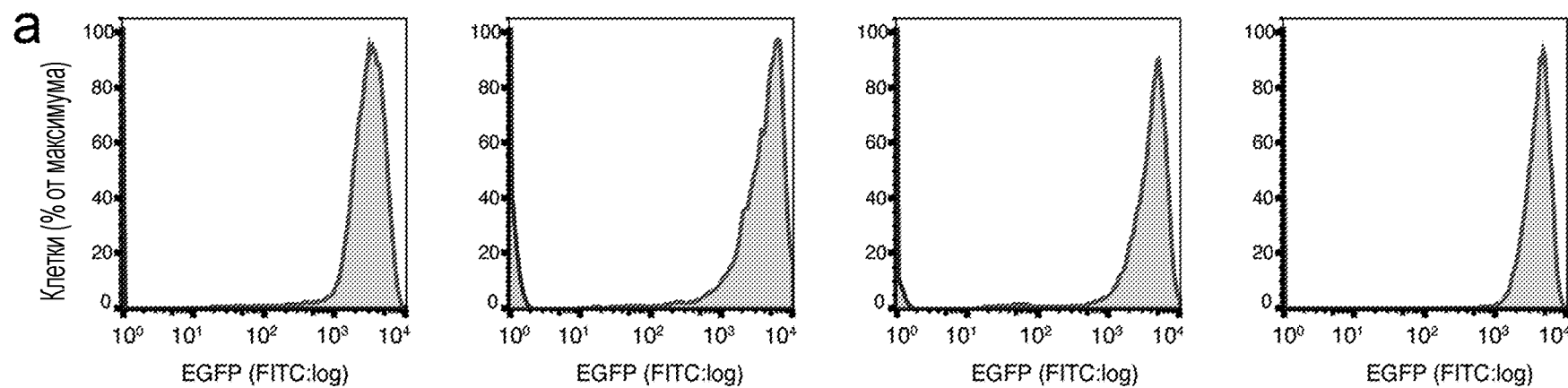
d



e



ФИГ. 5

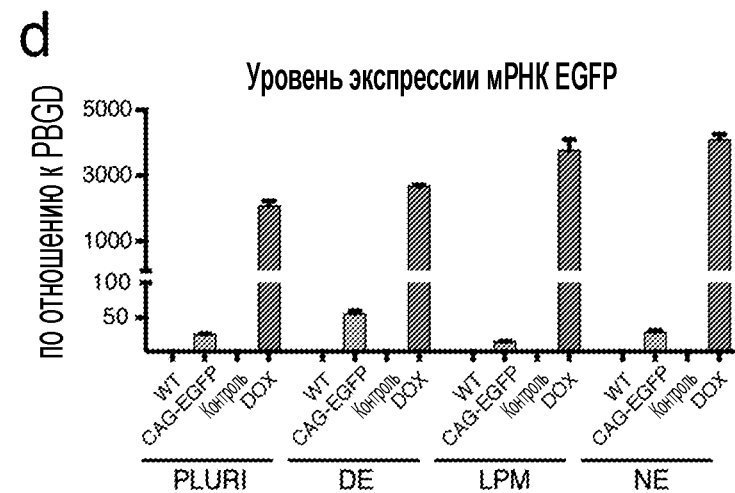
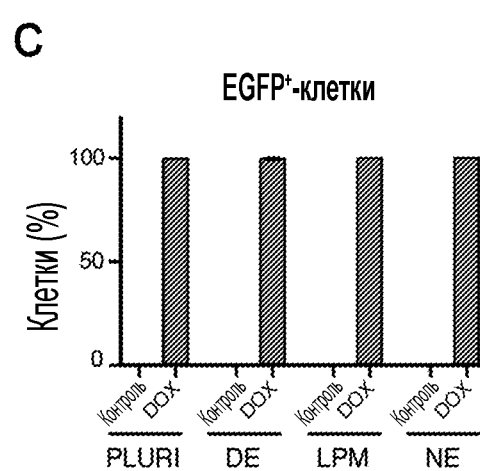
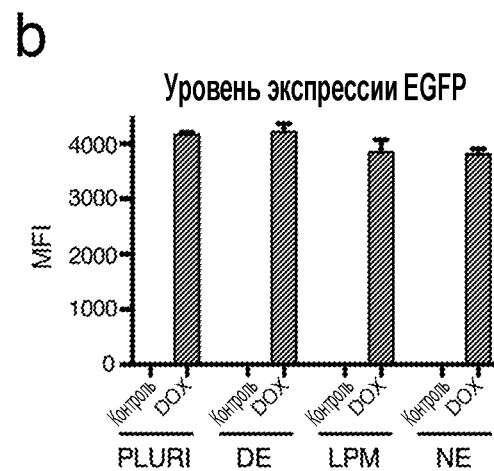


Образец	MFI	EGFP ⁺ (%)
Плюрипотентность, CTR	1	0
Плюрипотентность, Dox	3030	99.8

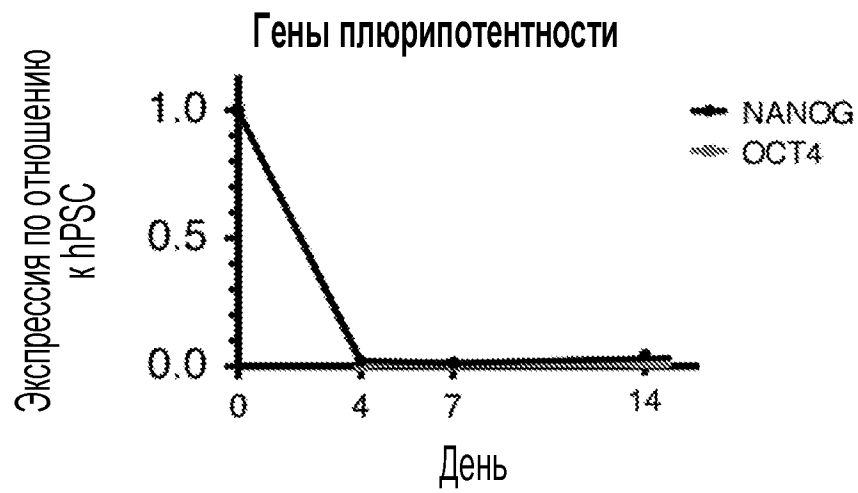
Образец	MFI	EGFP ⁺ (%)
Эндодерма, CTR	1	0
Эндодерма, Dox	4207	99.8

Образец	MFI	EGFP ⁺ (%)
Мезодерма, CTR	1	0
Мезодерма, Dox	4077	99.9

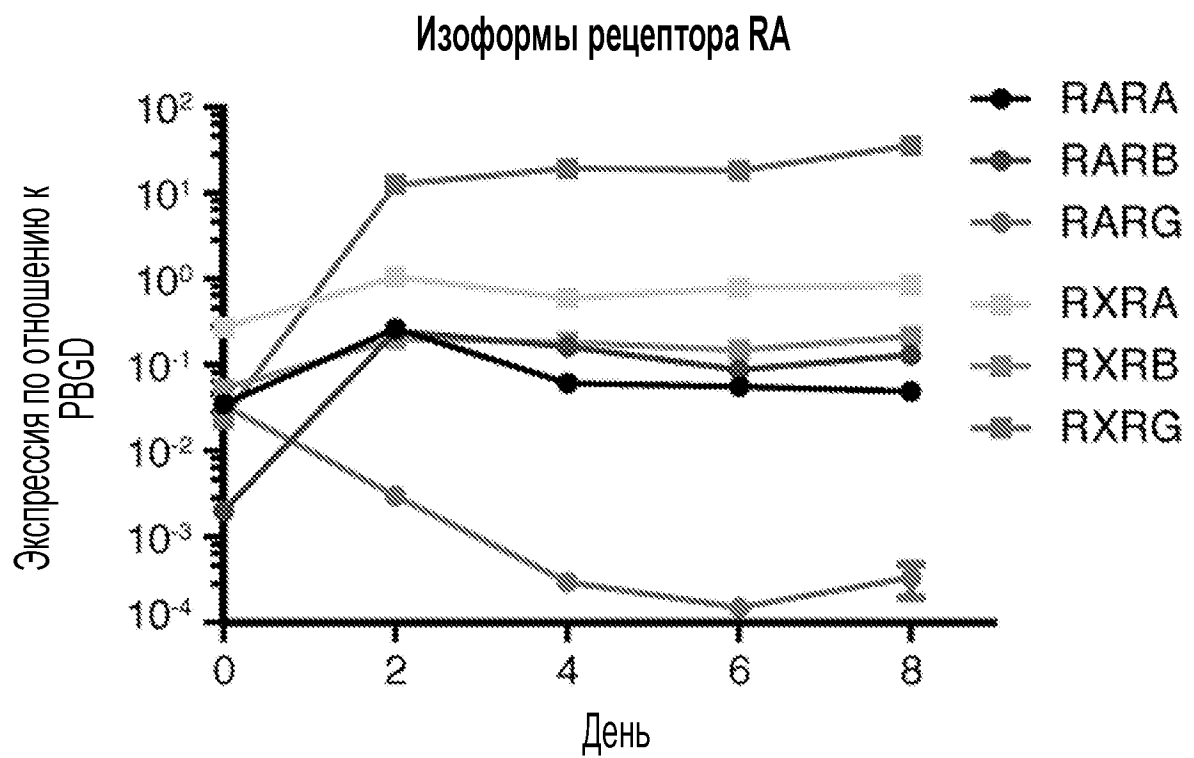
Образец	MFI	EGFP ⁺ (%)
Эктодерма, CTR	1	0
Эктодерма, Dox	4068	100



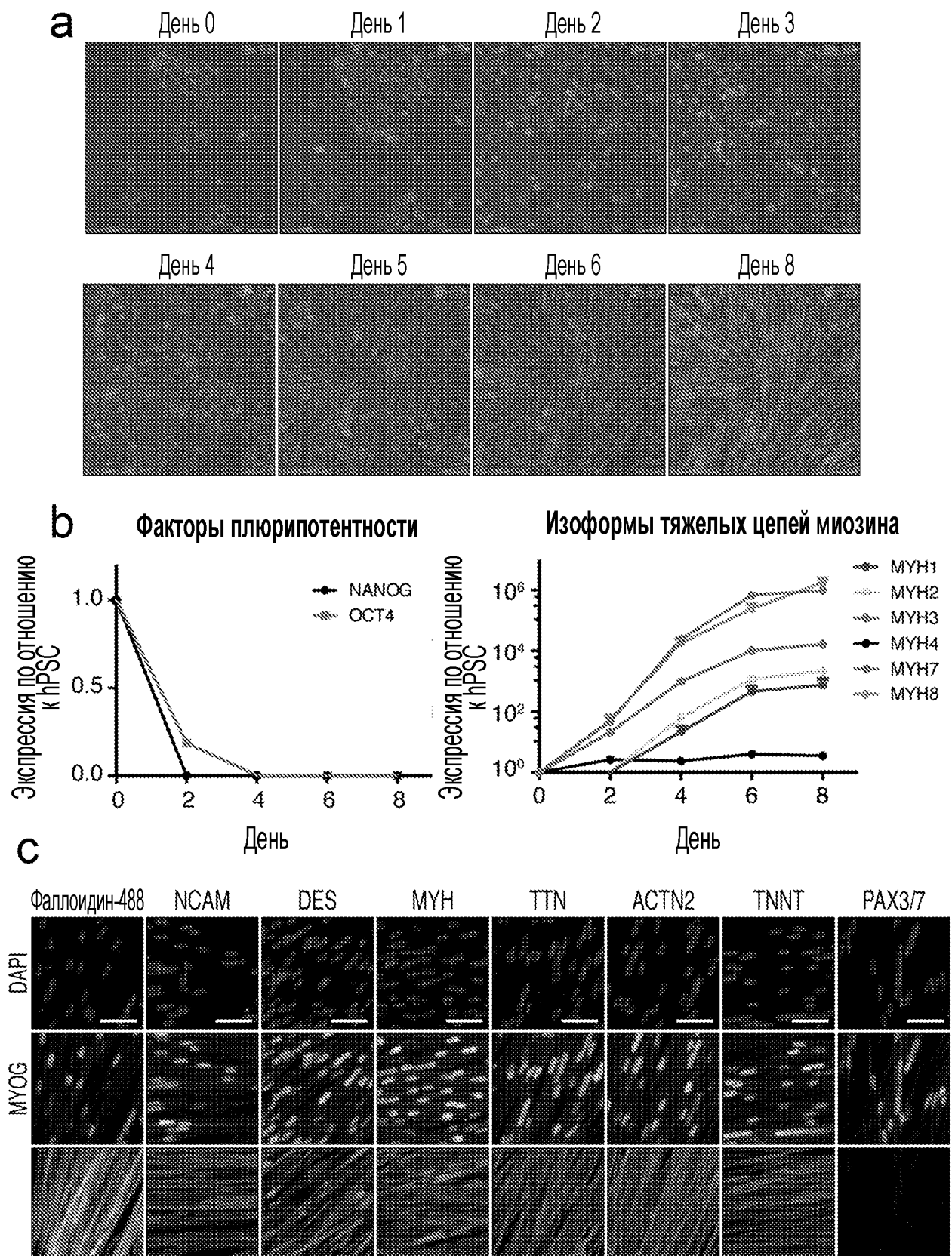
ФИГ. 6



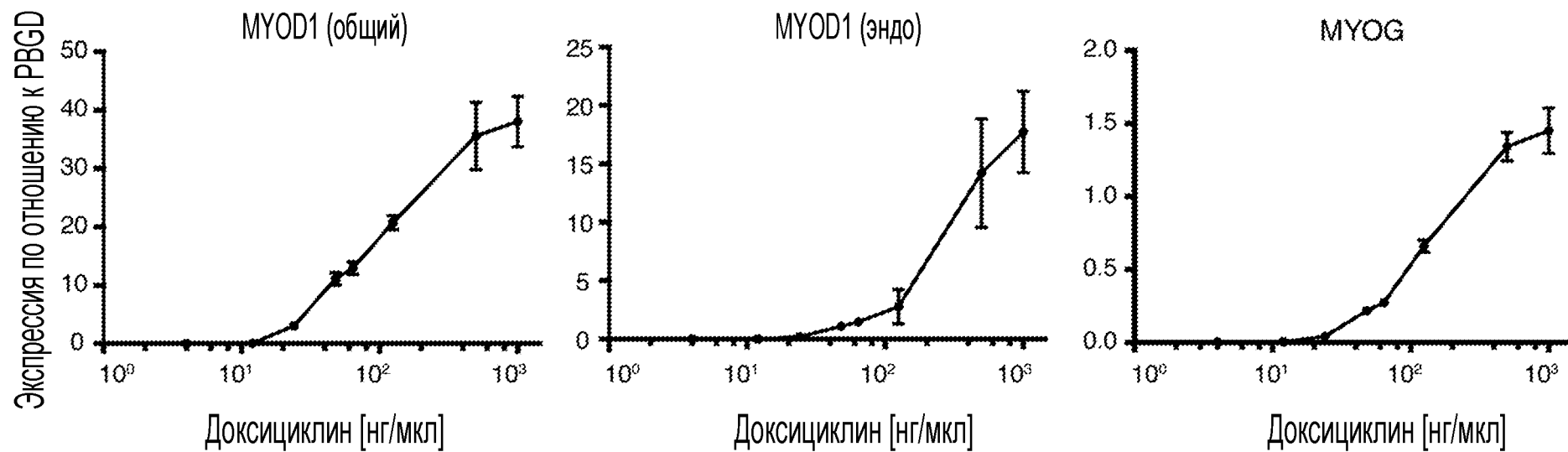
ФИГ. 7



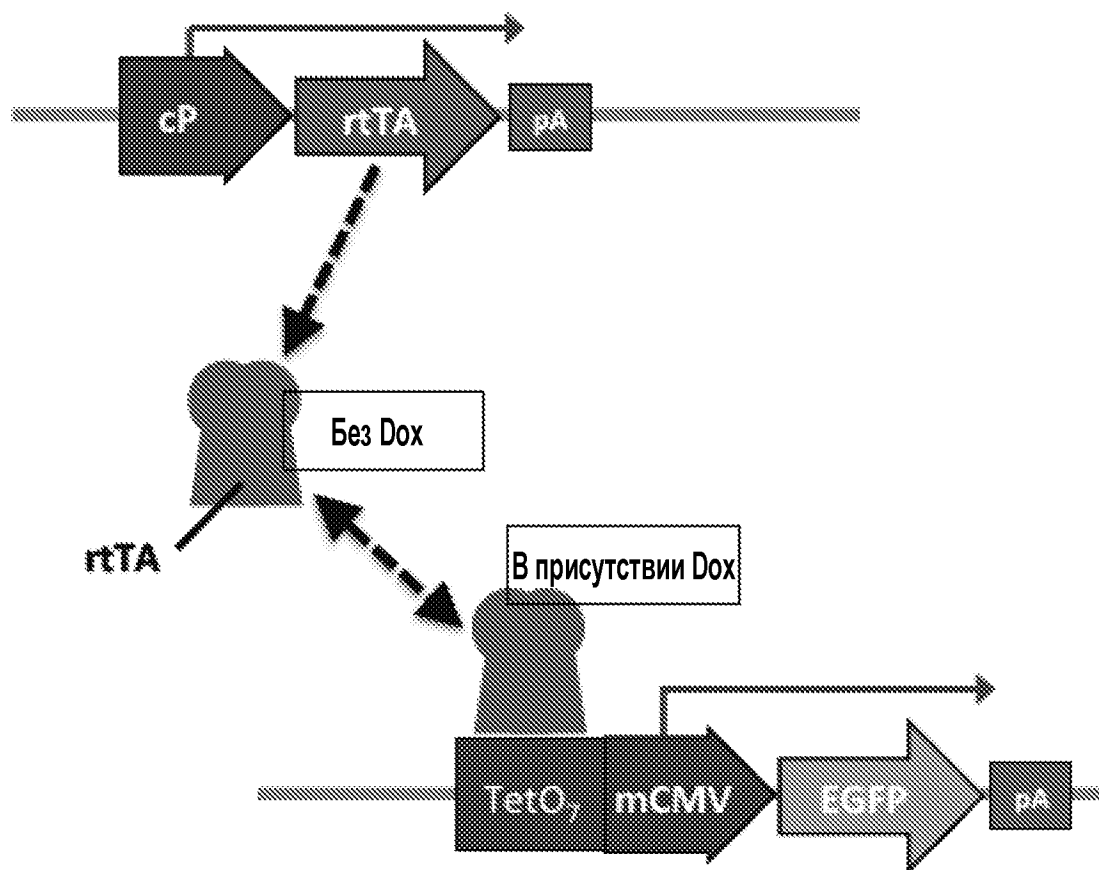
ФИГ. 8



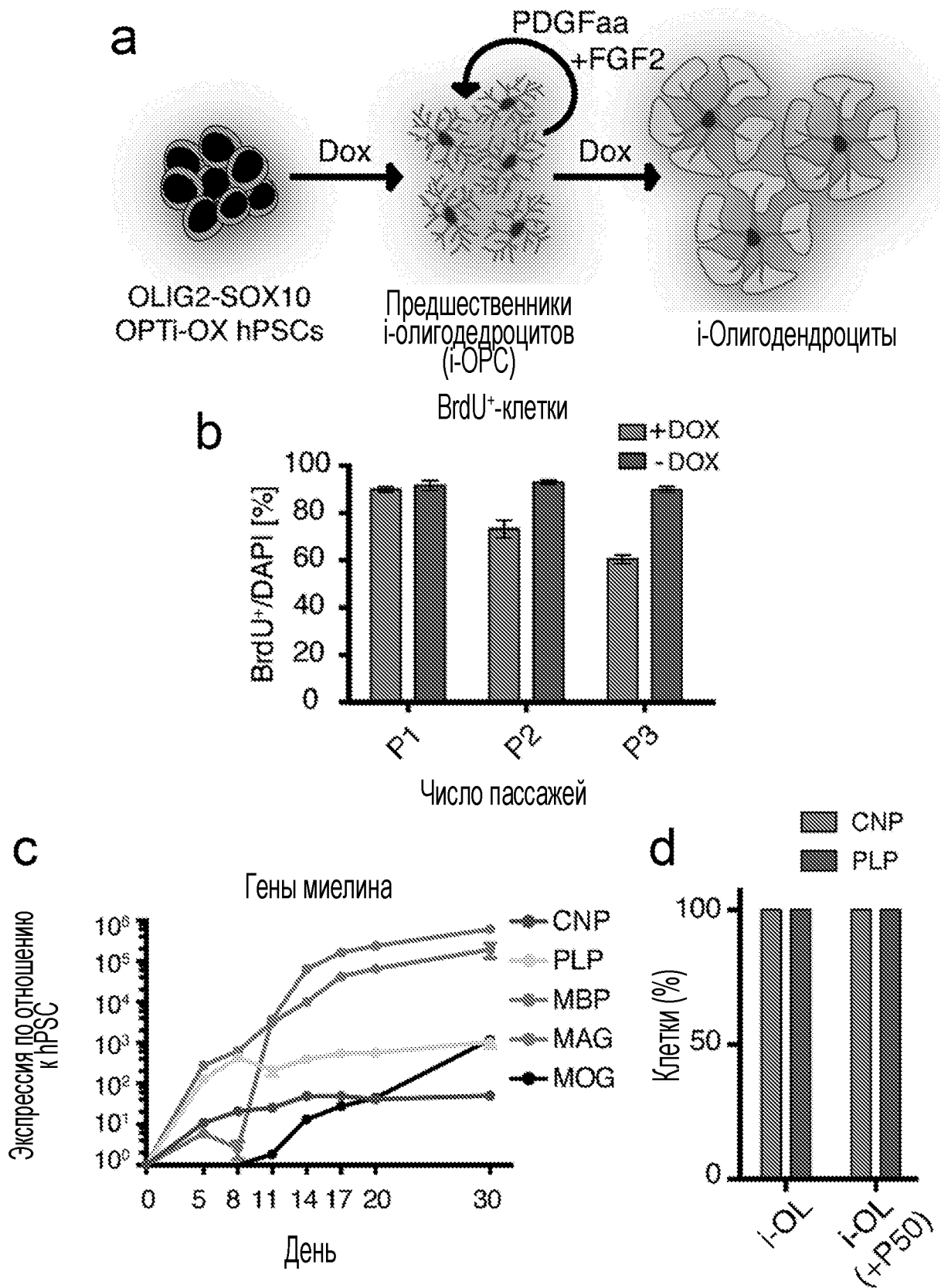
ФИГ.9



ФИГ. 10



ФИГ. 11



ФИГ. 12

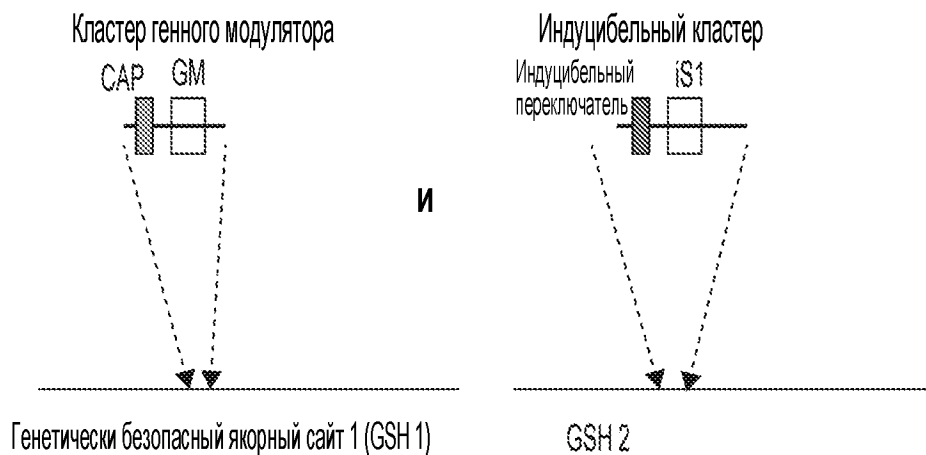
CAP=Конститутивно активный промотор

GM=Генный модулятор (например, rtTA)

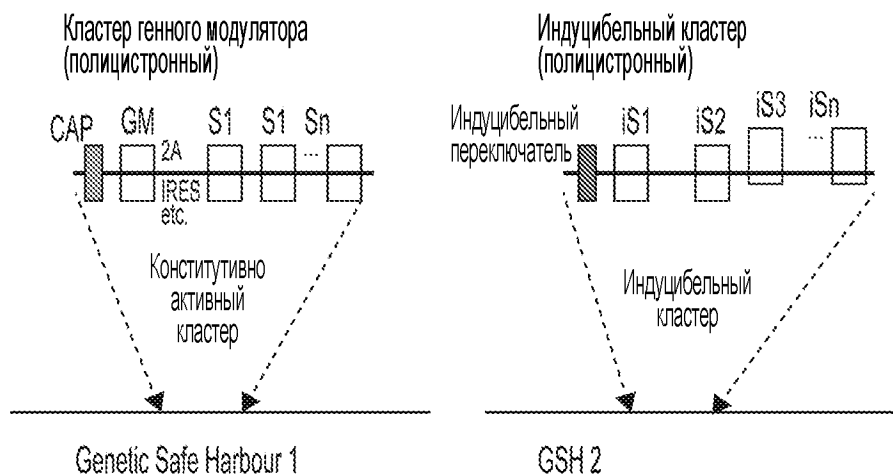
S1-n=Конститутивно экспрессируемая генная последовательность

iS1-n=Индукцибельно экспрессируемая генная последовательность

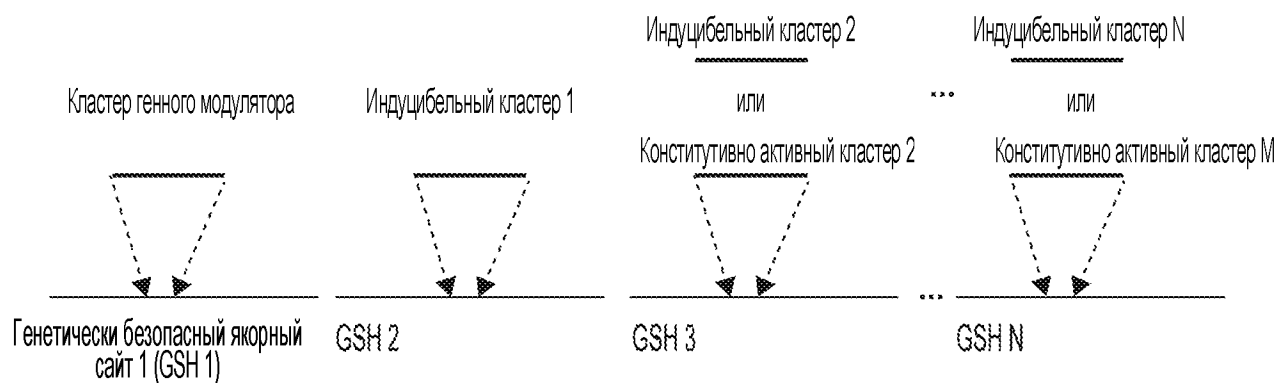
Принцип настоящего изобретения

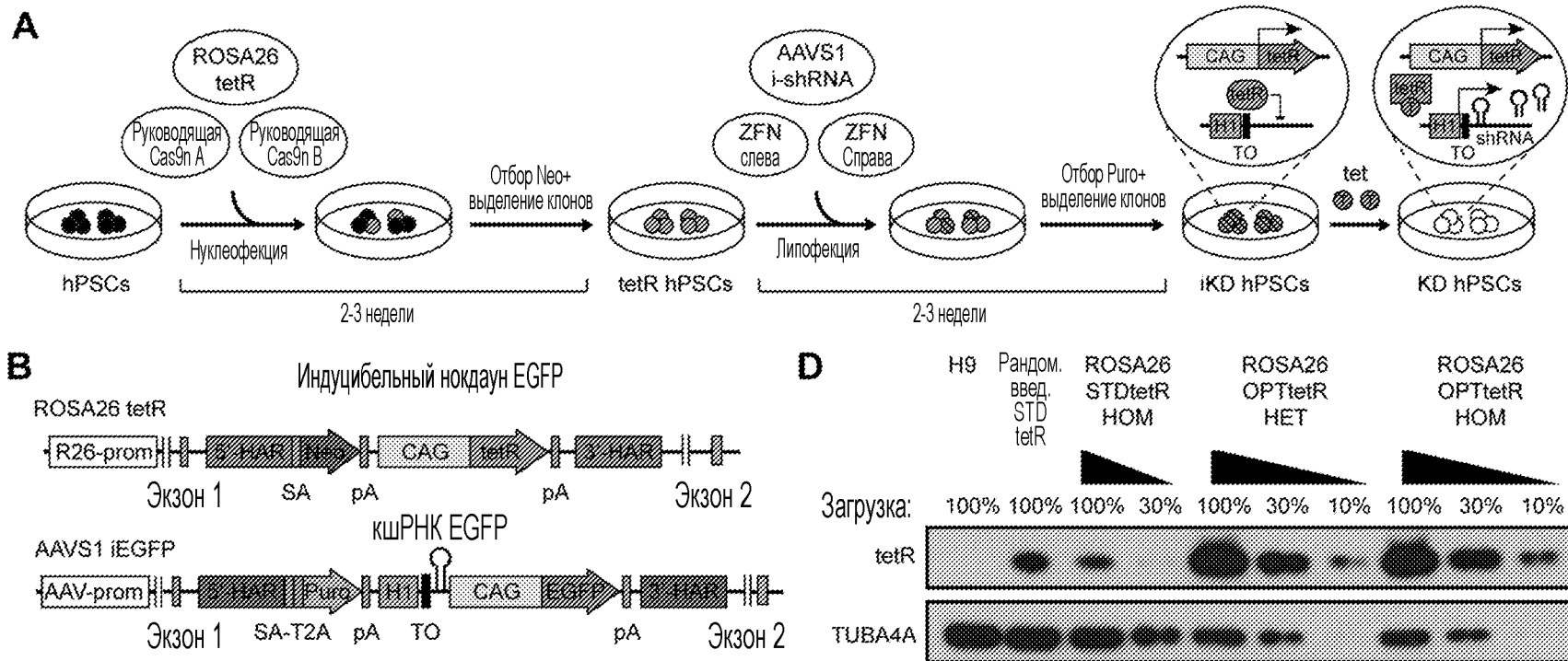


Введение полицистронных векторных конструкций

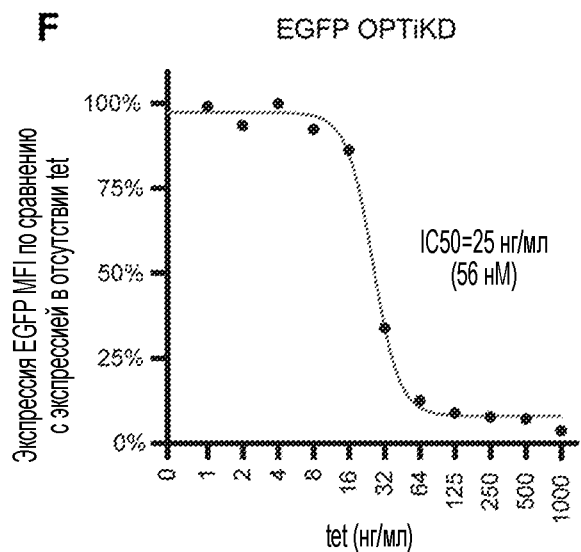
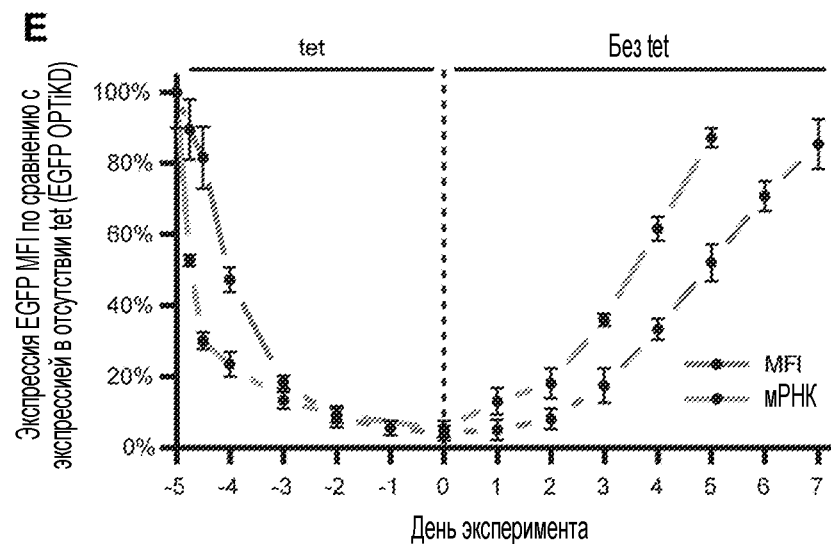
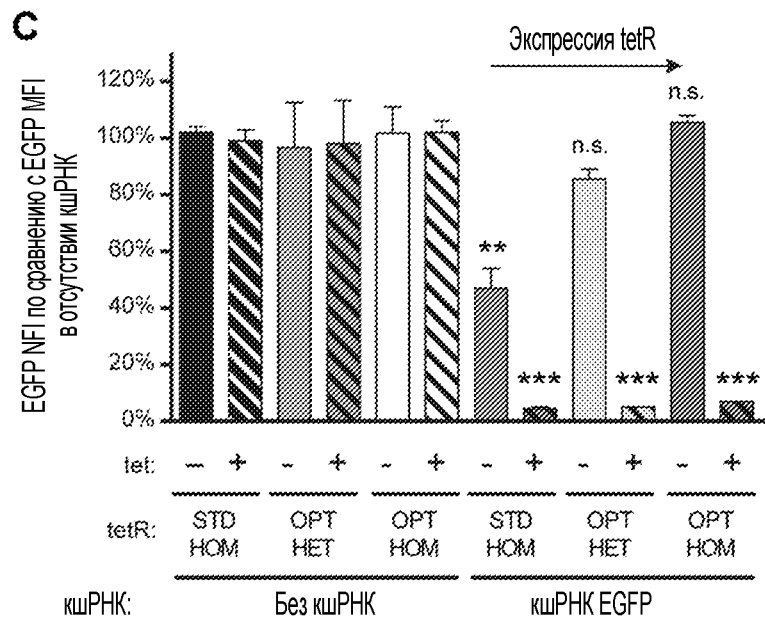


Введение множества GSH

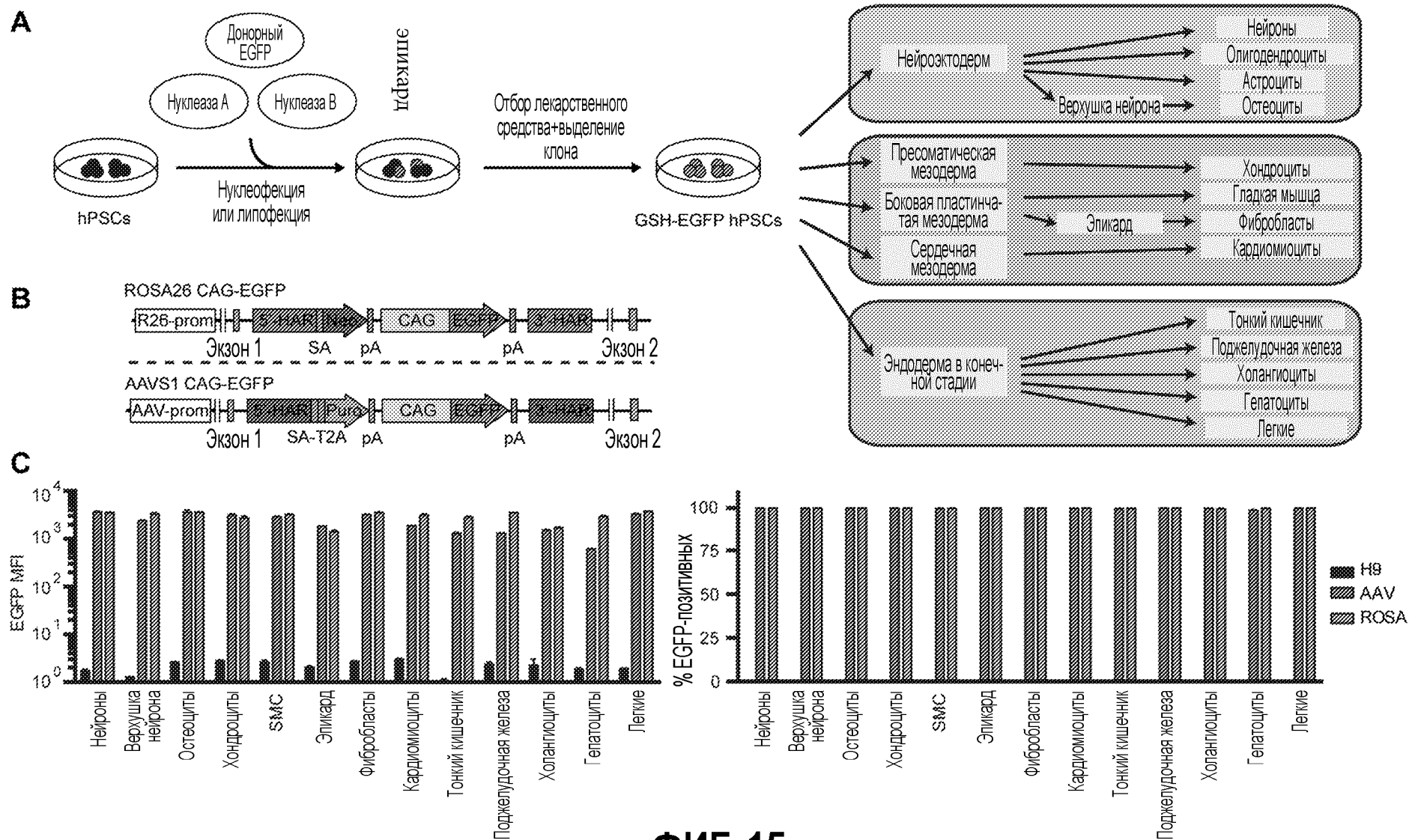




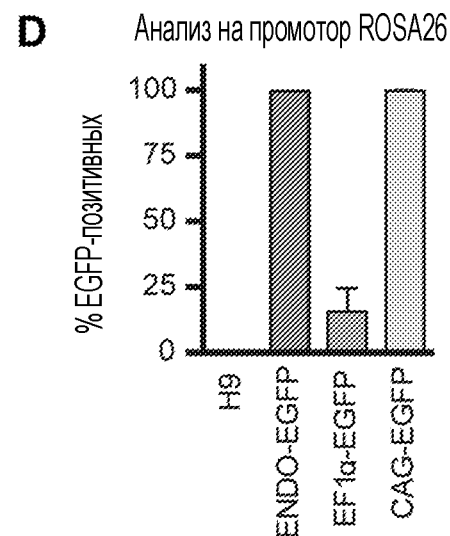
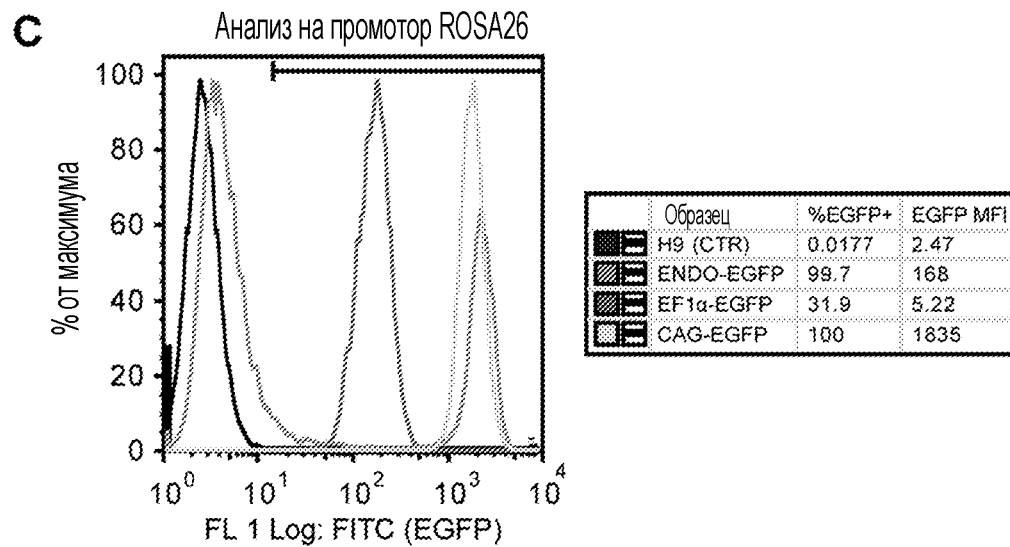
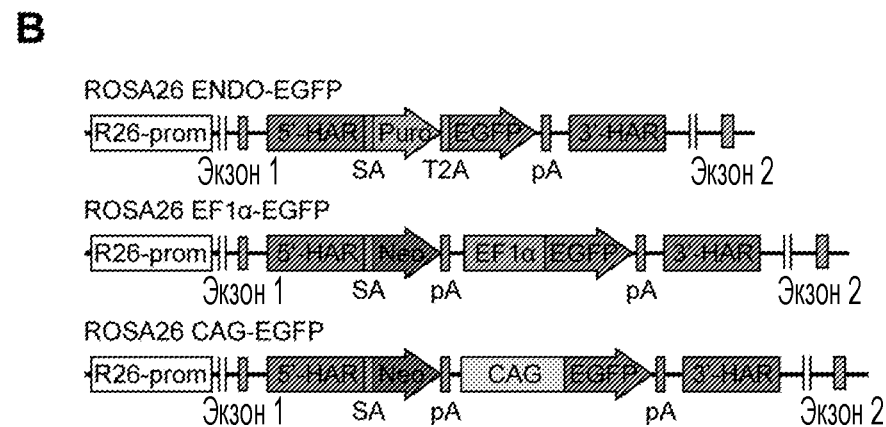
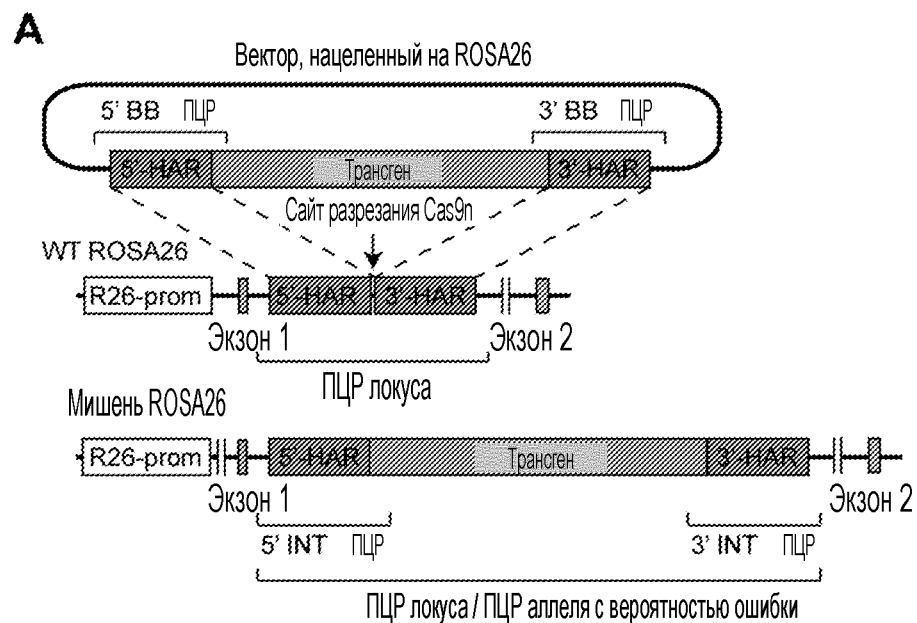
ФИГ. 14



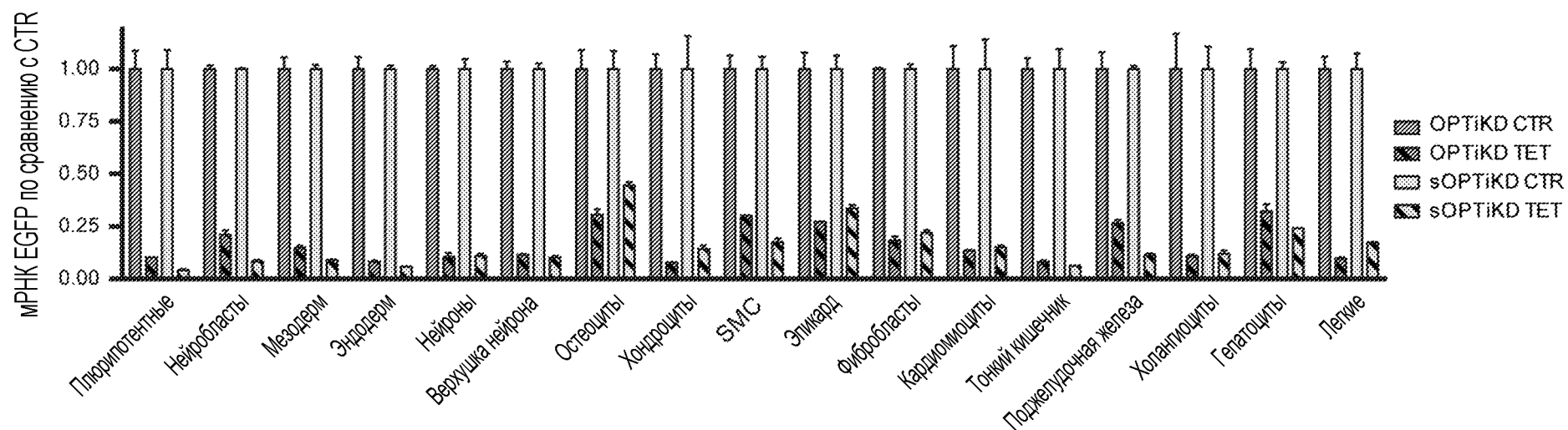
ФИГ. 14
(продолжение)



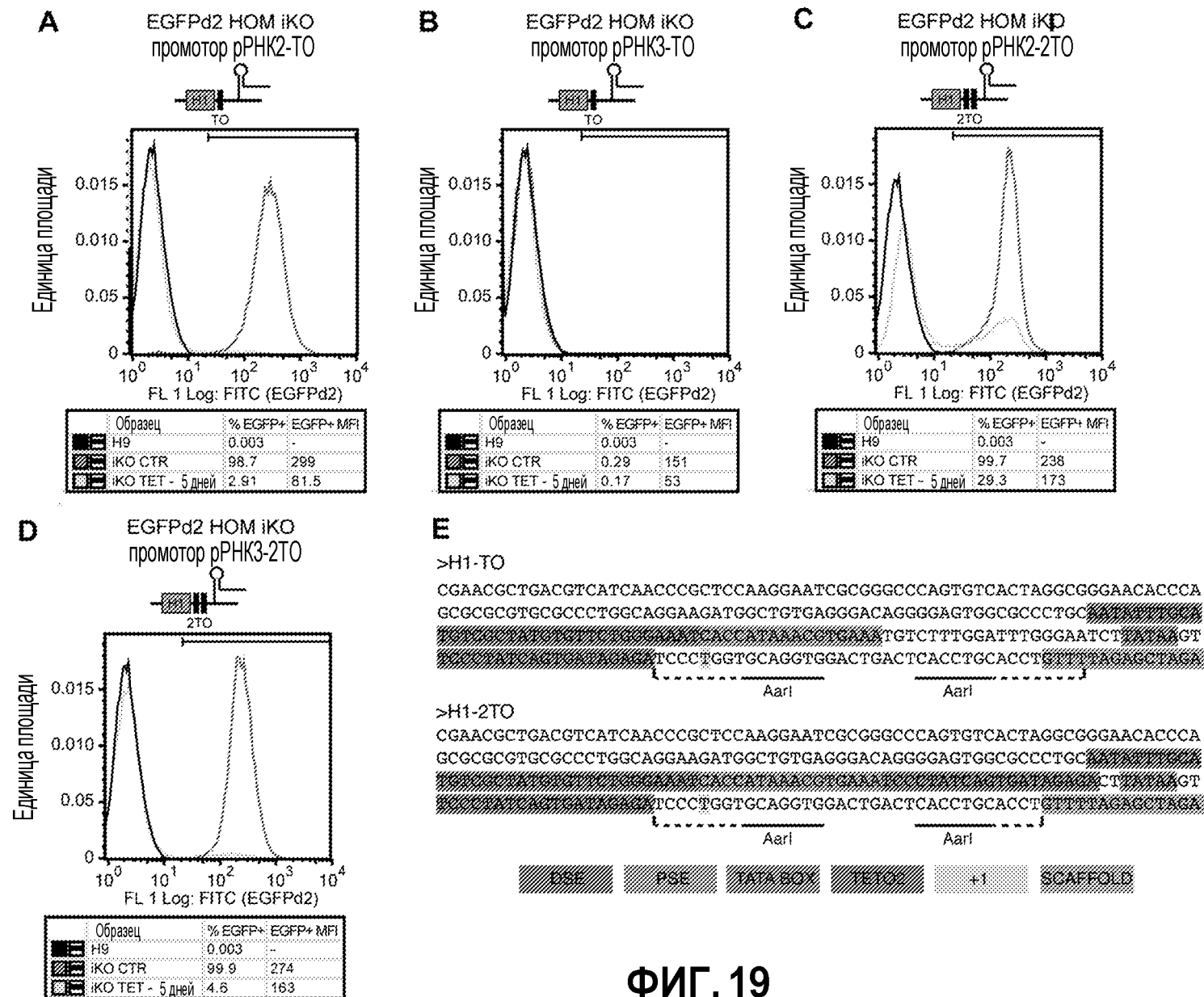
ФИГ. 15



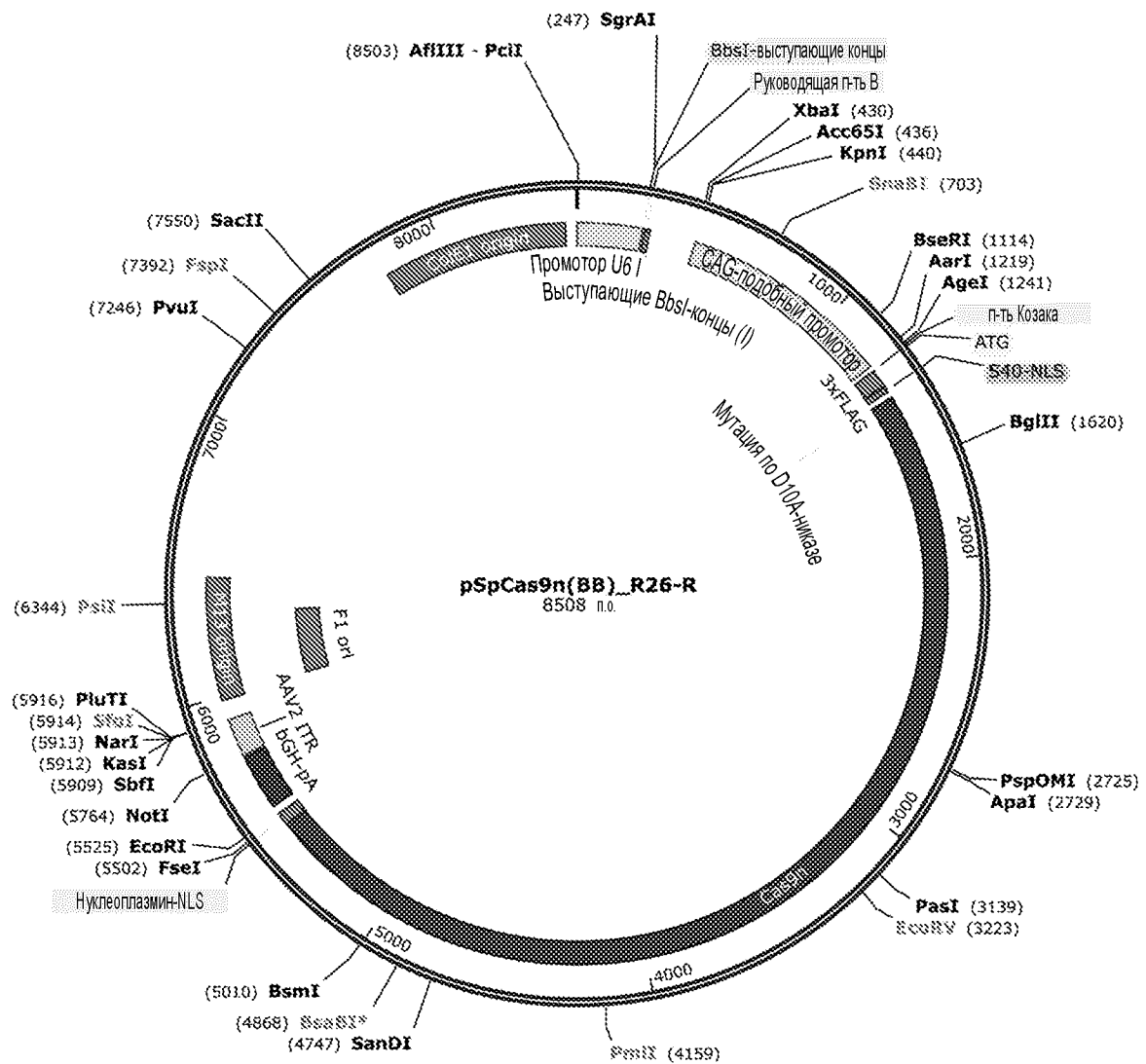
ФИГ. 16



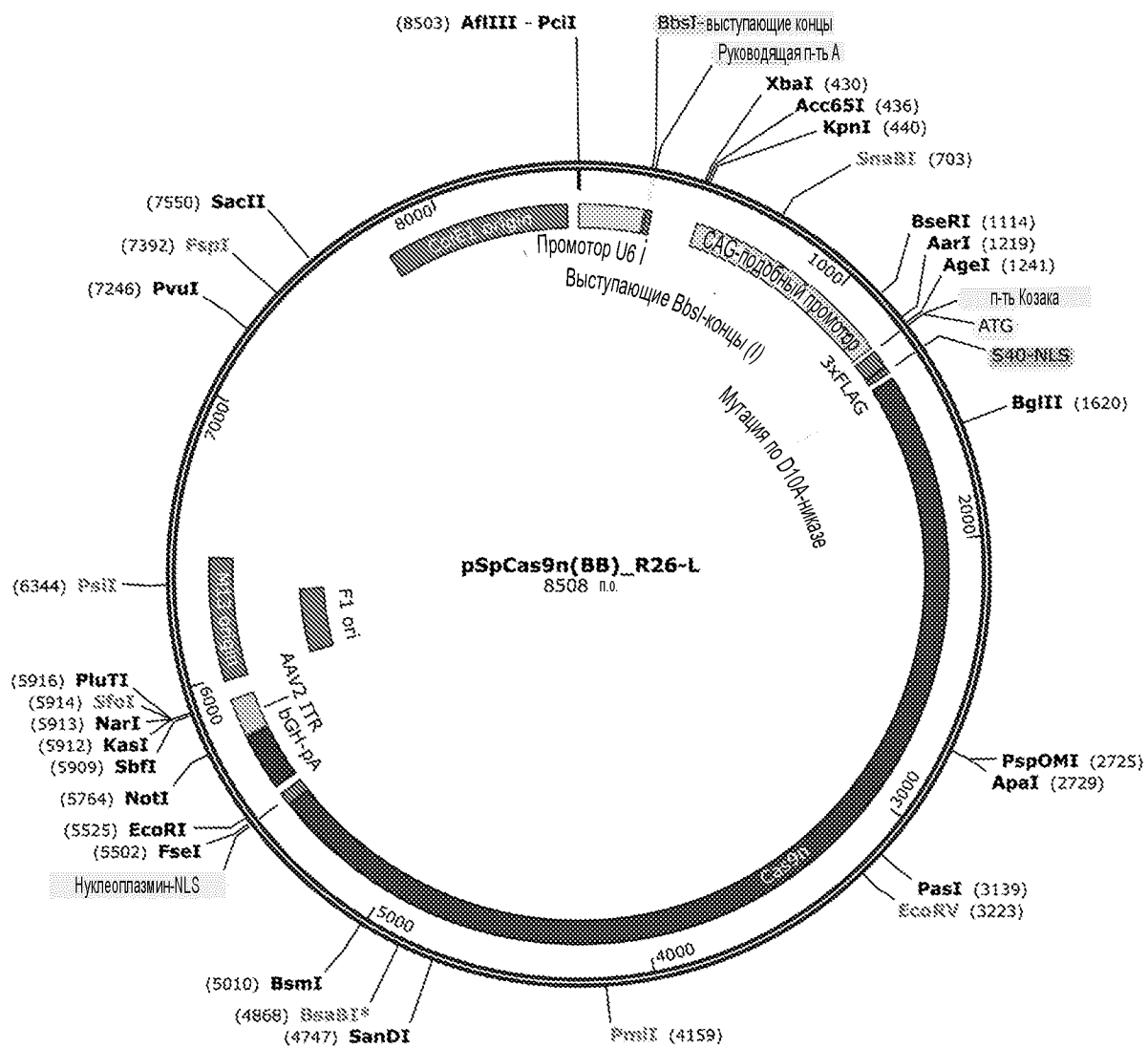
ФИГ. 17



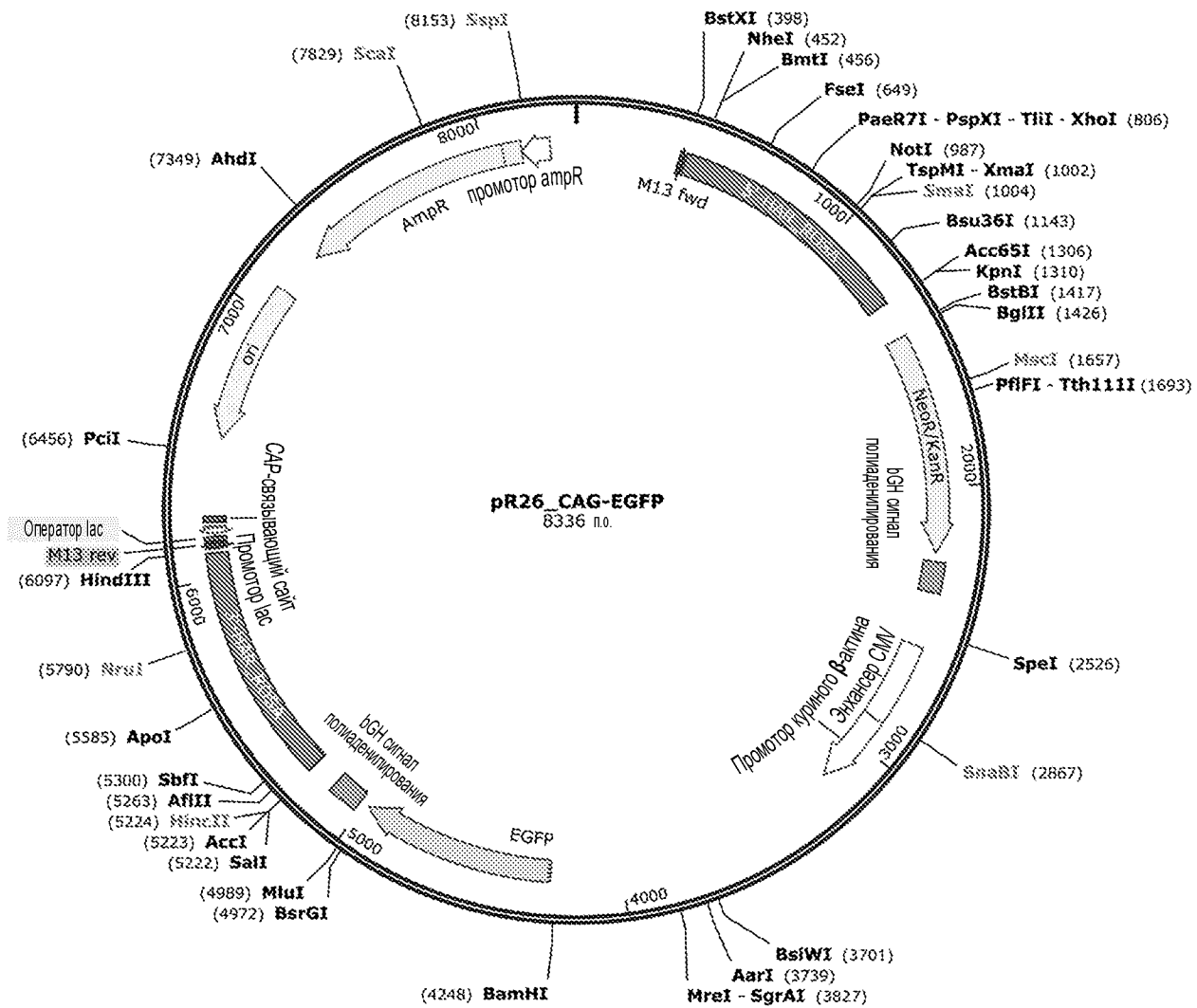
ФИГ. 19



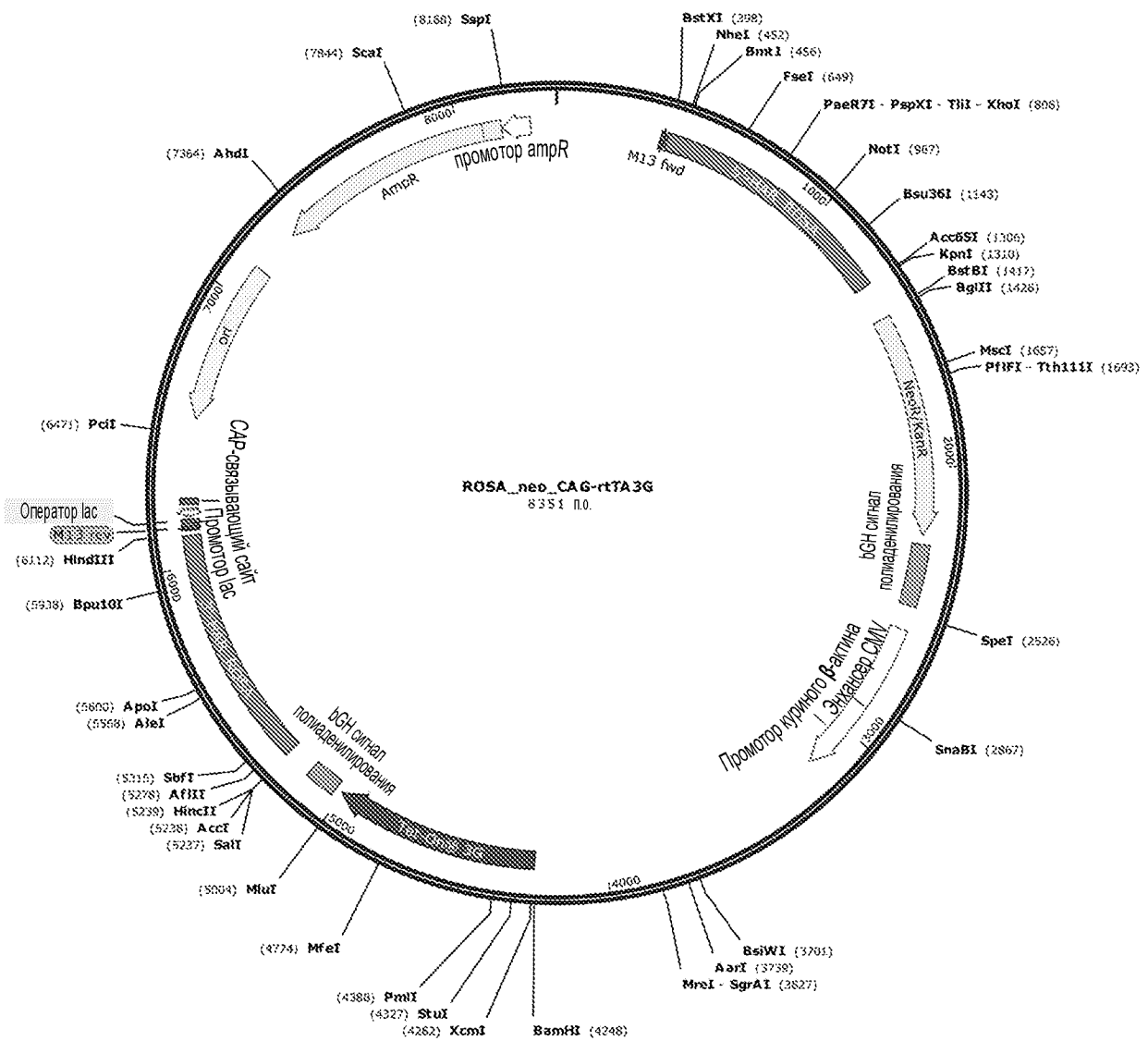
ФИГ. 20



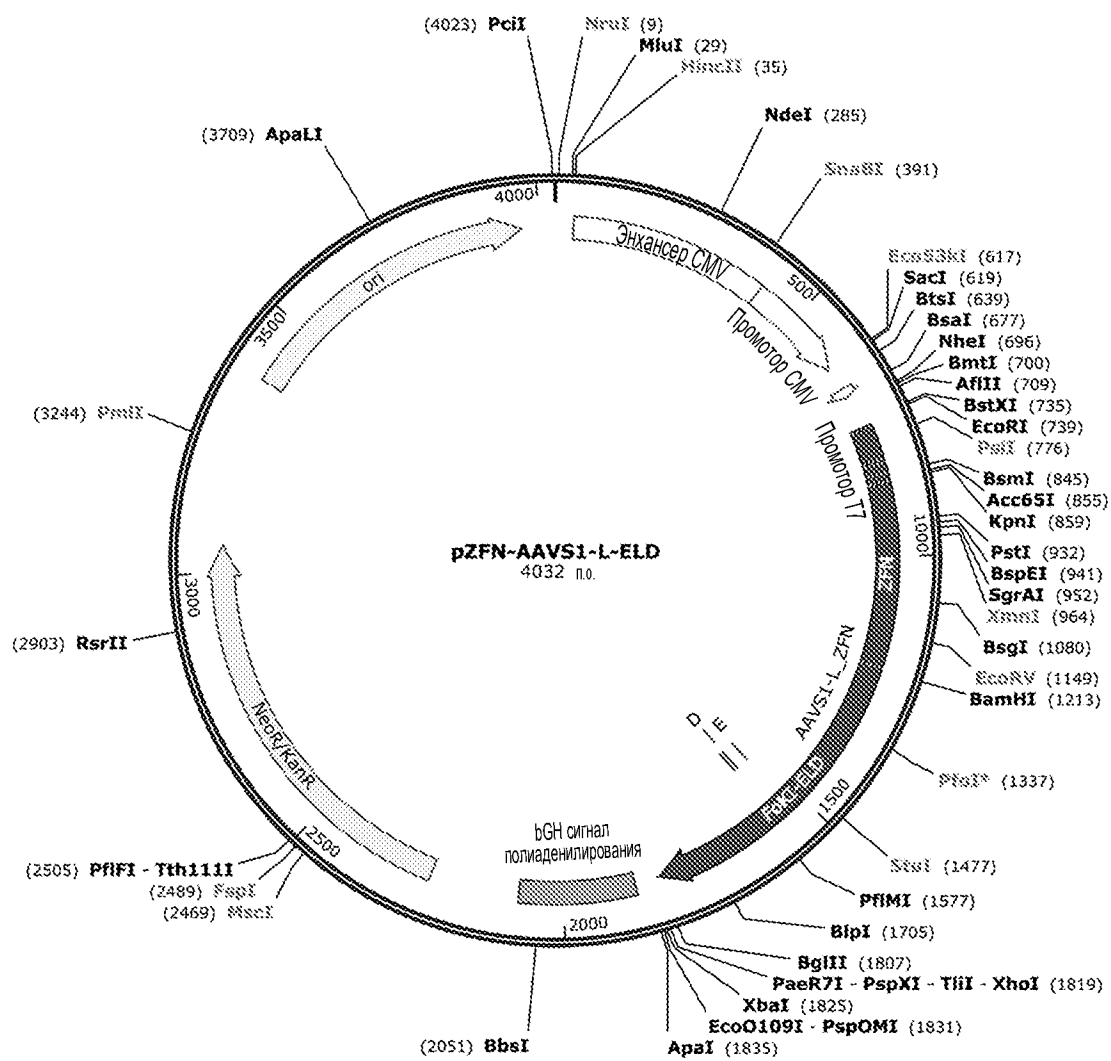
ФИГ. 21



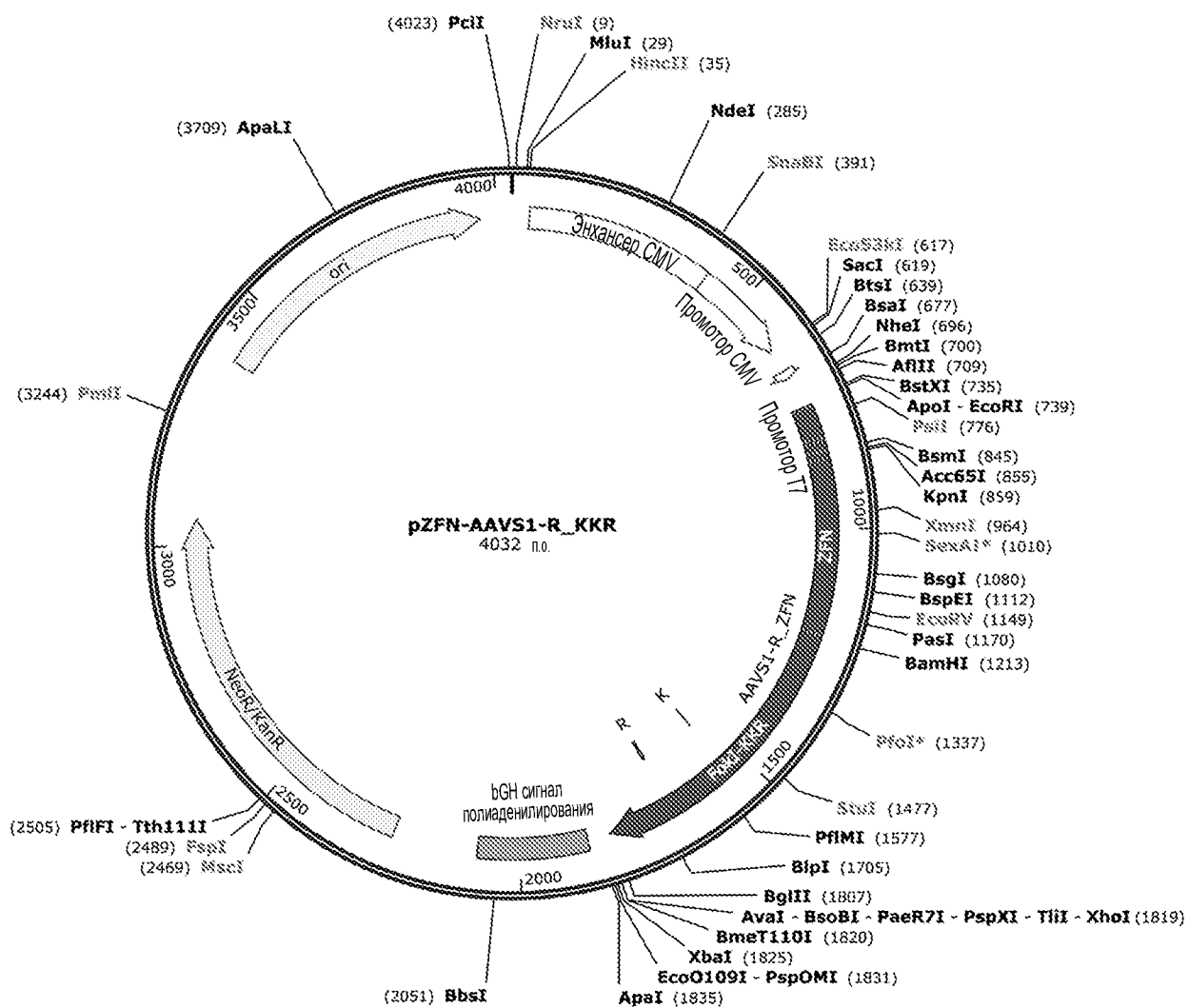
ФИГ. 22



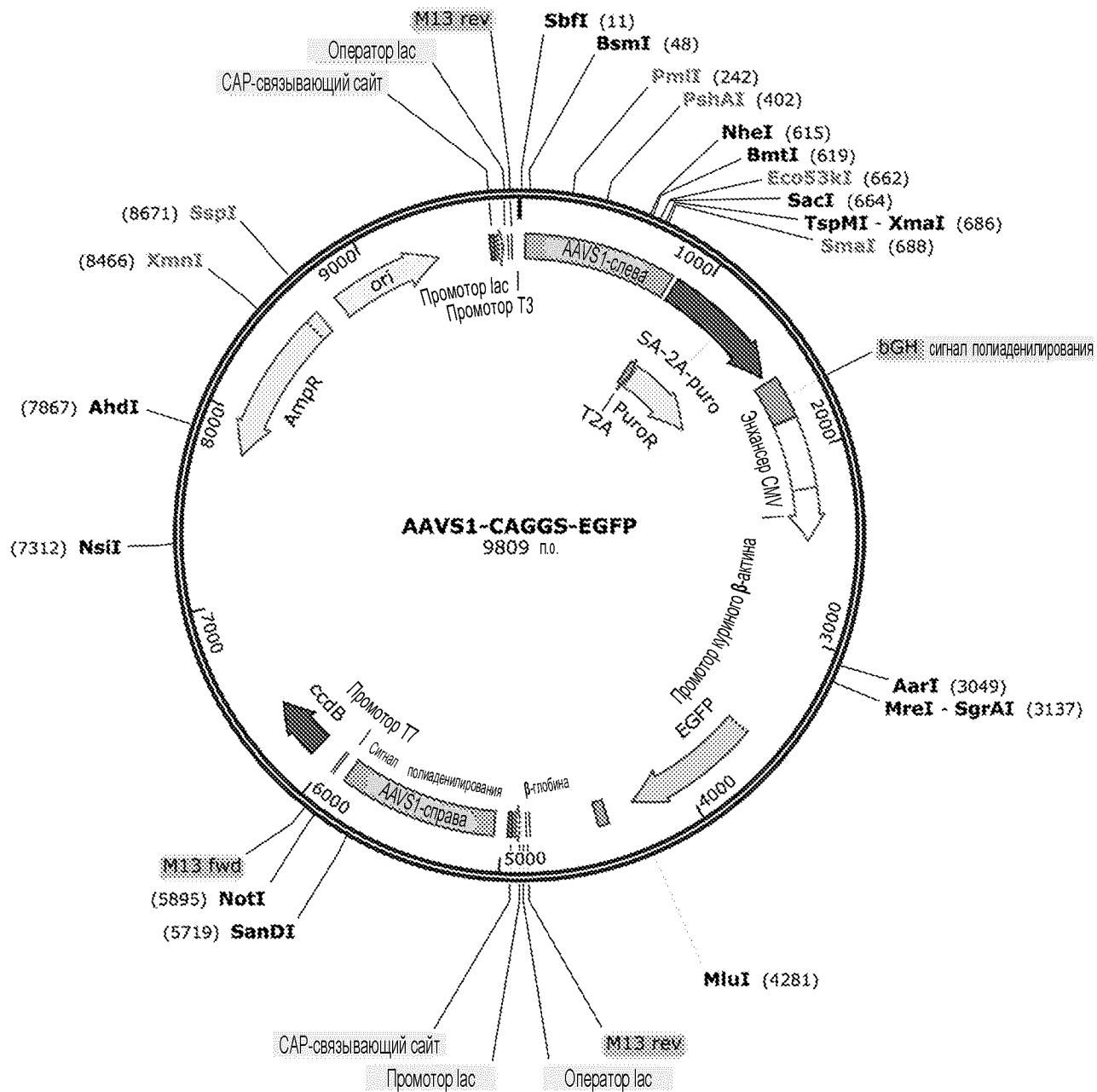
ФИГ. 23



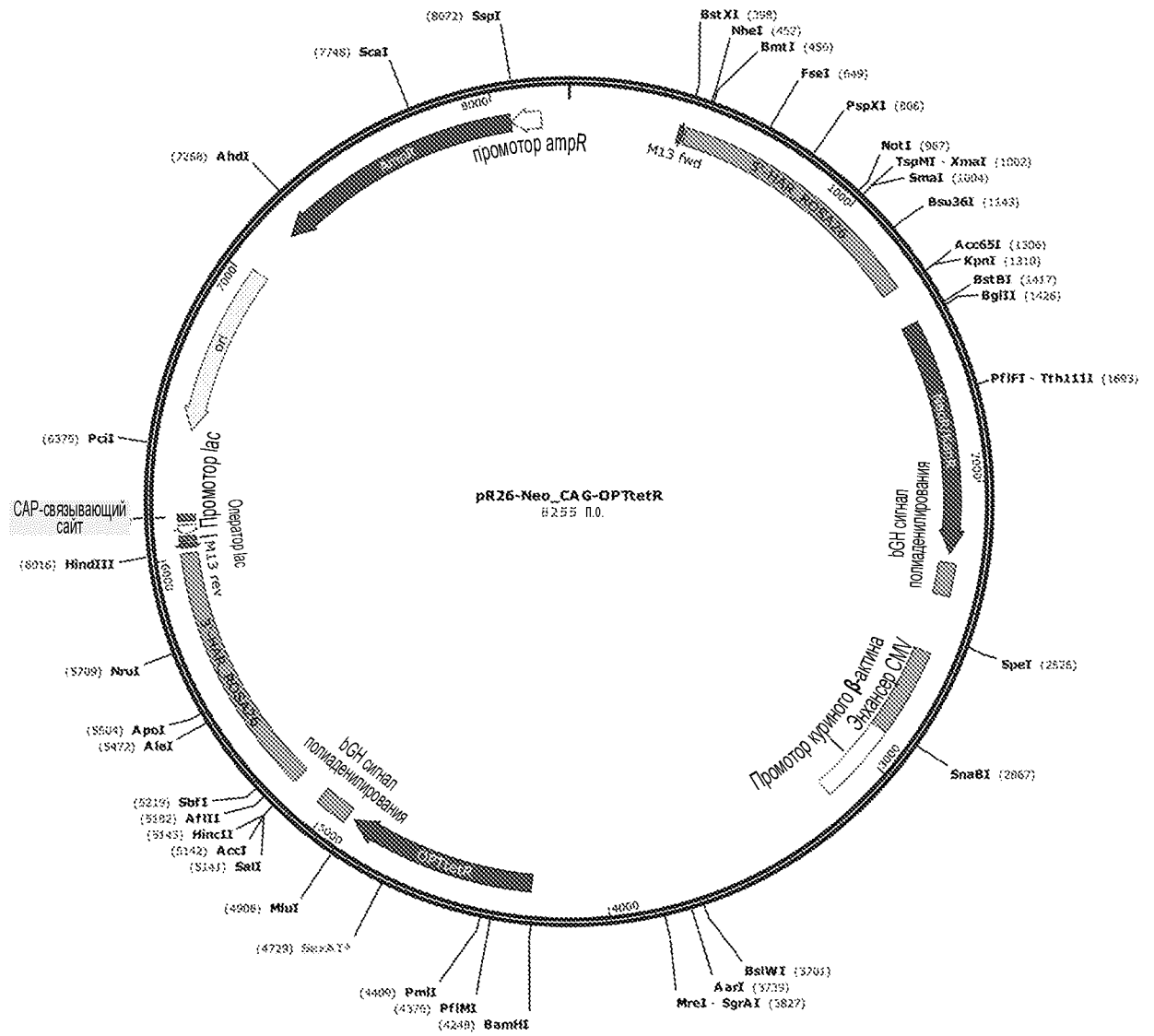
ФИГ. 24



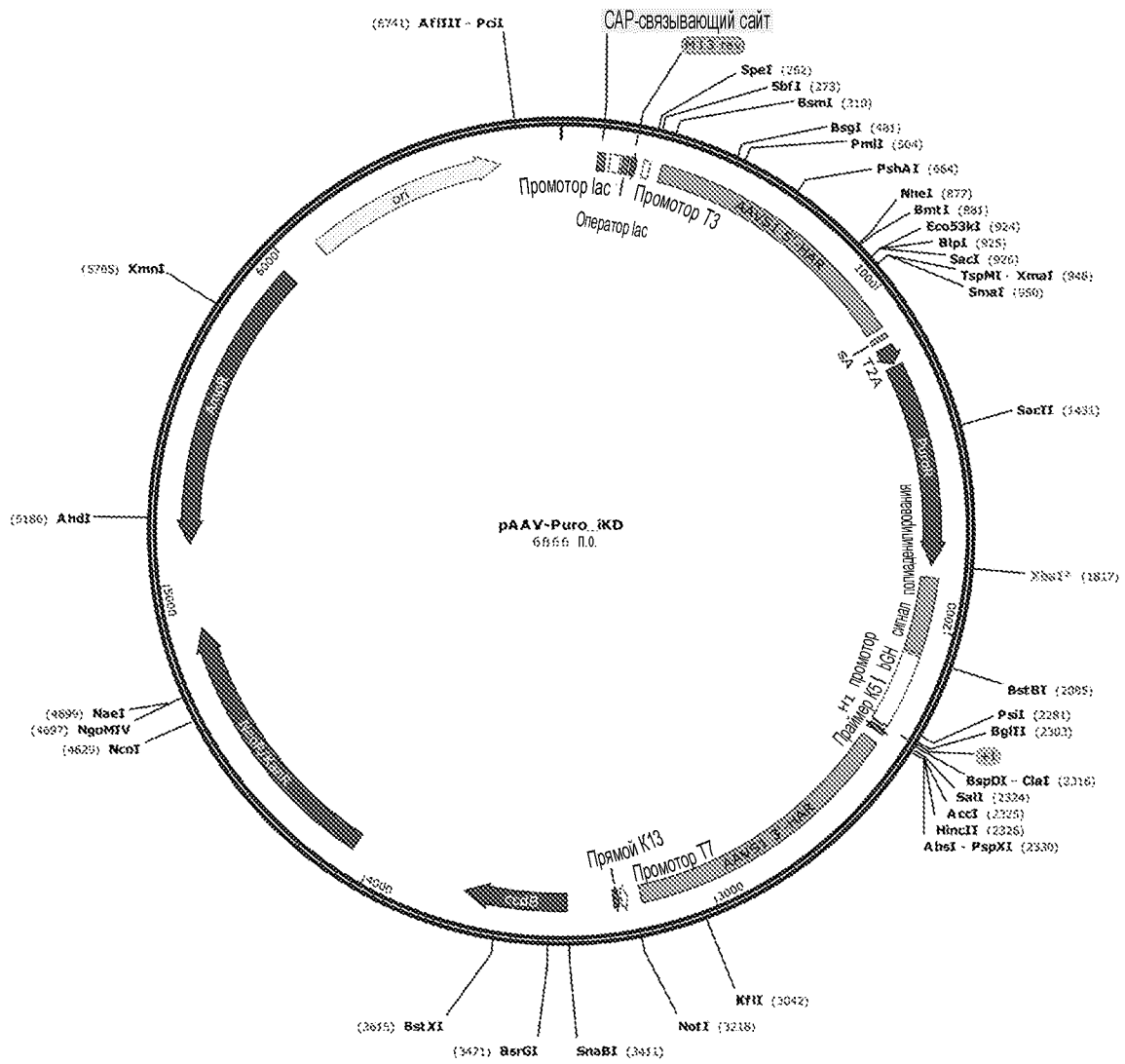
ФИГ. 25



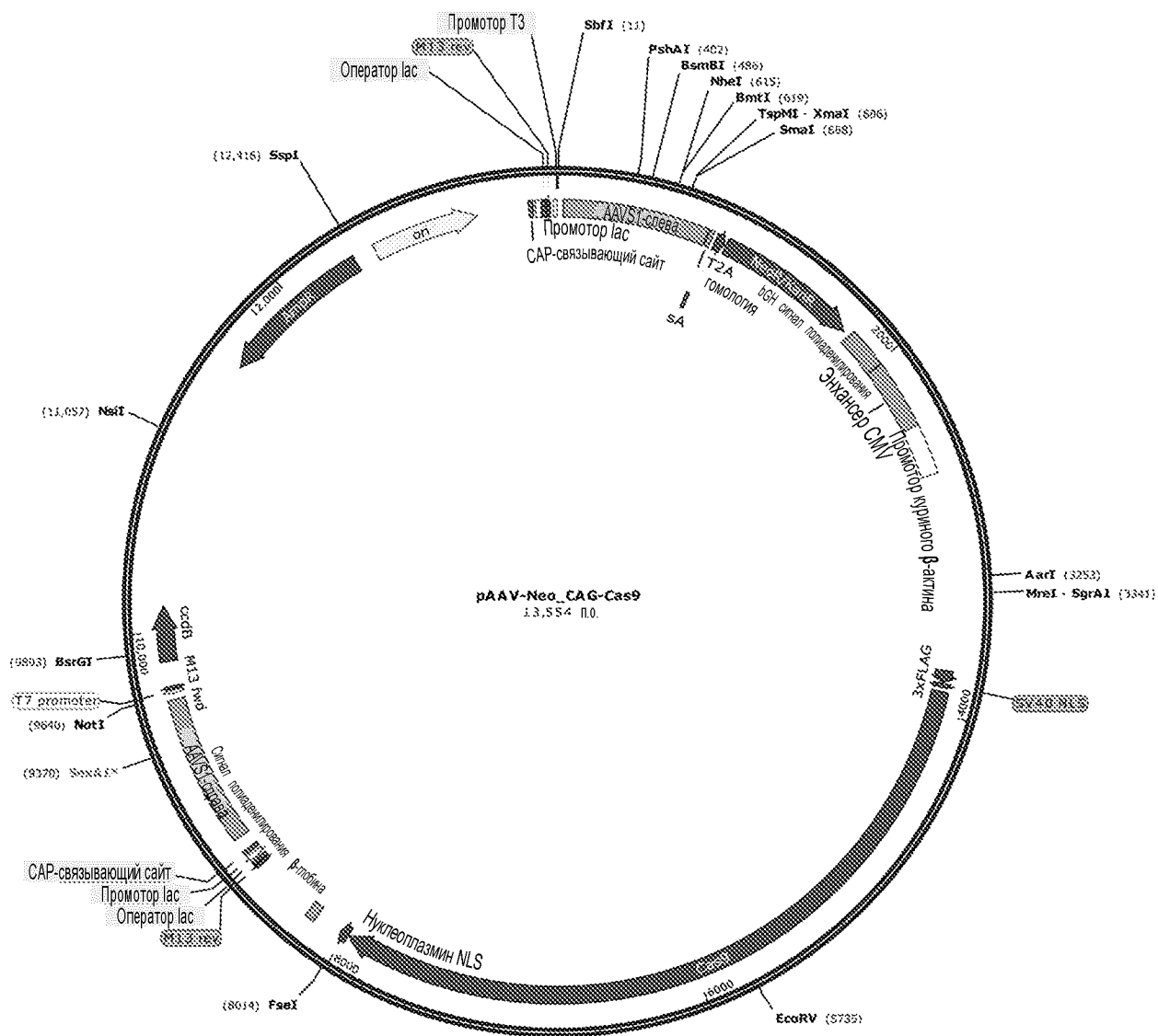
ФИГ. 26



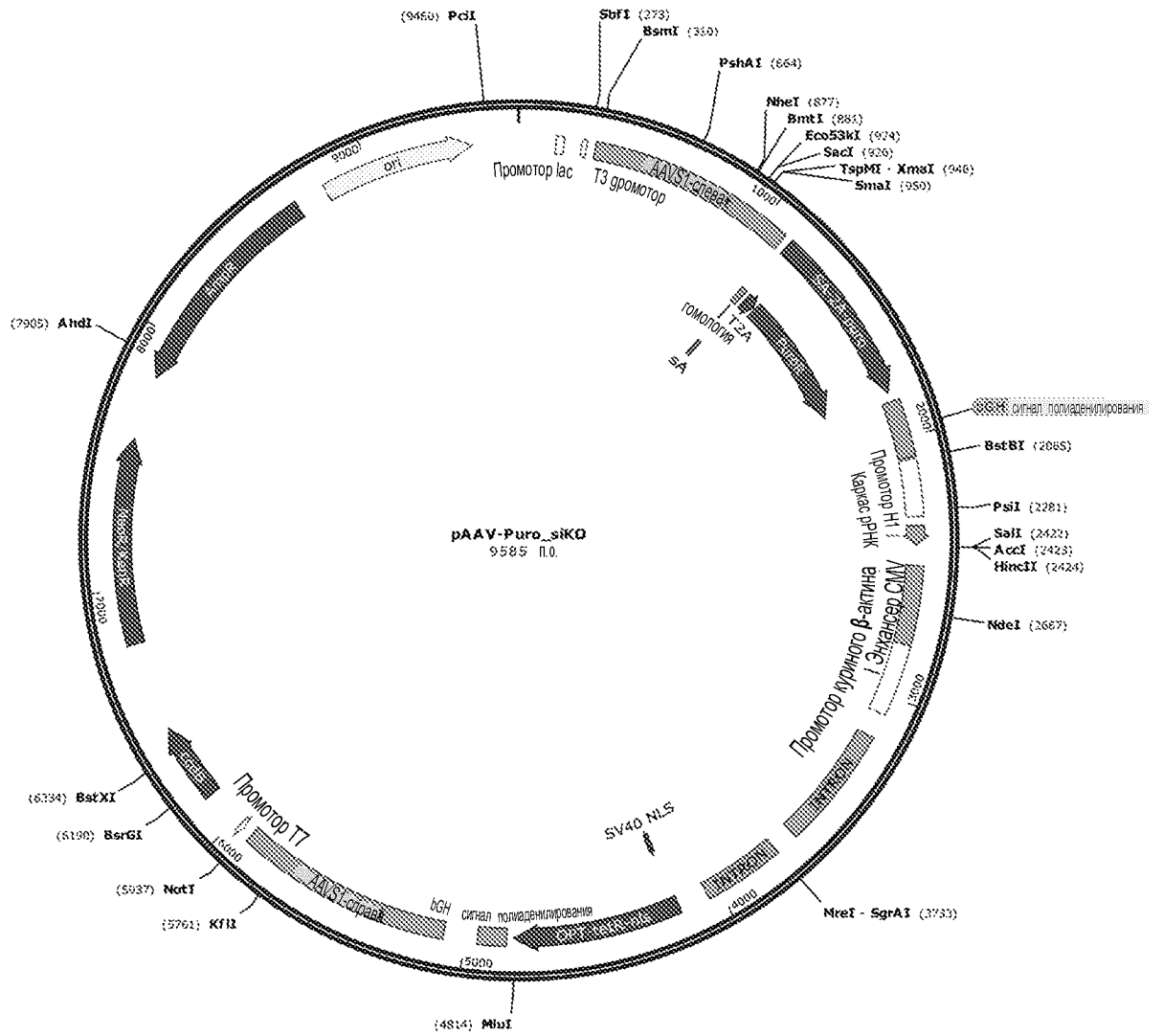
ФИГ. 27



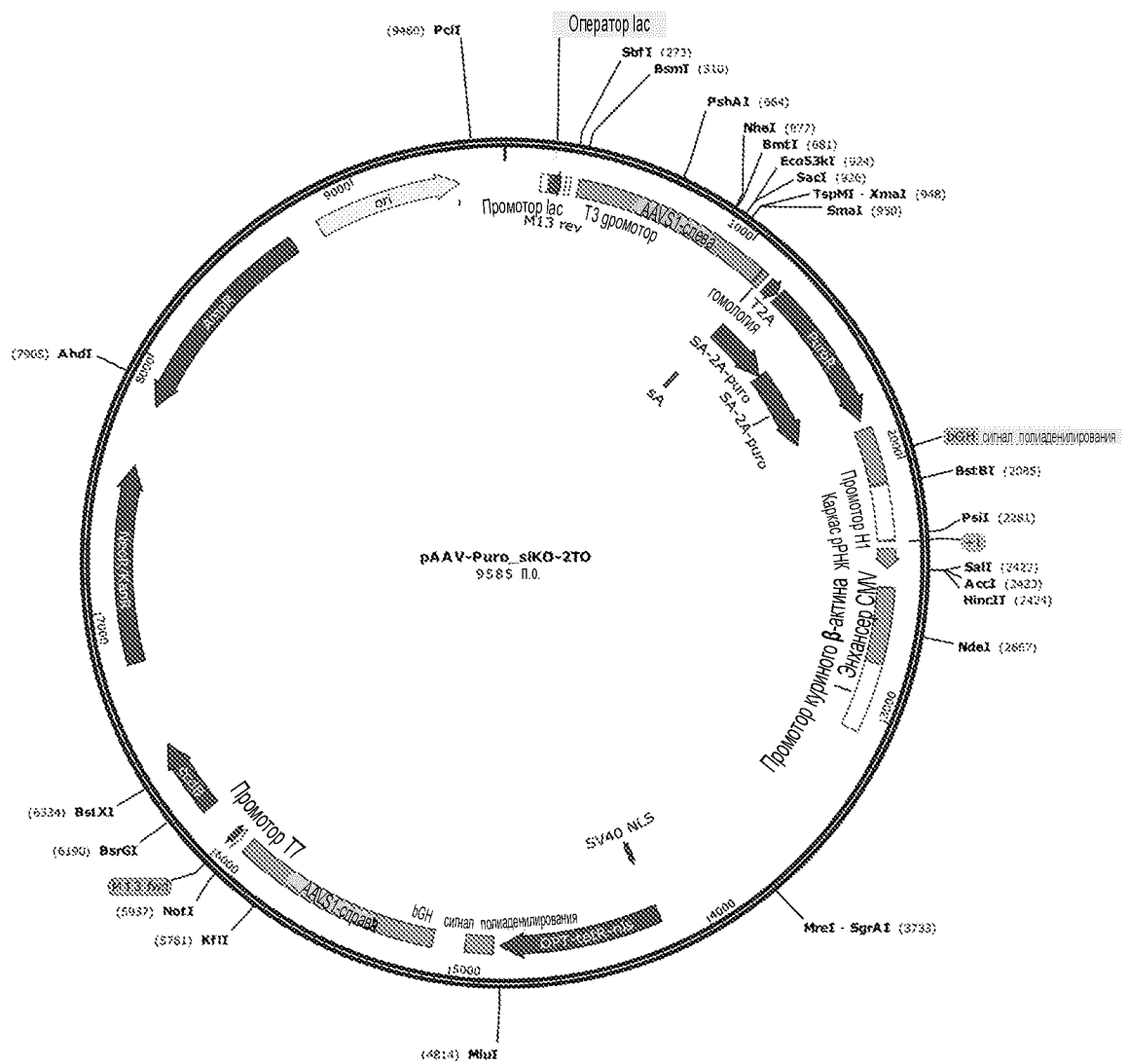
ФИГ. 28



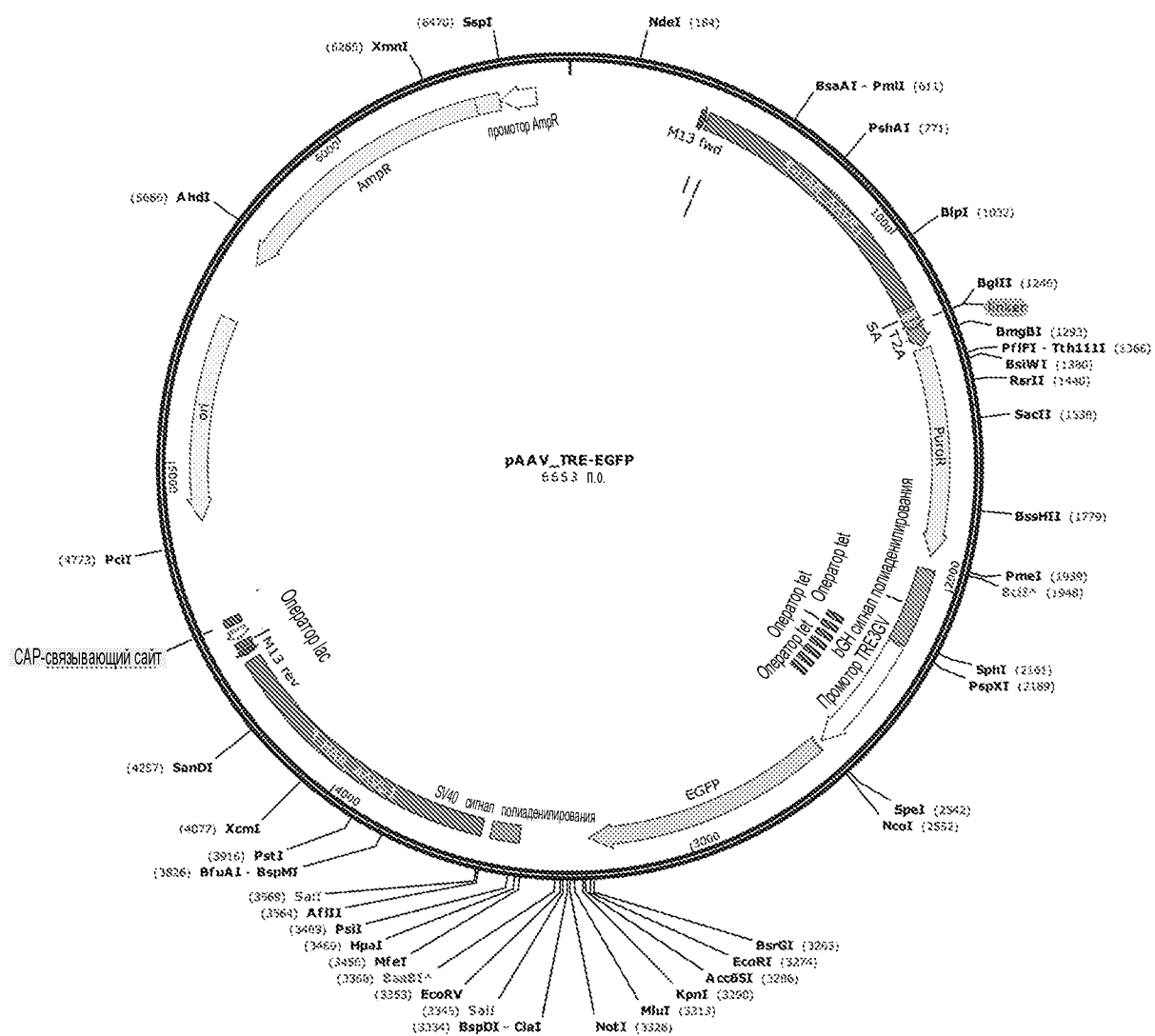
ФИГ. 29



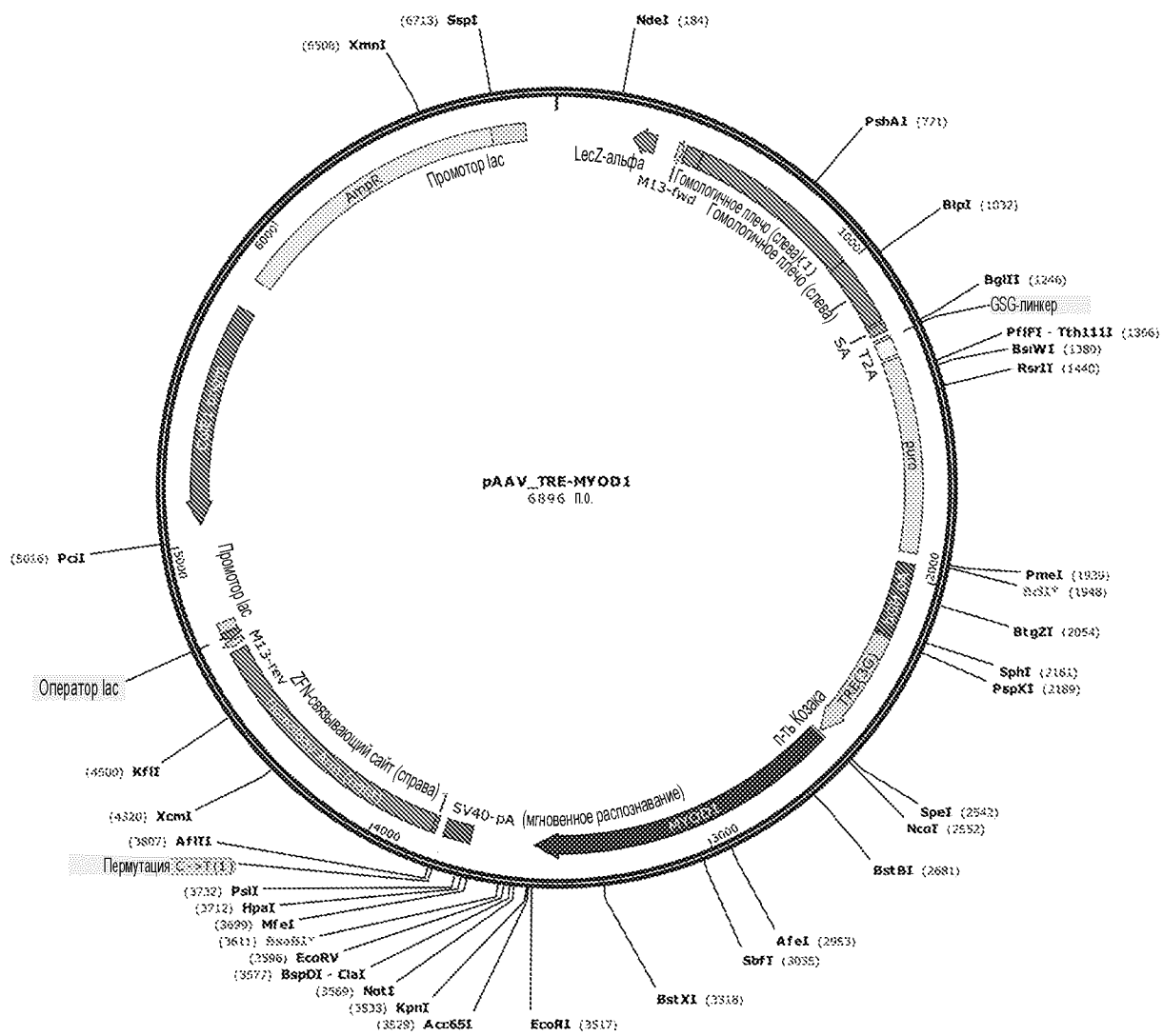
ФИГ. 30



ФИГ. 31



ФИГ. 32



ФИГ. 33