

(19)



Евразийское  
патентное  
ведомство

(21) 201691990 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки  
2017.05.31

(51) Int. Cl. C12N 15/10 (2006.01)  
C12N 15/90 (2006.01)  
C12N 9/22 (2006.01)

(22) Дата подачи заявки  
2015.05.01

(54) ПОЛУЧЕНИЕ БИБЛИОТЕК ВАРИАНТОВ БЕЛКОВ, ЭКСПРЕССИРОВАННЫХ  
В ЭУКАРИОТИЧЕСКИХ КЛЕТКАХ, И ПРИМЕНЕНИЕ ИХ ДЛЯ ОТБОРА  
СВЯЗЫВАЮЩИХ МОЛЕКУЛ

(31) 1407852.1

(32) 2014.05.02

(33) GB

(86) PCT/GB2015/051287

(87) WO 2015/166272 2015.11.05

(88) 2015.12.30

(71) Заявитель:  
ИОНТАС ЛИМИТЕД (GB)

(72) Изобретатель:  
Маккафферти Джон, Дайсон Майкл,  
Партибан Котай (GB)

(74) Представитель:  
Нилова М.И. (RU)

(57) Настоящее изобретение относится к способам получения библиотек эукариотических клеток, кодирующих набор связывающих молекул ("связывающие молекулы"), причем в указанных способах используется сайт-специфическая нуклеаза для направленного расщепления клеточной ДНК с целью улучшения сайт-специфического встраивания генов связывающих молекул посредством эндогенных клеточных механизмов репарации. Получают популяции эукариотических клеток, в которых набор генов, кодирующих связывающие молекулы, встроен в требуемый локус в клеточной ДНК (например, геномный локус), что обеспечивает экспрессию кодируемой связывающей молекулы, в результате чего получают популяцию клеток, экспрессирующих разные связывающие молекулы.



A1

201691990

201691990

A1

# **ПОЛУЧЕНИЕ БИБЛИОТЕК ВАРИАНТОВ БЕЛКОВ, ЭКСПРЕССИРОВАННЫХ В ЭУКАРИОТИЧЕСКИХ КЛЕТКАХ, И ПРИМЕНЕНИЕ ИХ ДЛЯ ОТБОРА СВЯЗЫВАЮЩИХ МОЛЕКУЛ**

## **ОБЛАСТЬ ТЕХНИКИ**

Настоящее изобретение относится к способам получения библиотек эукариотических клеток (например, млекопитающих) для скрининга и/или отбора связывающих молекул, таких как антитела. Библиотеки могут быть использованы для охвата и дисплея широкого набора связывающих молекул, что делает возможным осуществление скрининга связывающих молекул для отбора одной или более связывающих молекул, обладающих требуемым свойством, таким как специфичность по отношению к молекуле-мишени. Настоящее изобретение, в частности, относится к способам введения донорной ДНК, кодирующей связывающие молекулы, в эукариотические клетки для получения библиотеки клеток, в которой требуемое количество молекул донорной ДНК точно встроено в требуемый локус или локусы в клетках.

## **УРОВЕНЬ ТЕХНИКИ**

Методы белковой инженерии позволяют создать большие разнообразные популяции родственных молекул (например, антитела, белки, пептиды), из которых могут быть отобраны отдельные варианты с новой или улучшенной связывающей способностью или каталитическими свойствами. Возможность получения больших популяций эукариотических клеток, в особенности клеток млекопитающих, в которых каждая клетка экспрессирует отдельное антитело, пептид или генно-инженерный белок имеет большое значение для выявления связывающих молекул в требуемыми свойствами.

Основной принцип технологии дисплея заключается в связи связывающей молекулы с генетической информацией, кодирующей эту молекулу. Связывающую способность связывающей молекулы используют для выделения гена, который ее кодирует. Такие же основные принципы применимы ко всем формам технологии дисплея, включая бактериофаговый дисплей, бактериальный дисплей, ретровирусный дисплей, бакуловирусный дисплей, рибосомный дисплей, дрожжевой дисплей и дисплей на клетках высших эукариот, как, например, клетки млекопитающих [1, 2, 3, 4].

Примером технологии дисплея послужил дисплей антител на нитевидном бактериофаге (фаговый дисплей антител), который за последние 24 года обеспечил важные

инструменты для открытия и конструирования новых связывающих молекул, включая создание человеческих терапевтических антител. С использованием фагового дисплея осуществляется представление молекул антител на поверхности частиц нитевидного бактериофага путем клонирования гена, кодирующего антитело или фрагмент антитела в рамку считывания с геном, кодирующим белок оболочки фага. Гены антител первоначально клонируют в клетки *E.coli* таким образом, что каждая бактерия кодирует одно антитело. Нарботка бактериофагов в бактериях с использованием стандартных способов позволяет получить бактериофаги, несущие на поверхности фрагмент антитела, а внутри - ген, кодирующий антитело. Совокупность бактерий или бактериофагов, полученных из них, называется «библиотекой антител». Использование фагового дисплея делает возможным обогащение популяции антител и связанных с ними генов путем взаимодействия бактериофага, презентирующего антитело, с представляющей интерес молекулой-мишенью.

Для выделения бактериофага, несущего на своей поверхности связывающую молекулу, распознающую представляющую интерес мишень, молекулу-мишень необходимо иммобилизовать на поверхности подложки, используемой для отбора, или она должна поддаваться выделению из раствора с помощью вторичных реагентов, например, биотинилированный белок-мишень, выделенный из раствора с использованием микрогранул, покрытых стрептавидином. После инкубации библиотеки бактериофагов, экспрессирующих связывающую молекулу, с молекулой-мишенью несвязавшиеся фаговые частицы удаляют. Это включает отмывку матрицы, к которой прикреплена мишень (и ассоциированный бактериофаг), для удаления несвязавшихся бактериофагов. Связавшиеся бактериофаги с ассоциированными с ними генами антител могут быть выделены и и/или использованы для заражения бактериальных клеток-хозяев. С использованием подхода, обозначенного выше, становится возможным обогащение субпопуляции клонов бактериофагов, способных к связыванию с предпочтительной молекулой-мишенью. Библиотеки фагового дисплея, как было показано, служат богатым источником разнообразия антител, обеспечивая получение сотен уникальных антител к одной мишени [5,6,7].

В историческом плане системы дисплея для выделения антител с новой специфичностью связывания были основаны на прокариотических системах и, в частности, на дисплее одноцепочечных Fv (scFv) и в меньшей степени в виде Fab на бактериофаге. Дисплей связывающих молекул на поверхности бактерий был описан, но широко не использовался, а его применение в значительной мере было ограничено пептидным дисплеем или дисплеем фрагментов антител, предварительно обогащенных в отношении

связывающих молекул посредством иммунизации [8]. Несмотря на эффективность прокариотических систем дисплея, включая фаговый дисплей, существуют ограничения. После отбора с помощью фагового или рибосомного дисплея идентифицируют гены, кодирующие отдельные связывающие молекулы, путем введения отобранной популяции генов в бактерии, высева популяций бактерий, выделения колоний, экспрессии связывающих молекул в супернатанте или периплазме и идентификации положительных клонов в анализах связывания, таких как иммуноферментные или флуоресцентные иммуносорбентные анализы (ИФА). Несмотря на идентификацию связывающих молекул данный подход не дает информации о степени экспрессии и аффинности связывания полученных в результате клонов. Таким образом, несмотря на потенциальную возможность создания множества связывающих молекул, возможность скрининга вырабатываемого продукта ограничена необходимостью выделения колоний, манипуляций с жидкостью и т. д. в сочетании с ограниченной первичной информацией об относительном уровне экспрессии и аффинности.

Дисплей связывающих молекул на поверхности эукариотических клеток потенциально позволяет преодолеть некоторые из этих проблем. В сочетании с проточной цитометрией дисплей на поверхности эукариотических клеток обеспечивает быстрый и высокоэффективный отбор. Становится возможным исследование миллионов клеточных клонов, экспрессирующих разные связывающие молекулы на их поверхности. Дисплей на клеточной поверхности послужил примером дисплея фрагментов антител, имеющих формат scFv, на поверхности дрожжевых клеток. В широко распространенном методе дисплея на поверхности дрожжевых клеток используются дрожжевые белки-агглютинины (Aga1p и Aga2p). Как было описано Chao *et al.* [9], гены, кодирующие набор scFv, слиты с геном субъединицы дрожжевого агглютинина Aga2p. Субъединица Aga2p затем присоединяется к субъединице Aga1p, содержащейся в клеточной оболочке, посредством дисульфидных связей. Дрожжевые клетки, экспрессирующие связывающую молекулу, специфическую по отношению к мишени, могут быть идентифицированы с помощью проточной цитометрии с использованием прямо или опосредованно меченой молекулы-мишени. Например, к клеткам могут быть добавлены биотинилированные молекулы-мишени, а связывание с клеточной поверхностью может быть установлено с использованием стрептавидина-фикоэритрина. С использованием предельных концентраций становится возможным отличить в составе популяции те клоны, которые экспрессируют связывающие молекулы с более высокой аффинностью, поскольку эти клоны будут связываться с большим количеством молекул-мишеней и, следовательно, будут иметь более интенсивную флуоресценцию. Обычно каждая дрожжевая клетка будет



экспрессировать от 10 000 до 100 000 копий одного scFv на клеточной поверхности. Для контроля за изменением экспрессии scFv на поверхности разных клеток Chao et al. использовали флуоресцентно меченое антитело к метке tag, чтобы измерить уровень экспрессии антител на поверхности каждой клетки, что позволило нормализовать изменение уровня экспрессии. Данный подход, следовательно, позволяет отличить дрожжевые клетки, экспрессирующие высокоаффинные связывающие молекулы, от тех клеток, которые экспрессируют низкие уровни антител с более низкой аффинностью. Таким образом, с использованием сортировки флуоресцентно-активированных клеток (FACS) можно разделить клеточные клоны по аффинности и/или уровню экспрессии кодируемой связывающей молекулы.

Было доказано, что эукариотические системы также являются более эффективными, чем прокариотические системы в отношении дисплея многоцепочечных фрагментов антител и, в частности, больших фрагментов, таких как полноразмерные IgG, FAb или молекулы scFv, слитые с Fc-доменами (слитые молекулы scFv-Fc). Способы, основанные на применении микрогранул или проточной сортировки клеток, которые описаны выше для дрожжевых клеток, также могут быть использованы для отбора антител из дисплейных библиотек, основанных на применении клеток высших эукариот, таких как клетки млекопитающих. Возможность форматирования дисплейных библиотек и прямого отбора в форме IgG, Fab или в форме слитых молекул scFv-Fc на клетках млекопитающих может быть дополнительным преимуществом по сравнению с дрожжевым дисплеем. Гликозилирование, экспрессия и аппарат секреции бактериальных и дрожжевых клеток отличается от высших эукариот, что обуславливает получение антител с другими посттрансляционными модификациями, чем те, что встречаются на клетках млекопитающих. Поскольку производство антител для исследовательских, диагностических и терапевтических целей обычно осуществляется на клетках млекопитающих, дисплей на клетках млекопитающих (или других клетках высших эукариот, как, например, линии клеток беспозвоночных животных, птиц или растений) может дать лучшее представление о потенциальных проблемах или преимуществах для последующего производства, например, идентификации клонов с оптимальными свойствами экспрессии. В дополнение к этому, антитела, обнаруженные в рамках дисплея на клетках высших эукариот и, в особенности, клетках млекопитающих могут непосредственно применяться в анализах клеток, содержащих репортерные гены, в отсутствие глубокой очистки и осложнений, связанных с примесями из бактериальных и дрожжевых клеток. Кроме того, библиотеки связывающих молекул могут быть экспрессированы непосредственно в эукариотических клетках, содержащих репортерные

гены, как, например, клетках млекопитающих для идентификации клонов, которые непосредственно влияют на клеточный фенотип.

Несмотря на вышеупомянутые преимущества, связанные с дисплейными библиотеками, полученными в эукариотических клетках, имеются значительные проблемы с созданием библиотек связывающих молекул в эукариотических клетках, особенно у высших эукариот. Введение набора экзогенных генов («трансгенов») для экспрессии у высших эукариот является более сложным, чем у дрожжей и бактерий. С клетками высших эукариот сложнее работать и получить в большом количестве, а эффективность трансформации является более низкой. Обычно такие библиотеки намного меньше по размеру. В дополнение к этому, введенная ДНК встраивается случайным образом в геном, что обуславливает эффект положения мозаичного типа. Кроме того, донорная ДНК, введенная в клетки млекопитающих с помощью стандартных способов трансфекции или электропорации, встраивается в виде линейной матрицы с различным числом копий трансфецированного трансгена. Следовательно, при введении ДНК, кодирующей набор генов антител, потенциально может произойти попадание многих генов антител в каждую клетку, что приведет к экспрессии множества разных антител в одной клетке. В дополнение к этому, при наличии множества генов антител будет уменьшаться относительный уровень экспрессии того или иного антитела и происходить выделение нецелевых генов, вследствие чего степень обогащения специфическими клонами будет уменьшаться.

Хотя дисплей библиотек связывающих молекул на поверхности клеток высших эукариот является более затруднительным, все же ранее были описаны некоторые примеры. В ранней публикации, основанной на использовании дисплея на клетках млекопитающих полученных в результате иммунизации человека IgG, понадобилось 3 раунда отбора (включая временную трансфекцию, сортировку клеток, выделение ДНК и повторную трансфекцию) для достижения антиген-специфическими клетками в 450 раз, что дало в среднем обогащение в 7,6 раза каждом раунде [10]. Аналогичным образом, также была описана временная экспрессия полученных в результате иммунизации библиотек молекул, экспрессированных в составе векторов, реплицируемых в виде эписом, с антителами, имеющими формат scFv [11, 12] или IgG [13].

Был описан целый ряд подходов для введения одного или ограниченного количества генов антител в каждую клетку. К ним относится разведение ДНК или смешивание с несущей ДНК [13], но это относительно неконтролируемый способ регуляции количества копий вводимых генов, и уменьшение количества вводимой ДНК будет оказывать отрицательное влияние на размер библиотеки. Введение генов антител с помощью вирусных векторов оказалось другим решением для контроля введения множества генов

антител на клетку. Таким способом из нескольких сотен человеческих В-лимфоцитов, полученных путем иммунизации, были составлены библиотеки экспонированных на клеточной поверхности молекул и дополнительно обогащены с помощью проточной сортировки антиген-специфических В-клеток [14]. Гены антител из этого обогащенного пула были получены в формате scFv, клонированы в экспрессирующую систему альфавируса Синдбис и введены в клетки ВНК с использованием заражения с низкой множественностью.

Breous-Nyström *et al.* [15] использовали последовательное заражение ретровирусами для введения ограниченного набора, состоящего из 91 гена варибельной области каппацепи антител, за которым следовал набор генов тяжелой цепи антител 6 здоровых доноров, в линию мышинных пре-В-клеток (1624-5). Инфекционный ретровирус был получен с использованием системы V-Pack, основанной на вирусе мышинного лейкоза Молони (Stratagene). Для того, чтобы обеспечить преимущественное образование клонов с однокопийными вставками, была подобрана множественность заражения, которая обеспечивала заражение приблизительно 5 % клеток. Основным недостатком этих подходов является то, что встраивание в геном происходит случайным образом, что приводит к потенциальному варьированию уровня транскрипции на основе транскрипционной активности сайта встраивания. Другой недостаток во всех этих случаях заключается в том, что встраивание генов антител контролируется ограниченным заражением или трансфекцией, что влияет на размер библиотеки.

Ранее было описано сайт-специфическое встраивание трансгенов, осуществляемое рекомбиназой. Рекомбиназы представляют собой ферменты, что катализируют обменные реакции между молекулами ДНК, содержащими фермент-специфические распознаваемые последовательности. Например, рекомбиназа Cre (полученная из сайт-специфической системы рекомбинации *E. coli*) или рекомбиназа Flp (основанная на системе рекомбинации *Saccharomyces cerevisiae*) действуют на их специфические сайты распознавания IoxP длиной 34 п.о. и сайт-мишень рекомбинации Flp (FRT) длиной 34 п.о., соответственно [16]. Рекомбиназы главным образом были использованы в клеточной инженерии для катализа сайт-специфического встраивания. В целом ряде исследований из работы Chen Zhou [17, 18, US7884054] было описано опосредованное рекомбиназами сайт-специфическое встраивание генов антител в геном клеток млекопитающих с использованием рекомбиназы Flp в составе системы «Flp-In», ([http://tools.lifetechnologies.com/content/sfs/manuals/flpinsystem\\_man.pdf](http://tools.lifetechnologies.com/content/sfs/manuals/flpinsystem_man.pdf)). В системе Flp-In используется множество клеточных линий с одним сайтом FRT, ранее введенным в их геном. За счет экспрессии фермента рекомбиназы Flp становится возможным прямое

встраивание экспрессирующих плазмид, содержащих в своем составе сайт рекомбинации FRT, в этот прединтеграционный сайт FRT в клетках-мишенях.

С использованием системы Flp-In Zhou *et al.* [17] внесли экспрессирующую плазмиду, содержащую сайт FRT, в линию клеток яичника китайского хомячка (CHO), содержащую сайт FRT (клетки CHO-F). В их работе описано создание дисплейной библиотеки, в которой были мутированы 4 остатка в составе существующего антитела к лиганду OX40. Был проведен скрининг библиотеки с использованием FACS для идентификации антител, обладающих сродством к лигандам на клеточной поверхности. Общий успех в создании антител с улучшенными свойствами был ограничен выделением одного антитела с улучшенными свойствами. Отсутствуют сообщения касательно полученного количества уникальных клонов клеток млекопитающих.

В последующей статье Li *et al.* в 2012 году [18] было описано использование лимфоцитов пациента с гепатитом В для создания дисплейных библиотек антител. Были получены отдельные библиотеки с генами тяжелой и легкой цепей, полученными у донора, который был иммунизирован HBSAg, причем отдельно сообщалось о библиотеках с размером  $1,02 \times 10^6$  и  $1,78 \times 10^5$ , соответственно. Затем была получена вторичная библиотека, включающая как тяжелые, так и легкие цепи, размер которых по имеющимся сообщениям составлял  $4,32 \times 10^5$ . По имеющимся сообщениям FACS-анализ показал, что у около 40 % клеток на клеточной поверхности были выявлены полноразмерные антитела. Скрининг библиотеки с помощью FACS позволил выявить антитела, связывающиеся с HBSAg. Как было установлено, из выборки, состоящей из 8 отобранных антител библиотеки, которые связывались с антигеном, шесть из них были одинаковыми, так что в целом было выявлено три уникальных клона, обладающих специфичностью к HBSAg.

Довольно ограниченный успех этой работы может быть следствием того, что система Flp-In предназначена для точного встраивания в ограниченное количество клонов, а не для создания большой библиотеки. Следовательно, имеется потенциальное противоречие между достижением точности встраивания и достижением максимального размера библиотеки. В системе Flp-In используется мутантная форма рекомбиназы Flp в плазмиде pOG44, которая при 37 °C обладает только 10 % активности нативной рекомбиназы Flp [19]. Был выявлен вариант рекомбиназы Flp (Flp<sub>e</sub>) с лучшей термостабильностью и более высокой активностью, чем дикий тип [19, 20]. Он дополнительно был улучшен с помощью оптимизации кодонов с образованием Flp<sub>o</sub>, кодируемого в составе плазмиды cCAGGS-Flp<sub>o</sub> (GeneBridges номер по каталогу A203). Однако в руководстве по применению Flp-In отмечено: «при создании экспрессирующих клеточных линий Flp-In™ важно помнить, что речь идет об относительно редком явлении рекомбинации, поскольку желательно, чтобы

рекомбинация и встраивание конструкции  $\text{pcDNA}^{\text{TM}}5/\text{FRT}$  осуществлялись только через сайт FRT и за ограниченный период времени. В данном случае использование высокоэффективной рекомбиназы F1r является выгодным и может снизить частоту возникновения других нежелательных явлений рекомбинации... Для увеличения вероятности получения клонов с одной интегрированной последовательностью понадобится снижение эффективности трансфекции путем ограничения количества плазмидной ДНК, которую трансфецируют». Это было повторно отражено в источнике Buchholz *et al.*, 1996 [19]: «FLP может быть особенно полезной для применений, которые не зависят от эффективности, но зависят от строгой регуляции».

В модельных экспериментах и с использованием «инструкций, описанных в руководстве» Zhou *et al.* (2010) [17] действительно было продемонстрировано, что в >90 % клонов встречаются однокопийные вставки. При создании этой библиотеки, однако было использовано относительно большое количество экспрессирующей плазмиды (2,5-3,2 пг на  $10^6$  клеток) и избыточное количество донорных молекул по сравнению с количеством плазмиды, кодирующей рекомбиназу pOG44 [17, 18]. Рекомендуется использовать систему F1r-In в соотношении по меньшей мере 9:1 в пользу рекомбиназы, кодирующей эту плазмиду, по сравнению с экспрессирующей плазмидой. Однако при попытке увеличить размер библиотеки за счет трансфекции большого количества ДНК наблюдается тенденция к случайному встраиванию вносимой плазмиды [21]. Во всех этих исследованиях отсутствуют сообщения о точности встраивания и количестве интегрированных последовательностей на клетку в условиях «создания библиотеки».

При направляемом нуклеазой встраивании генов используется сайт-специфическая нуклеаза для расщепления клеточной ДНК в определенном положении. Ранее было показано, что это приводит к повышению скорости гомологической рекомбинации по меньшей мере в 40000 раз, а также репарации посредством механизмов негомологического соединения концов. Такое повышение эффективности сайт-специфического встраивания ранее не было использовано или не предполагалось, что оно способно решить проблемы, связанные с созданием библиотек связывающих молекул.

В источнике US20100212035 описаны способы получения грызунов, клетки которых способны к экспрессии экзогенного антитела, путем целенаправленного воздействия на локус иммуноглобулина эмбриона млекопитающих с использованием мегануклеазы для прямого встраивания донорной ДНК. Описана возможность создания библиотек вариантов мегануклеаз для создания молекул с новой специфичностью расщепления ДНК, но не предполагается использование мегануклеаз с целью создания библиотек связывающих молекул.

В источнике WO 2013/190032 A1 описано встраивание генов в определенный локус (Fer1L4), ранее модифицированный экзогенной ДНК («a site specific integration» SSI host cell, клетки-реципиенты для проведения «сайт-специфического встраивания»), для включения сайтов рекомбиназ, как, например, сайтов IoxP и FRT для опосредованного рекомбиназами сайт-специфического введения генов. Создание библиотеки с использованием сайт-направленных нуклеаз не описано.

В источнике WO 2012/167192 A2 описано направленное введение генов в локус, который затем может быть подвергнут избирательной амплификации. Способы, основанные на использовании сайт-направленных нуклеаз, используются для направленного воздействия на данный локус. Создание библиотеки с использованием сайт-направленных нуклеаз не описано.

В источнике US 2009/0263900A1 описаны молекулы ДНК, содержащие гомологичные участки, и их применение в способах гомологичной рекомбинации. Создание библиотеки с использованием сайт-направленных нуклеаз не описано.

В источнике WO 2011/100058 описаны способы встраивания нуклеиновой кислоты в геном, которые исключают необходимость использования длинных гомологичных участков, а вместо этого они основаны на микрогомологии или «липких концах» в геноме и донорной молекуле для обеспечения прямого встраивания. Создание библиотеки с использованием сайт-направленных нуклеаз не описано.

В источнике WO 20122/090804 описаны способы встраивания множества генов или множества копий одного и того же гена с использованием разных нуклеаз с доменом «цинковые пальцы» (ZFN) в последовательных раундах. Создание библиотеки с использованием сайт-направленных нуклеаз не описано.

В источнике WO2014/039872 описаны способы инженерии растительных клеток, содержащих «сайт посадки», в которое встраивается донорная ДНК путем гомологичной рекомбинации или негомологичного соединения концов с использованием сайт-направляемых нуклеаз. Библиотеки бактериальных искусственных хромосом (BAC) используются для первоначального клонирования донорной ДНК. Эти библиотеки упоминаются применительно к способам секвенирования с использованием Illumina. Создание библиотеки с использованием сайт-направленных нуклеаз не описано.

В источнике W02007/047859 A2 описаны способы конструирования мегануклеаз с заданной специфичностью и их применение для направленного изменения локусов генома. Описаны библиотеки мутантных форм мегануклеаз, которые могут содержать мегануклеазы с новой специфичностью. Создание библиотеки с использованием сайт-направленных нуклеаз не описано.

В источнике US2014/0113375 A1 описана система временной экспрессии для создания одноцепочечных последовательностей ДНК, гомологичных геномной последовательности-мишени, которая может быть транспортирована в ядро для изменения генетической информации геномной последовательности-мишени через пути репарации ДНК или гомологичной рекомбинации. Предполагается, что может быть создана «библиотека» мутаций путем неточной обратной транскрипции вводимой (не относящейся к библиотечной) ДНК. Дисплей на клетках млекопитающих и отбор молекул со связывающей активностью не описаны.

В источнике US2012/0277120 описаны способы и композиции для одновременного встраивания множества экзогенных нуклеиновых кислот в ходе одной реакции трансформации с использованием природной системы гомологичной рекомбинации в клетках дрожжей, причем указанная рекомбинация дополнительно может быть улучшена путем индукции направленных двухнитевых разрывов в геноме клеток-хозяев в заданных сайтах встраивания. Указанные способы позволяют избежать необходимость в осуществлении многократных циклов генно-инженерных манипуляций для встраивания множества молекул ДНК, например, для конструирования функциональных метаболических путей в промышленно важных микроорганизмах, таких как дрожжи. Дисплей или экспрессия библиотек связывающих молекул, применение высших эукариот и отбор молекул со связывающей активностью не описаны.

Для полной реализации потенциала дисплея антител на клетках млекопитающих и других высших эукариот существует необходимость в системе для создания больших библиотек, в которой бы объединялись свойства, обеспечивающие точность встраивания в заранее заданный сайт и эффективность, которая позволяет создавать большие библиотеки.

## КРАТКОЕ ОПИСАНИЕ СУЩНОСТИ ИЗОБРЕТЕНИЯ

Авторы решили указанную проблему создания больших библиотек связывающих молекул, включающих один или два гена связывающих молекул на клетку, путем использования направляемого нуклеазой встраивания популяции генов, кодирующих связывающие молекулы. Настоящее изобретение Таким образом, позволяет получить популяции эукариотических клеток, причем набор генов, кодирующих связывающие молекулы, встраивается в фиксированный локус в геноме, что обеспечивает экспрессию кодируемой связывающей молекулы, в результате чего получают популяцию клеток, экспрессирующих разные связывающие молекулы.

Настоящее изобретение относится к способам получения библиотек эукариотических клеток, кодирующих набор связывающих молекул («связывающие молекулы»), причем в способах используется сайт-специфическая нуклеаза для направленного расщепления клеточной ДНК для улучшения сайт-специфического встраивания генов связывающих молекул посредством эндогенных клеточных механизмов репарации. Сайт-специфические нуклеазы обеспечивают точное введение донорной ДНК, кодирующей связывающие молекулы, в один или более определенных локусов в геноме эукариотических клеток или других ДНК эукариотических клеток. Настоящее изобретение относится к способам получения популяции эукариотических клеток, в которых набор генов, кодирующих связывающие молекулы, встроен в требуемый локус в клеточной ДНК (например, геномный локус), что обеспечивает экспрессию кодируемой связывающей молекулы, в результате чего получают популяцию клеток, экспрессирующих разные связывающие молекулы.

Создание библиотек связывающих молекул в эукариотических клетках согласно настоящему изобретению обладает преимуществами по сравнению с подходами, основанными на использовании рекомбиназ, для сайт-направленного включения экспрессирующих конструкций. В настоящем изобретении используется расщепление клеточной ДНК сайт-специфическими нуклеазами для решения проблем, которые ранее были связаны с составлением больших наборов генов связывающих молекул в эукариотических клетках и, в частности, высших эукариот. Настоящее изобретение позволяет эффективно получать большие популяции клеточных клонов, при этом каждый клон экспрессирует отдельные связывающие молекулы, встроенные в фиксированный локус в клеточной ДНК. Из этих библиотек клеточных клонов становится возможным выделить гены, кодирующие новые связывающие или модифицирующие функцию белки и пептиды.

В отличие от направляемого рекомбиназой обмена ДНК в подходе настоящего изобретения используется сайт-специфическое расщепление клеточной (например, геномной) ДНК с последующим применением природных механизмов репарации для встраивания донорной ДНК, кодирующей связывающие молекулы. После расщепления клеточной ДНК в последовательности, распознаваемой сайт-специфической нуклеазой («распознаваемая последовательность»), разрывы в клеточной ДНК подвергаются репарации с использованием механизмов, таких как гомологичная рекомбинация или негомологичное соединение концов (NHEJ). Создание сайт-специфических разрывов в клеточной ДНК повышает эффективность включения экзогенной донорной ДНК,



обеспечивая получение больших популяций клеток с генами связывающих молекул, встроенными в фиксированный локус.

До настоящего времени сайт-специфические нуклеазы, такие как мегануклеазы, ZFN, нуклеазы TALE и системы CRISPR/Cas, были использованы для эффективного получения клеток с модификациями эндогенных генов или для введения репортерных генов с целью изучения функции клеток. Имеются также примеры, когда для получения антител (путем выделения из культуральной среды) было использовано направляемое нуклеазой целенаправленное встраивание генов, кодирующих одиночные секретируемые антитела, [21, 22,].

Настоящее изобретение позволяет упростить создание больших библиотек, при этом обеспечивая целенаправленное встраивание ДНК в один локус или ограниченное количество определенных локусов генов. Встраивание донорной ДНК в один или более фиксированных локусов нормализует транскрипцию по сравнению со случайным встраиванием различного количества трансгенов и обеспечивает отбор клонов антител, исходя из свойств трансляции и стабильности связывающей молекулы как таковой. Точное встраивание донорной ДНК в заранее заданное местоположение или местоположения в клеточной ДНК приводит к относительно равномерным уровням транскрипции связывающих молекул в библиотеке и высокой эффективности введения донорной ДНК, делает клеточные популяции, полученные с помощью способов по настоящему изобретению, особенно пригодными в качестве библиотек для дисплея и отбора связывающих молекул.

Таким образом, способы по настоящему изобретению позволяют получить высококачественные библиотеки связывающих молекул в эукариотических клетках, которые могут подвергаться скринингу для идентификации клеток, кодирующих и экспрессирующих специфическую молекулу, связывающуюся с представляющей интерес мишенью.

В различных аспектах настоящее изобретение относится к новым и усовершенствованным способам получения библиотек эукариотических клеток, библиотек самих по себе, выделению требуемых связывающих молекул, кодирующих нуклеиновую кислоту, и клеток из библиотек, и применениям библиотек, как, например, для экспрессии и скрининга связывающих молекул и для скрининга эффектов связывающих молекул. Описаны различные способы получения библиотек *in vitro* и применение библиотек *in vitro* или *in vivo*.

Настоящее изобретение относится к способу получения библиотеки клонов эукариотических клеток, содержащих ДНК, кодирующую широкий набор связывающих

молекул, причем способ включает применение сайт-специфической нуклеазы для направленного расщепления ДНК эукариотических клеток для улучшения сайт-специфического встраивания генов связывающих молекул в клеточную ДНК посредством эндогенных клеточных механизмов репарации ДНК.

Способ получения библиотеки клонов эукариотических клеток, содержащих ДНК, кодирующую широкий набор связывающих молекул, может включать:

получение молекул донорной ДНК, кодирующих связывающие молекулы, и эукариотических клеток, введение донорной ДНК в клетки и обеспечение присутствия в клетке сайт-специфической нуклеазы, причем указанная нуклеаза расщепляет клеточную ДНК с образованием сайта встраивания, в котором донорная ДНК встраивается в клеточную ДНК, при этом встраивание осуществляется посредством эндогенных клеточных механизмов репарации ДНК.

В случае мультимерных связывающих молекул, содержащих по меньшей мере первую и вторую субъединицу (т.е. отдельные полипептидные цепи, как, например, VH- и VL-домены антител, представленные в формате Fab или IgG), многие субъединицы могут кодироваться одной и той же молекулой донорной ДНК. Однако возможно потребуются встраивание разных субъединиц в отдельные локусы, в случае чего субъединицы могут быть введены в составе отдельных молекул донорной ДНК. Они могут быть встроены в ходе одного и того же цикла направляемого нуклеазой встраивания или могут быть встроены последовательно с использованием направляемого нуклеазой встраивания в течение одного или обоих этапов встраивания.

Способы получения библиотек клонов эукариотических клеток, кодирующих мультимерные связывающие молекулы, могут включать:

получение эукариотических клеток, содержащих ДНК, кодирующую первую субъединицу, и получение молекул донорной ДНК, кодирующих вторую субъединицу связывающей молекулы,

введение донорной ДНК в клетки и обеспечение присутствия в клетке сайт-специфической нуклеазы, причем указанная нуклеаза расщепляет распознаваемую последовательность в клеточной ДНК с образованием сайта встраивания, в котором донорная ДНК встраивается в клеточную ДНК, при этом встраивание осуществляется посредством эндогенных клеточных механизмов репарации ДНК, в результате чего получают рекомбинантные клетки, которые содержат донорную ДНК, встроенную в клеточную ДНК. Эти рекомбинантные клетки будут содержать ДНК, кодирующую первую и вторую субъединицы мультимерной связывающей молекулы, и могут быть

культивированы для экспрессии обеих субъединиц. Мультимерные связывающие молекулы получают путем экспрессии и сборки отдельно кодируемых субъединиц.

В описанном выше примере направляемое нуклеазой встраивание используется для встраивания ДНК, кодирующей вторую субъединицу, в клетки уже содержащие ДНК, кодирующую первую субъединицу. Первая субъединица может быть введена заранее с использованием способов по настоящему изобретению или любых других подходящих способов встраивания ДНК. Альтернативный подход заключается в применении направляемого нуклеазой встраивания в первом цикле введения донорной ДНК для встраивания первой субъединицы с последующим введением второй субъединицы либо с использованием того же самого подхода, либо же любого другого подходящего способа. Если подход, основанный на использовании нуклеаз, используется во многих циклах встраивания, необязательно могут быть использованы разные сайт-специфические нуклеазы для запуска направляемого нуклеазой встраивания донорной ДНК в разные сайты распознавания. Способ создания библиотеки может включать:

получение молекул первой донорной ДНК, кодирующих первую субъединицу, и получение эукариотических клеток,

введение первой донорной ДНК в клетки и обеспечение присутствия в клетке сайт-специфической нуклеазы, причем указанная нуклеаза расщепляет распознаваемую последовательность в клеточной ДНК с образованием сайта встраивания, в котором донорная ДНК встраивается в клеточную ДНК, при этом встраивание осуществляется посредством эндогенных клеточных механизмов репарации ДНК, в результате чего получают первый набор рекомбинантных клеток, содержащих первую донорную ДНК, встроенную в клеточную ДНК,

культивирование первого набора рекомбинантных клеток для получения первого набора клонов, содержащих ДНК, кодирующую первую субъединицу,

введение молекул второй донорной ДНК, кодирующих вторую субъединицу, в клетки первого набора клонов, причем вторая донорная ДНК встраивается в клеточную ДНК первого набора клонов, в результате чего получают второй набор рекомбинантных клеток, содержащих первую и вторую донорную ДНК, встроенную в клеточную ДНК, и

культивирование второго набора рекомбинантных клеток для получения второго набора клонов, причем эти клоны содержат ДНК, кодирующую первую и вторую субъединицы мультимерной связывающей молекулы,

в результате чего получают библиотеку клонов эукариотических клеток, содержащих донорную ДНК, кодирующую набор мультимерных связывающих молекул.

Сайт-специфическое встраивание донорной ДНК в клеточную ДНК обеспечивает получение рекомбинантных клеток, которые могут быть культивированы для получения клонов. Таким образом, происходит размножение отдельных рекомбинантных клеток, в которые была встроена донорная ДНК, с образованием клональной популяции клеток – «клонов» - каждый клон получают из одной исходной рекомбинантной клетки. Таким образом, с помощью указанного способа можно получить целый ряд клонов, соответствующих количеству клеток, в которые была успешно встроена донорная ДНК. Группа клонов формирует библиотеку, кодирующую набор связывающих молекул (или на промежуточной стадии, когда субъединицы связывающей молекулы встраиваются в отдельных раундах, клоны могут кодировать ряд субъединиц связывающих молекул). Таким образом, способы по настоящему изобретению могут обеспечивать получение библиотеки клонов эукариотических клеток, содержащих донорную ДНК, кодирующую набор связывающих молекул.

С помощью способов по настоящему изобретению можно создать библиотеки клонов, содержащих донорную ДНК, встроенную в фиксированный локус или во многие фиксированные локусы в клеточной ДНК. Под термином «фиксированный» подразумевается, что локус является одинаковым в разных клетках. Следовательно, клетки, используемые для создания библиотеки, могут содержать распознаваемую нуклеазой последовательность в фиксированном локусе, представляющим собой универсальный сайт посадки в клеточной ДНК, в которую может встроиться донорная ДНК. Распознаваемая последовательность для сайт-специфической нуклеазы может располагаться в одном или нескольких положениях в клеточной ДНК.

Библиотеки, полученные согласно настоящему изобретению, могут быть использованы различными способами. Клетки библиотеки могут быть культивированы для экспрессии связывающих молекул, Таким образом, обеспечивая получение широкого набора связывающих молекул. Библиотека может подвергаться скринингу для выявления клеток с требуемым фенотипом, при этом фенотип обусловлен экспрессией связывающей молекулы клеткой. Представляется возможным скрининг фенотипических свойств, в котором клетки библиотеки культивируют для экспрессии связывающих молекул с последующим обнаружением проявления требуемого фенотипа в клонах библиотеки. Критерии оценки клеток могут быть основаны на изменении поведения клеток, как, например, изменение экспрессии эндогенных или экзогенных репортерных генов, дифференцировочный статус, пролиферация, жизнеспособность, размер клеток, метаболизм или изменение взаимодействий с другими клетками. В случае обнаружения требуемого фенотипа клетки клона, который проявляет требуемый фенотип, затем могут

быть выделены. Затем из выделенного клона обязательно выделяют ДНК, кодирующую связывающую молекулу, обеспечивая получение ДНК, кодирующей связывающую молекулу, с получением требуемого фенотипа при экспрессии в клетке.

Основная цель, с которой библиотеки эукариотических клеток были использованы, заключается в способах скрининга связывающих молекул, которые распознают представляющую интерес мишень. В таких способах библиотеку клеток культивируют для экспрессии связывающих молекул, а связывающие молекулы экспонируют с мишенью для обеспечения распознавания мишени одной или более когнатных связывающих молекул, в случае наличия, и обнаружение распознавания мишени когнатной связывающей молекулой. В таких способах, связывающие молекулы могут быть экспрессированы на клеточной поверхности, а те клоны библиотеки, которые экспрессируют связывающие молекулы с требуемыми свойствами, могут быть выделены. Таким образом, клетки включающие гены, кодирующие связывающие молекулы с требуемыми функциональными или связывающими свойствами, могут быть идентифицированы в составе библиотеки. Гены могут быть выделены и использованы для получения связывающей молекулы или использованы для дальнейшего конструирования с образованием производных библиотек связывающих молекул для получения связывающих молекул с улучшенными свойствами.

Настоящее изобретение предоставляет преимущества по сравнению с предыдущими подходами к созданию библиотек у высших эукариот. В некоторых исследованиях для введения генов антител в репортерные клетки млекопитающих было использовано заражение лентивирусами [106]. Этот подход обладает тем преимуществом, что могут быть созданы большие библиотеки, но при этом отсутствует контроль за сайтом встраивания, а количество копий контролируется с использованием заражения с низкой множественностью (как отмечалось выше). В альтернативном подходе гены антител были введены посредством гомологичной рекомбинации, не используя направляемое нуклеазой встраивание и используя гомологичные участки длиной 10 т.п.о., но эффективность направленного встраивания была относительно низкой, что означает, что потенциальный размер библиотеки был ограниченным [105]. В отличие от этого применение нуклеаз, направленных на определенные последовательности, сохраняет полезные свойства направленного встраивания в один или несколько предпочтительных локусов, при этом обеспечивая эффективное создание больших библиотек. Направляемое нуклеазой встраивание имеет то преимущество, что трансгены направленно встраиваются в фиксированный локус или фиксированные локусы в составе клеточной ДНК. Это означает, что активность промотора, запускающего транскрипцию генов связывающих молекул, во всех клонх будет одинаковой, а функциональные свойства каждой связывающей молекулы

будут скорее отражением свойственной ей эффективности, трансляционной эффективности и стабильности, а не следствием изменчивости, связанной с сайтом встраивания. целенаправленное встраивание в один или ограниченное количество локусов также обеспечит лучший контроль экспрессии, в случае необходимости, например, с использованием индуцибельных промоторов.

Различные признаки настоящего изобретения дополнительно описаны ниже. Следует отметить, что заглавия, использованные по всему настоящему описанию, служат только для навигации и не должны интерпретироваться как исчерпывающие, и что варианты реализации, описанные в разных разделах, в соответствующих случаях могут сочетаться.

## ПОДРОБНОЕ ОПИСАНИЕ СУЩНОСТИ ИЗОБРЕТЕНИЯ

### *Эукариотические клетки*

Потенциальные популяции эукариотических клеток, экспрессирующих широкий набор связывающих молекул, представлены и обсуждаются в Примерах в данном документе в отношении экспрессии наборов антител на поверхности клеток млекопитающих. Преимущества настоящего изобретения не ограничиваются клетками млекопитающих и включают всех эукариот.

Дрожжи (например, *Saccharomyces cerevisiae*) имеют меньший геном, чем клетки млекопитающих, а гомологичная рекомбинация, определяемая гомологичными участками (в отсутствие направляемого нуклеазой расщепления), является эффективным способом введения чужеродной ДНК по сравнению с высшими эукариотами. Таким образом, особое преимущество направляемого нуклеазой встраивания по настоящему изобретению заключается во встраивании генов связывающих молекул в клетки высших эукариот с геномом большего размера, в которых гомологичная рекомбинация в отсутствие расщепления нуклеазой является менее эффективной. Направляемое нуклеазой встраивание было использовано в дрожжевых клетках для решения проблемы эффективного встраивания множества генов в отдельные дрожжевые клетки, Например, для конструирования метаболических путей (US2012/0277120), но в этой работе не охвачено введение библиотек связывающих молекул, а также не рассмотрены проблемы создания библиотек у высших эукариот.

Библиотеки эукариотических клеток согласно настоящему изобретению предпочтительно представляют собой клетки высших эукариот, определенные в данном документе как клетки с геномом большим, чем геном *Saccharomyces cerevisiae*, размер

которого составляет  $12 \times 10^6$  пар оснований (п.о.). Размер генома клеток высших эукариот может, например, составлять более чем  $2 \times 10^7$  пар оснований. К ним относятся, например, клетки млекопитающих, птиц, насекомых или растений. Предпочтительно клетки представляют собой клетки млекопитающих, Например, мыши или человека. Клетки могут представлять собой первичную культуру клеток или могут представлять собой клеточные линии. Обычно для экспрессии антител и белков используются клетки яичника китайского хомячка (СНО), но в настоящем изобретении могут быть использованы любые альтернативные стабильные клеточные линии. В данном документе в Примерах используются клетки НЕК293. Для эффективного введения чужеродной ДНК в клетки первичной культуры доступны различные способы, позволяющие их использовать (например, путем электропорации, где показатели эффективности и жизнеспособности составляли до 95 %, <http://www.maxcyte.com/technology/primary-cells-stem-cells.php>).

Клетки линии Т-лимфоцитов (например, первичная культура Т-клеток или линия Т-клеток) или клетки линии В-лимфоцитов являются одними из предпочтительных типов клеток. Особый интерес представляют первичная культура Т-клеток или клеточные линии, полученные из Т-клеток, для применения в библиотеках TCR, включая клеточные линии, в которых отсутствует экспрессия TCR [23, 24, 25]. Примеры линии клеток В-лимфоцитов включают В-клетки, пре-В-клетки или при-В-клетки и клеточные линии, полученные из любых из них.

Создание библиотек в первичной культуре В-клеток или линиях В-клеток представляет особую ценность для создания библиотек антител. Breous-Nystrom *et al.* [15] создали библиотеки в линии мышинных пре-В-клеток (1624-5). Клеточная линия DT40, полученная из В-клеток курицы (ATCC CRL-2111), является особо перспективной для создания библиотек связывающих молекул. DT40 представляет собой линию небольших клеток с относительно высокой скоростью деления клеток. Наборы связывающих молекул могут быть целенаправленно встроены в специфические локусы с использованием нуклеаз ZFN, TALE или CRISPR/Cas9, которые целенаправленно расщепляют эндогенные последовательности, или путем целенаправленного встраивания в предварительно встроенные гетерологичные сайты, которые могут содержать сайты распознавания мегануклеазой. Клетки DT40 экспрессируют антитела и, таким образом, предпочтительным было бы целенаправленное изменение генов антител в локусе, содержащем гены, кодирующие антитела, с разрушением эндогенных переменных доменов антител курицы или в отсутствие такового. DT40 клетки также были использованы в качестве основы *in vitro* системы для создания IgM курицы, которая называется системой создания библиотек с использованием автономной диверсификации генов (Autonomously Diversifying Library

system, система ADLib), в которой используется преимущество собственной системы диверсификации локуса антитела курицы. В результате этой эндогенной диверсификации возникает возможность получения антител с новой специфичностью. Подход, основанный на использовании нуклеаз, описанный в данном документе, может быть использован в комбинации с ADLib для объединения разнообразных библиотек связывающих молекул из гетерологичных источников (например, наборы вариабельных областей антител человека или полученные синтетическим методом альтернативные каркасные белки) с возможностью дальнейшей диверсификации с использованием локуса IgG курицы. Аналогичные преимущества могут касаться линий В-клеток человека, таких как Nalm6 [26]. Другие представляющие интерес линии В-клеток включают такие линии, как, например, линия мышинных пре-В-клеток 1624-5 и линия про-В-клеток Va/F3. Va/F3 зависит от использования ИЛ-3 [27], и ее применение рассматривается в других разделах данного документа. В заключение необходимо отметить, что может быть использован целый ряд человеческих клеточных линий, включая линии, приведенные в источниках «Cancer Cell Line Encyclopaedia» [2 8] или «COSMIC catalogue of somatic mutations in cancer» [2 9].

Обычно библиотека состоит из одного типа клеток, полученных путем введения донорной ДНК в популяцию клональных эукариотических клеток, например, путем введения донорной ДНК в клетки определенной клеточной линии. Основное значительное различие между разными клонами библиотек будет тогда обусловлено встраиванием донорной ДНК.

#### *Вирусные системы, используемые в эукариотических клетках*

Преимущества использования системы при создании библиотек связывающих молекул в эукариотических клетках могут быть применимы к вирусным системам дисплея, основанным на эукариотических экспрессирующих системах, Например, бакуловирусный дисплей или ретровирусный дисплей [1, 2, 3, 4]. В данном подходе геном каждой клетки будет кодировать связывающую молекулу, способную к встраиванию в геном вирусной частицы. В случае ретровирусной системы кодирующая мРНК будет упакована, и кодируемая связывающая молекула будет представлена на клеточной поверхности. В случае бакуловирусных систем понадобится заключение генов, кодирующих связывающую молекулу, гены, кодирующие связывающую молекулу, должны быть включены в вирусный геном, чтобы обеспечить физическую связь между геном и белком, который он кодирует. Этого можно достичь с использованием клеток-хозяев, несущих эписомальные копии генома бакуловирусов. В качестве альтернативы, встроенные копии могут быть высвобождены под действием специфической нуклеазы (отличающейся от нуклеазы,



использованной для осуществления сайт-специфического встраивания). В случае мультимерных связывающих молекул часть субъединиц может кодировать клеточная ДНК, а гены одной или нескольких субъединиц могут содержаться в геноме вируса.

#### *Сайт-специфическая нуклеаза*

Настоящее изобретение включает применение сайт-специфической нуклеазы для направленного расщепления клеточной ДНК при составлении библиотеки эукариотических клеток, содержащих ДНК, кодирующую набор связывающих молекул, причем опосредованное нуклеазой расщепление ДНК улучшает сайт-специфическое встраивание генов связывающих молекул посредством эндогенных клеточных механизмов репарации ДНК. Сайт-специфическая нуклеаза расщепляет клеточную ДНК после специфического связывания с распознаваемой последовательностью, в результате чего получают сайт встраивания для донорной ДНК. Нуклеаза может создавать двухцепочечные разрывы или одноцепочечный разрыв («ник»). Клетки, используемые для создания библиотек, могут содержать эндогенные последовательности, распознаваемые сайт-специфической нуклеазой, или распознаваемая последовательность может быть встроена в клеточную ДНК.

Сайт-специфическая нуклеаза может быть чужеродной для клеток, *т.е.* не встречающейся в клетках выбранного типа в естественных условиях.

Сайт-специфическая нуклеаза может быть введена до, после или одновременно с введением донорной ДНК, кодирующей связывающую молекулу. Допускается, что донорная ДНК кодирует нуклеазу наряду со связывающей молекулой или в составе отдельной молекулы нуклеиновой кислоты, которая котрансфецируется или вводится другим образом вместе с донорной ДНК. Клоны библиотеки необязательно могут сохранять нуклеиновую кислоту, кодирующую сайт-специфическую нуклеазу, или такая нуклеиновая кислота может быть только транзитно трансфецирована в клетки.

В настоящем изобретении может быть использована любая подходящая сайт-специфическая нуклеаза. Она может представлять собой природный фермент или сконструированный вариант. Известен целый ряд особенно подходящих нуклеаз, как, например, нуклеазы, что распознают или нуклеазы, которые могут быть сконструированы с целью распознавания последовательностей, которые только изредка встречаются в клеточной ДНК. Предпочтительным является расщепление нуклеазой только в одном или двух сайтах, поскольку только одна или две молекулы донорной ДНК должны встраиваться на клетку. Низкая частота распознавания последовательности сайт-специфической нуклеазой является более вероятной, если длина распознаваемой последовательности является относительно большой. Например, последовательность, специфически

распознаваемая нуклеазой, может представлять собой последовательность с длиной по меньшей мере 10, 15, 20, 25 или 30 нуклеотидов.

Примерами подходящих нуклеаз являются мегануклеазы, нуклеазы с доменом «цинковые пальцы» (ZFN), нуклеазы TALE и нуклеазы, направляемые нуклеиновой кислотой (например, направляемые РНК), как, например, система CRISPR/Cas. Каждая из них создает двухцепочечные разрывы, хотя известны сконструированные формы, которые создают одноцепочечные разрывы.

Мегануклеазы (также известные как хоминг-эндонуклеазы) представляют собой нуклеазы, которые встречаются во всех царствах живых организмов и распознают относительно длинные последовательности (12-40 п.о.). Учитывая длинную распознаваемую последовательность, они либо отсутствуют, либо встречаются относительно редко в геномах эукариотических клеток. Мегануклеазы подразделяются на 5 семейств на основе последовательности/структуры. (LAGLIDADG, GIY-YIG, HNH, His-Cys box и PD-(D/E)XK). Наиболее изученным семейством является семейство LAGLIDADG, которое включает хорошо изученную мегануклеазу I-SceI из *Saccharomyces cerevisiae*. I-SceI распознает и расщепляет распознаваемую последовательность длиной 18 п.о. (5' TAGGGATAACAGGGTAAT), оставляя выступ на 3'-конце длиной 4 п.о. Другим широко используемым примером является I-CreI, которую получают из хлоропластов одноклеточной зеленой водоросли *Chlamydomonas reinhardtii*, и которая распознает последовательность длиной 22 п.о. [30]. Был сконструирован целый ряд вариантов с измененными распознаваемыми последовательностями [31]. Мегануклеазы являются первым примером применения сайт-специфических нуклеаз в геномной инженерии [49, 50]. Как и в случае подходов, основанных на использовании рекомбиназ, применение I-SceI и других мегануклеаз требует предварительного встраивания соответствующего сайта распознавания, для направленного распознавания в геноме или конструирования мегануклеаз для распознавания эндогенных сайтов [30]. С помощью данного подхода эффективное целенаправленное встраивание в клетках HEK293 (судя по основанной на гомологичной рекомбинации «репарации» встроенного дефектного гена GFP) было достигнута у 10-20 % клеток за счет применения I-SceI [32].

Предпочтительным классом мегануклеаз для применения в настоящем изобретении являются эндонуклеазы LAGLIDADG. К ним относятся I-Sce I, I-Chu I, I-Cre I, Csm I, Pl-Sce I, Pl-Tli I, Pl-Mtu I, I-Ceu I, I-Sce II, I-Sce III, HO, Pi-Civ I, Pl-Ctr I, Pl-Aae I, Pl-Bsu I, Pl-Dha I, Pl-Dra I, Pl-Mav I, Pl-Mch I, Pl-Mfu Pl-Mfl I, Pl-Mga I, Pl-Mgo I, Pl-Min I, Pl-Mka I, Pl-Mle I, Pl-Mma I, Pl-Msh I, Pl-Msm I, Pl-Mth I, Pl-Mtu Pl-Mxe I, Pl-Npu I, Pl-Pfu I, Pl-Rma I, Pl-Spb I, Pl-Ssp I, Pl-Fac I, Pl-Mja I, Pl-Pho I, Pi-Tag I, Pl-Thy I, Pl-Tko I, I-Mso I, и Pl-Tsp I;

предпочтительно I-Sce I, I-Cre I, I-Chu I, I-Dmo I, I-Csm I, Pl-Sce I, Pl-Pfu I, Pl-Tli I, Pl-Mtu I и I-Ceu I.

В последние годы был разработан целый ряд способов, с помощью которых можно сконструировать новые нуклеазы, специфические к определенным последовательностям, путем слияния ДНК-связывающих доменов, специфических к определенным последовательностям, с неспецифическими нуклеазами для создания сконструированных нуклеаз, специфических к определенным последовательностям, направляемых ДНК-связывающими доменами с заданными свойствами. Специфичность связывания может определяться сконструированными связывающими доменами, как, например, доменами типа «цинковые пальцы». Они представляют собой небольшие модульные домены, стабилизированные ионами цинка, которые вовлечены в молекулярное распознавание и используются в природе для распознавания последовательностей ДНК. Было сконструировано множество доменов типа «цинковые пальцы» для обеспечения специфического связывания с определенной последовательностью, и они были связаны с неспецифическим расщепляющим ДНК доменом рестриктазы FokI типа II с образованием нуклеаз с доменом «цинковые пальцы» (ZFN). ZFN могут быть использованы для создания двухцепочечного разрыва в определенных сайтах генома. FokI представляет собой облигатный димер и требует наличия двух ZFN для связывания в непосредственной близости для осуществления расщепления. Специфичность сконструированных нуклеаз была увеличена, а их токсичность была уменьшена за счет создания двух разных вариантов FokI с целью образования гетеродимеров друг с другом [33]. Было показано встраивание путем гомологичной рекомбинации таких облигатных гетеродимерных ZFN в 5-18 % клеток-мишеней в отсутствие необходимости в отборе по чувствительности к лекарственному препарату [21, 34, 35]. Было продемонстрировано включение вставок до 8 т. п.о. с частотой >5 % в отсутствие отбора.

Недавно было показано, что одноцепочечные выступы на 5'-конце, создаваемые нуклеазами, такими как ZFN, способствуют эффективному встраиванию трансгенов в сайты расщепления [45]. Эти исследования были расширены, чтобы показать, что расщепление донорной ДНК *in vivo* (посредством включения специфического сайта распознавания нуклеазой в донорную плазмиду) повышает эффективность негомологичного встраивания. Механизм не совсем хорошо понятен, но не исключено, что уменьшение воздействия клеточных нуклеаз за счет линеаризации *in vivo* могло способствовать наблюдаемому повышению эффективности [45]. Также возможно, что совпадения выступов на 5'-конце донорной и акцепторной ДНК, создаваемых нуклеазами, запускает лигирование. Однако анализ последовательностей в местах соединений показал

наличие делеций. Возможно, что точное совпадение соединений продолжает действовать в качестве субстрата для сайт-направляемых нуклеаз до того, пока не происходит делеция распознаваемой последовательности. Для преодоления этой потенциальной проблемы Magesca *et al.* [36] поменяли сайты распознавания левой и правой ZFN местами в донорной ДНК, так что лигирование донорной ДНК в геномный локус приводило к дупликации двух левосторонних ZFN на одном фланге встраивания и дупликации двух правосторонних ZFN на другом фланге. Применение облигатной гетеродимерной нуклеазы (как описано для FokI) означает, что ни одна из этих новосозданных фланкирующих последовательностей не может расщепляться целевой нуклеазой.

Возможность конструирования ДНК-связывающих доменов определенной специфичности была дополнительно упрощена вследствие открытия в бактериях *Xanthomonas* молекул, называемых эффекторы, подобные активаторам транскрипции (TALE). Эти молекулы TALE состоят из наборов мономеров из 33-35 аминокислот, при этом каждый мономер распознает одно основание в составе последовательности-мишени [37]. Эта модульное соотношение 1:1, относительно облегчая возможность конструирования молекул TALE для связывания любой представляющей интерес ДНК-мишени. В результате присоединения этих сконструированных TALE к FokI стало возможным создание новых TALE-нуклеаз, специфических к определенной последовательности. В настоящее время были сконструированы нуклеазы TALE, также известные как TALEN, специфические к большому количеству сайтов и характеризующиеся высоким уровнем успеха в отношении эффективной геном-модифицирующей активности [38]. В Примерах в данном документе авторы настоящего изобретения продемонстрировали улучшенное встраивание донорной ДНК вследствие применения нуклеаз TALE. Другие варианты и усовершенствованные подходы технологии на основе нуклеаз TALE были разработаны и могут быть использованы для создания библиотек связывающих молекул посредством направляемого нуклеазой встраивания. К ним относятся «мега-TALEN», когда связывающий домен нуклеазы TALE, слит с мегануклеазой [39], и «компактные TALEN», когда для осуществления расщепления используется один распознающий домен нуклеазы TALE [40].

На протяжении последних лет была описана другая система для осуществления двух- или одноцепочечных разрывов в специфических последовательностях в геноме. Эта система, называемая «короткие палиндромные повторы, регулярно расположенные группами (CRISPR) и система, ассоциированная с CRISPR (Cas)», основана на защитном механизме бактерий [41]. Система CRISPR/Cas целенаправленно воздействует на ДНК для расщепления с помощью короткой, комплементарной одноцепочечной РНК (CRISPR RNA

или crRNA), присоединенной к короткому палиндромному повтору. В широко используемой системе «тип II», процессинг РНК-мишени зависит от наличия трансдействующей crRNA (tracrRNA), которая имеет последовательность, комплементарную палиндромному повтору. Гибридизация tracrRNA с последовательностью палиндромного повтора запускает процессинг. Процессированная РНК активирует домен Cas9 и направляет ее действие на комплементарную последовательность в остове ДНК. Система была упрощена для осуществления расщепления с помощью Cas9 одного РНК-транскрипта и была направлена на многие разные последовательности в составе генома [42, 43]. Данный подход к расщеплению геномной ДНК обладает преимуществом, заключающемся в направлении с помощью короткой последовательности РНК, делая относительно простым конструирование молекул со специфичностью расщепления. Таким образом, существует целый ряд разных способов для достижения сайт-специфического расщепления геномной ДНК. Как описано выше, эта повышает показатель встраивания донорной плазмиды посредством эндогенных клеточных механизмов репарации ДНК.

Применение мегануклеаз, ZFN, нуклеазы TALE или систем, направляемых нуклеиновыми кислотами, как, например, системы CRISPR/Cas9, обеспечит возможность направленного воздействия на эндогенные локусы в геноме. В примерах в данном документе авторы настоящего изобретения продемонстрировали целенаправленное встраивание в локус AAVS, но целенаправленному воздействию также могут подвергаться альтернативные локусы. Например, локус гена коллагена типа I был использован для эффективной экспрессии трансгенов [44].

В качестве альтернативы гетерологичные сайты распознавания для целенаправленно воздействующих нуклеаз, включая мегануклеазы, нуклеазы ZFN и TALE, могут быть введены заранее для последующего целенаправленного получения библиотеки. В примерах в данном документе авторы настоящего изобретения описывают применение нуклеазы TALE, распознающей последовательность в составе локуса AAVS для введения путем гомологичной рекомбинации, последовательности, распознаваемой мегануклеазой I-SceI, и гетерологичных сайтов распознавания нуклеазой TALE в составе локуса AAVS. Осуществляемое нуклеазой целенаправленное встраивание может быть использовано для осуществления вставки последовательностей-мишеней путем гомологичной рекомбинации или NHEJ с использованием векторной ДНК или даже двухцепочечных олигонуклеотидов [45]. В качестве альтернативы, способы неспецифического направленного встраивания могут быть использованы для введения сайтов-мишеней посредством применения встраивания, направляемого транспозонами [46] для введения сайтов распознавания для сайт-специфических нуклеаз. Системы на основе вирусов, таких как лентивирус,

используемые при низком титре также могут быть использованы для введения сайтов-мишеней. Трансфекция ДНК в сочетании со скринингом в отношении однокопийной вставки также может быть использована для идентификации уникальных сайтов встраивания [17]. Такие неспецифические подходы будут особо полезными в случае клеток, в которых явно отсутствует сайт для направленного воздействия или в случае геномов, которые не были секвенированы, или в случае геномов, для которых недоступны существующие нуклеазы TALE, ZFN или системы Cas9/CRISPR. После установления клеточной линии вследствие случайной вставки сайта распознавания нуклеазой, впоследствии клеточная линия может быть использована для создания библиотек связывающих молекул, в которых все клоны библиотеки содержат трансген в фиксированном локусе, что обусловлено использованием направляемого нуклеазой встраивания.

В представленных примерах используются три разные плазмиды, включающие в себя пары нуклеаз TALE или ZFN на отдельных плаزمидках с отдельной плазмидой для донорной ДНК. В случае мегануклеазы сайт-специфическая нуклеаза кодируется одним геном, и он вводится в составе одной плазмиды, при этом донорная ДНК содержится в составе второй плазмиды. Разумеется, могут быть использованы комбинации, включающие два или более этих элементов в составе одной и той же плазмиды, и это может повысить эффективность направленного встраивания за счет уменьшения количества плазмид, которые подлежат введению. В дополнение к этому, представляется возможным предварительное встраивание нуклеазы(нуклеаз), которая(которые) также может(могут) быть индуцибельной(индуцибельными) для обеспечения временного контроля активности нуклеазы, что было продемонстрировано для транспозаз [46]. В конечном итоге, нуклеаза может быть введена в виде рекомбинантного белка или в виде комплекса белка и РНК (например, в случае РНК-направляемой нуклеазы, такой как CRISPR:Cas9).

#### *Локус*

Последовательность, распознаваемая сайт-специфической нуклеазой, может содержаться в геномной ДНК или эписомальной ДНК, которая стабильно наследуется клетками. Следовательно, донорная ДНК может быть встроена в геномный или эписомальный локус клеточной ДНК.

В самом простом виде один ген, кодирующий связывающую молекулу (ген связывающей молекулы), целенаправленно встраивается в один сайт в составе генома эукариотических клеток. Идентификация клетки, характеризующейся определенной активностью связывания или клеточным фенотипом, обеспечивает возможность прямого выделения гена, кодирующего молекулу с требуемым свойством (например, путем ПЦР

мРНК или геномной ДНК). Этому способствует использование уникальной распознаваемой последовательности для сайт-специфической нуклеазы, встречающейся один раз в клеточной ДНК. Таким образом, клетки, используемые для создания библиотеки, могут содержать распознаваемую нуклеазой последовательность в одном фиксированном локусе, т.е. один идентичный локус во всех клетках. Библиотека, полученная из таких клеток, будет содержать донорную ДНК, встроенную в фиксированном локусе, т.е. встречающуюся в том же самом локусе в клеточной ДНК всех клонов в библиотеке.

Необязательно распознаваемые последовательности могут встречаться много раз в клеточной ДНК, вследствие чего клетки имеют несколько потенциальных сайтов встраивания для донорной ДНК. Эта ситуация типична для диплоидных или полиплоидных клеток, когда распознаваемая последовательность содержится в соответствующих положениях в паре хромосом, т.е. дублированных локусах. В библиотеке, полученной из таких клеток, может содержаться донорная ДНК, встроенная в дублированные фиксированные локусы. Например, в библиотеке, полученной из диплоидных клеток, может содержаться донорная ДНК, встроенная в дубликатные фиксированные локусы, а в библиотеке, полученной из триплоидных клеток, может содержаться донорная ДНК, встроенная в трипликатные фиксированные локусы. Многие подходящие клетки млекопитающих являются диплоидными, а клоны библиотек клеток млекопитающих согласно настоящему изобретению могут содержать донорную ДНК, встроенную в дубликатные фиксированные локусы.

Последовательность, распознаваемая сайт-специфической нуклеазой, может встречаться в нескольких независимых локусах в клеточной ДНК. Следовательно, донорная ДНК может встраиваться во многие независимые локусы. Библиотека диплоидных или полиплоидных клеток может включать донорную ДНК, встроенную во многие независимые фиксированные локусы, и/или в дублированные фиксированные локусы.

В клетках, содержащих распознаваемые последовательности во многих локусах, (либо в дублированных, либо независимых локусах), каждый локус представляет собой потенциальный сайт встраивания для молекулы донорной ДНК. Введение донорной ДНК в клетки может обеспечивать встраивание во все распознаваемые нуклеазой последовательности, присутствующие в клетке, или донорная ДНК может встраиваться в некоторые, а не все эти потенциальные сайты. Например, при получении библиотеки из диплоидных клеток, содержащих распознаваемые последовательности в первом и втором фиксированных локусах (например, дубликатные фиксированные локусы), полученная в результате библиотека может включать клоны, в которых донорная ДНК встроена в первый фиксированный локус, клоны, в которых донорная ДНК встроена во второй

фиксированный локус, и клоны, в которых донорная ДНК встроена как в первый, так и второй фиксированные локусы.

Следовательно, способы получения библиотек могут включать расщепление многих фиксированных локусов в клетке сайт-специфической нуклеазой и встраивание донорной ДНК во многих фиксированных локусах. Как отмечалось выше, в тех случаях, когда имеется много копий одной и той же распознаваемой последовательности (например, что происходит при направленном воздействии на эндогенные локусы в диплоидных или полиплоидных клетках), существует вероятность встраивания двух генов связывающих молекул, особенно при использовании эффективных механизмов направленного воздействия, причем только один ген является специфическим в отношении мишени. Это может быть решено на протяжении последующего скрининга, после выделения генов связывающих молекул.

В некоторых случаях возможно потребуется введения более одной связывающей молекулы на клетку. Например, вследствие объединения двух разных антител могут быть получены биспецифические связывающие молекулы, и они могут обладать свойствами, которые отсутствуют у отдельных связывающих молекул [47]. Этого можно достичь путем введения разных генов антител в оба аллеля в дубликатных фиксированных локусах или путем целенаправленного введения генов разных популяций антител в независимые фиксированные локусы с использованием способов, описанных в данном документе. Кроме того, связывающая молекула, как таковая, может состоять из многих цепей (например, VH- и VL-домены антител, представленные в формате Fab или IgG). В данном случае возможно потребуется встраивание разных субъединиц в разные локусы. Они могут быть встроены в пределах одного и того же цикла направляемого нуклеазой встраивания, они могут быть встроены последовательно с использованием направляемого нуклеазой встраивания на протяжении одного или обоих этапов встраивания.

#### *Введение донорной ДНК*

Были описаны многочисленные способы для введения донорной ДНК в эукариотические клетки, включая трансфекцию, заражение или электропорацию. Осуществление трансфекции больших количеств клеток возможно за счет использования стандартных способов, включая полиэтиленимин-опосредованную трансфекцию, как описано в данном документе. В дополнение к этому, доступны способы для высокоэффективной электропорации  $10^{10}$  клетки за 5 минут, например, <http://www.maxcyte.com>.

Могут быть созданы комбинаторные библиотеки, в которых компоненты пар мультимерных связывающих молекул (например, гены VH и VL антител) или даже разные



части одной и той же связывающей молекулы вводятся в составе разных плазмид. Введение отдельных молекул донорной ДНК, кодирующих отдельные связывающие молекулы или субъединицы связывающих молекул, может осуществляться одновременно или последовательно. Например, легкая цепь антитела может быть введена путем трансфекции или заражения, клетки могут быть выращены и отобраны в случае необходимости. Затем в последующем этапе заражения или трансфекции могут быть введены другие компоненты. Один или оба этапа могут включать направляемое нуклеазой встраивание в специфические геномные локусы.

#### *Встраивание донорной ДНК*

Донорная ДНК встраивается в клеточную ДНК, образуя рекомбинантную ДНК, имеющую непрерывную последовательность ДНК, в которую встраивается донорная ДНК в сайте встраивания. В настоящем изобретении встраивание опосредовано природными механизмами репарации ДНК, которые относятся к собственным механизмам клетки. Таким образом, возможность встраивания может обеспечиваться просто за счет введения донорной ДНК в клетку, что обеспечивает формирование сайта встраивания-специфической нуклеазой и встраивание донорной ДНК. Клетки могут поддерживаться в культуре в течение периода времени, достаточного для интеграции ДНК. Это обычно приводит к образованию смешанной популяции клеток, включающей (i) рекомбинантные клетки, в которых донорная ДНК была встроена в сайте встраивания, образующегося под воздействием сайт-специфической нуклеазы, и необязательно (ii) клетки, в которых донорная ДНК была встроена в сайтах, отличных от требуемого сайта встраивания, и/или необязательно (iii) клетки, в которые донорная ДНК не была встроена. Требуемые рекомбинантные клетки и полученные в результате клоны библиотеки, таким образом, могут быть получены в смешанной популяции с другими эукариотическими клетками. Способы отбора, описанные в других разделах в данном документе, могут быть использованы для обогащения клеток библиотеки.

Эндогенные механизмы репарации ДНК, характерные для эукариотических клеток, включают гомологичную рекомбинацию, негомологичное соединение концов (NHEJ) и соединение концов с помощью микромологичной рекомбинации. Эффективность модификации ДНК посредством таких процессов может быть увеличена путем введения двухцепочечных разрывов (DSB) в ДНК, и, как сообщалось, показатели эффективности увеличивались в 40000 раз при использовании редкощепящей эндонуклеазы (мегануклеазы), такой как I-SceI [48, 49, 50].

В отличие от сайт-специфической рекомбинации, используемой в системах, таких как Flp-In [16], для реализации настоящего изобретения не требуется наличие экзогенных

рекомбиназ или встраивание сайтов распознавания рекомбиназой. Следовательно, в настоящее изобретение необязательно не включен этап опосредованного рекомбиназами встраивания ДНК в процесс создания библиотеки, и/или необязательно в эукариотических клетках, в которые вводится донорная ДНК, отсутствует сайт рекомбинации для сайт-специфической рекомбиназы. Механизмы и практические аспекты направляемого встраивания донорной ДНК в клеточную ДНК с помощью рекомбиназ и нуклеаз очень различаются. В обсуждении в источнике Jasin 1996 [50]: «... реакция, катализируемая сайт-специфическими рекомбиназами, достаточно отличается от клеточной репарации DSB. Сайт-специфические рекомбиназы, таких как Cre, объединяют два сайта распознавания и создают одноцепочечные разрывы в этих сайтах, тем самым приводя к образованию промежуточных соединений Холлидея. Промежуточные соединения распадаются с образованием делеций, инверсий и вставок (клоны с несколькими интегрированными последовательностями), все из которых восстанавливают два сайта распознавания. Реакция является абсолютно точной и поэтому обратимой. Разрывы никогда не подвергаются воздействию системы клеточной репарации».

В отличие от этого, действие сайт-специфической нуклеазы заключается в образовании разрывов или «ников» в клеточной ДНК (например, геномной или эписомальной), которые подвергаются воздействию эндогенных клеточных механизмов репарации, таких как гомологичная рекомбинация или NHEJ. Одним из абсолютных требований к подходам, основанным на использовании рекомбиназ, является предварительное встраивание их сайтов распознавания, таким образом, такие способы требуют сайта встраивания «горячая точка» в клеточную ДНК в качестве предварительного этапа. С использованием направляемого нуклеазой встраивания становится возможным конструировать нуклеазы или направлять с помощью направляющей РНК в случае CRISPR:Cas9 для распознавания эндогенных локусов, т.е. природных последовательностей нуклеиновой кислоты, встречающихся в клеточной ДНК. В конечном итоге, с практической точки зрения подходы, основанные на использовании нуклеаз, являются более эффективными для прямого встраивания трансгенов на уровнях, требуемых для создания больших библиотек связывающих молекул.

Механизм репарации ДНК, с помощью которого донорная ДНК встраивается согласно способам настоящего изобретения, может быть определен заранее или изменен до некоторой степени за счет дизайна донорной ДНК и/или выбора сайт-специфической нуклеазы.

Гомологичная рекомбинация является природным механизмом, используемым клетками для репарации двухцепочечных разрывов с использованием гомологичной

последовательности (например, из другого аллель) в качестве матрицы для репарации. Гомологичную рекомбинацию использовали в клеточной инженерии для введения вставок (включая трансгены), делеций и точечных мутаций в геном. Гомологичной рекомбинации способствует обеспечение наличия гомологичных участков на донорной ДНК. В первоначальном подходе к конструированию клеток высших эукариот обычно использовались гомологичные участки длиной 5-10 т.п.о. в составе донорной плазмиды для увеличения эффективности направленного встраивания в представляющий интерес сайт. Вопреки этому, гомологичная рекомбинация, определяемая исключительно длинными гомологичными участками, является менее эффективной, чем рекомбинация, направляемая Flp и Cre, особенно у высших эукариот с большими геномами. Гомологичная рекомбинация особенно подходит для эукариот, таких как дрожжи, размер генома которых составляет всего лишь  $12,5 \times 10^6$  п.о., у которых она является более эффективной по сравнению с высшими эукариотами с геномами большего размера, например, клетками млекопитающих с  $3000 \times 10^6$  п.о.

Гомологичная рекомбинация также может осуществляться за счет наличия [52] одноцепочечных разрывов в геномной ДНК, а также выступать в качестве способа направляемого нуклеазой встраивания в геномную ДНК. Два различных пути, как было показано, способствуют гомологичной рекомбинации в местах одноцепочечных разрывов ДНК. Один, по сути, аналогичен репарации в местах двухцепочечных разрывов, когда используются Rad51/Brcs2, тогда как другой подавляется Rad51/Brcs2 и в нем предпочтительно используется одноцепочечная ДНК или донорная двухцепочечная ДНК с одноцепочечными разрывами [51].

Негомологичное соединение концов (NHEJ) является альтернативным механизмом репарации двухцепочечных разрывов в геноме, при котором концы ДНК непосредственно подвергаются повторному лигированию в отсутствие необходимости в гомологичной матрице. Расщепление геномной ДНК с использованием сайт-направленных нуклеаз также позволяет улучшить встраивание трансгенов за счет механизмов негомологичной рекомбинации. Данный подход к репарации ДНК является менее точным и может приводить к вставкам или делециям. Тем не менее, NHEJ является простым способом встраивания находящихся в рамке считывания экзонов в интрон или обеспечивает встраивание промотора и генных кассет в геном. Применение способов негомологичной рекомбинации дает возможность применения донорных векторов, в которых отсутствуют гомологичные участки, тем самым упрощая создание донорной ДНК.

Отмечалось, что короткие области с гомологичными концами используют для повторного соединения концов ДНК, и была выдвинута гипотеза, что микрогомологичные

участки длиной 4 п.о. могут быть использованы для осуществления репарации ДНК в местах двухцепочечных разрывов, называемой соединением концов с помощью микрогомологичной рекомбинации [50].

#### *Донорная ДНК*

Донорная ДНК, как правило, представляет собой кольцевую ДНК и может быть представлена в виде плазмиды или вектора. Другой альтернативой является линейная ДНК. Молекулы донорной ДНК могут включать области, которые не встраиваются в клеточную ДНК, помимо одной или более последовательностей донорной ДНК, которые встраиваются в клеточную ДНК. Обычно используется двухцепочечная ДНК, хотя в некоторых случаях может быть использована одноцепочечная ДНК. Донорная ДНК содержит один или более трансгенов, кодирующих связывающую молекулу, например, она может содержать промотор и генную кассету.

Для запуска гомологичной рекомбинации может быть использована самая простая по форме двухцепочечная, кольцевая плазида ДНК. Это требует наличия областей ДНК, фланкирующих трансгены, которые гомологичны последовательности ДНК, фланкирующей сайт расщепления в геномной ДНК. Линеаризованная двухцепочечная плазида ДНК или ПЦР-продукт, или синтетические гены могут быть использованы для запуска как гомологичной рекомбинации, так и путей репарации NHEJ. В качестве альтернативы двухцепочечной ДНК существует возможность применения одноцепочечной ДНК для запуска гомологичной рекомбинации [52]. Стандартный подход к созданию одноцепочечной ДНК заключается во введении одноцепочечной точки начала репликации из нитевидного бактериофага в плазмиду.

Вирусы, содержащие одноцепочечную ДНК, такие как аденоассоциированный вирус (AAV), были использованы для запуска эффективной гомологичной рекомбинации, при этом было показано, что эффективность повышалась на несколько порядков величин [53, 54]. Системы, как, например, на основе AAV, могут быть использованы в сочетании с расщеплением с использованием сайт-направленных нуклеаз для создания больших библиотек связывающих молекул. Преимущества обеих систем могут применяться для целенаправленного создания библиотек связывающих молекул. Лимит упаковки векторов на основе AAV составляет 4.7 т.п.о., но применение расщепления нуклеазой геномной ДНК-мишени позволит уменьшить данный лимит, обеспечивая встраивание больших по размеру трансгенных конструкций.

Молекула донорной ДНК может кодировать одну связывающую молекулу или много связывающих молекул. Необязательно одна молекула донорной ДНК может

кодировать много субъединиц связывающей молекулы. В некоторых вариантах реализации изобретения донорная ДНК кодирует субъединицу мультимерной связывающей молекулы.

#### *Промоторы и генетические элементы для отбора*

Транскрипция донорной ДНК, кодирующей связывающую молекулу, обычно осуществляется путем помещения последовательности, кодирующей связывающую молекулу, под контроль промотора и необязательно один или более энхансерных элементов для транскрипции. Промотор (и необязательно другие элементы контроля экспрессии генов) могут быть включены в состав молекулы донорной ДНК как таковой. В качестве альтернативы, в последовательности, кодирующей связывающую молекулу, может отсутствовать промотор на донорной ДНК, и вместо этого она может быть функционально связана с промотором на клеточной ДНК, например, эндогенным промотором или предварительно встроенным экзогенным промотором, в результате его вставки в сайт встраивания, создаваемый сайт-специфической нуклеазой.

Донорная ДНК может дополнительно содержать один или более дополнительных кодирующих последовательностей, как, например, генетических элементов, обеспечивающих отбор клеток, содержащих или экспрессирующих донорную ДНК. Как и в случае с последовательностью, кодирующей связывающую молекулу, что обсуждалось выше, такие элементы могут быть связаны с промотором на донорной ДНК или могут быть помещены под контроль промотора в результате встраивания донорной ДНК в фиксированный локус. Последнее расположение является удобным инструментом для отбора именно тех клеток, которые содержат встроенную донорную ДНК в требуемом сайте, поскольку эти клетки должны экспрессировать генетический элемент для отбора. В качестве такового может быть использован, например, ген, обеспечивающий устойчивость к средству негативного отбора, такому как бластицидин или пурамицин. Могут быть использованы один или более этапов отбора для удаления нежелательных клеток, как, например, клеток, в которых отсутствует донорная ДНК или в которых не произошло встраивание донорной ДНК в правильное положение. В качестве альтернативы эти клетки могут быть смешаны с клонами библиотеки.

Экспрессия заякоренной в мембране связывающей молекулы как таковая может быть использована в качестве селективного маркера. Например, если производится введение библиотеки генов антител, получивших формат слитых белков IgG или scFv-Fc, то клетки, которые экспрессируют антитело, могут быть отобраны с использованием вторичных реагентов, которые распознают экспрессированный на поверхности Fc, с использованием способов, описанных в данном документе. При первоначальной трансфекции донорной ДНК, кодирующей трансен под контролем экзогенного промотора,

происходит временная экспрессия (и экспрессия на клеточной поверхности) связывающей молекулы, и необходимо подождать до снижения уровня временной экспрессии (для достижения направленного встраивания, например, 1-2 генов антител/клетку).

В качестве альтернативы конструкция, кодирующая закрепленный в мембране элемент (например, Fc-домен согласно настоящему примеру, слитый с трансмембранным доменом рецептора PDGF), может быть предварительно встроена до введения библиотеки связывающих молекул. Если в этом закрепленном в мембране элементе отсутствует промотор или он кодируется в составе экзона, который находится вне рамки считывания предшествующего экзона, то будет нарушаться экспрессия на поверхности. Затем с помощью направленного встраивания вносимой донорной молекулы можно устранить данный дефект (например, путем целенаправленного помещения промотора или находящегося «в рамке считывания» экзона в интрон, который находится выше дефектного элемента, отвечающего за закрепление в мембране). Если находящийся в рамке считывания «корректирующий экзон», также кодирует связывающую молекулу, то произойдет слияние генов связывающей молекулы и элемента, отвечающего за закрепление в мембране, что приведет к экспрессии их обоих на поверхности. Таким образом, точно направляемое встраивание приведет к экспрессии в рамке считывания элемента, отвечающего за закрепление в мембране, отдельно или в составе слитого белка, содержащего экзогенную связывающую молекулу. Кроме того, если в библиотеке вносимых связывающих молекул отсутствует закрепленный в мембране элемент, и он неправильно встроен, отбор осуществляться не будет. Таким образом, экспрессия связывающей молекулы как таковой на клеточной поверхности может быть использована для отбора популяции клеток, характеризующихся точно направляемым встраиванием.

#### *Количество клонов и разнообразие библиотеки*

Библиотеки, полученные с помощью дрожжевого дисплея, размером  $10^7$ - $10^{10}$  клонов были ранее созданы, и было показано, что они позволяют получить связывающие молекулы в отсутствие иммунизации или предварительного отбора популяции [9, 55, 56, 57]. Во многих из библиотек, полученных с помощью дисплея на клетках млекопитающих, о которых ранее сообщалось, были использованы гены антител, полученные у иммунизированных доноров, или даже обогащенные антиген-специфические В-лимфоциты, учитывая ограничения размера библиотеки и вариабельность при использовании клеток высших эукариот. Благодаря эффективности направленного воздействия на ген, описанного в настоящем изобретении, могут быть созданы большие, библиотеки наивных антител с использованием клеток высших эукариот, как, например,

клеток млекопитающих, которые соответствуют библиотекам, описанным для низших эукариот, таких как дрожжи.

После встраивания донорной ДНК в клеточную ДНК, полученные в результате рекомбинантные клетки культивируют для размножения, получают клон клеток из каждой первоначально полученной рекомбинантной клетки. Каждый клон, таким образом, получают из одной первоначальной клетки, в которую была встроена донорная ДНК, в сайте встраивания, созданном сайт-специфической нуклеазой. Способы согласно настоящему изобретению отличаются высокой эффективностью и высокой точностью встраивания донорной ДНК, а библиотека согласно настоящему изобретению может содержать по меньшей мере  $100$ ,  $10^3$ ,  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ ,  $10^8$ ,  $10^9$  или  $10^{10}$  клонов.

С использованием направляемого нуклеазой встраивания существует возможность нацеленного воздействия на 10 % или больше трансфицированных клеток млекопитающих. С практической точки зрения целесообразным является выращивание и трансформация  $>10^{10}$  клеток (например, от 5 литров клеток, выращиваемых в концентрации  $2 \times 10^6$  клеток/мл). Трансфекция такого большого количества клеток может быть осуществлена с использованием стандартных способов, включая трансфекцию, опосредованную полиэтиленгликолем, как описано в данном документе. В дополнение к этому, доступны способы для высокоэффективной электропорации  $10^{10}$  клеток за 5 минут, например, <http://www.maxcyte.com>. Таким образом, с использованием подхода по настоящему изобретению представляется возможным создание библиотек свыше  $10^9$  клонов.

Если популяция молекул донорной ДНК, которые используются для образования библиотеки, содержит множество копий одной и той же последовательности, могут быть получены два или более клонов, которые содержат ДНК, кодирующую ту же самую связывающую молекулу. Также может возникнуть ситуация, когда клон может содержать донорную ДНК, кодирующую несколько разных связывающих молекул, например, если имеется одна распознаваемая последовательность для сайт-специфической нуклеазы, как подробно изложено в других разделах в данном документе. Таким образом, количество разных кодируемых или экспрессируемых связывающих молекул библиотеки может отличаться от количества полученных клонов.

Клоны в библиотеке предпочтительно содержат донорную ДНК, кодирующую одну или две молекулы из набора связывающих молекул, и/или предпочтительно экспрессируют только одну или две молекулы из набора связывающих молекул. Наличие ограниченного количества разных связывающих молекул на клетку является преимуществом, что касается идентификации клона и/или ДНК, кодирующей определенную связывающую молекулу, идентифицированную при скрининге библиотеки в отношении конкретной мишени.

Наиболее простая ситуация наблюдается тогда, когда клоны кодируют одну молекулу из набора связывающих молекул. Однако также не вызывает затруднений идентификация релевантной, кодирующей ДНК для требуемой связывающей молекулы, если клон, отобранный из библиотеки клонов, содержит фрагменты, кодирующие небольшое количество разных связывающих молекул, например, клон может кодировать две молекулы из набора связывающих молекул. Как обсуждалось в других разделах в данном документе, клоны, кодирующие одну или две связывающие молекулы, удобно получать путем отбора последовательности, распознаваемой сайт-специфической нуклеазой, которая встречается один раз на копию хромосом в диплоидном геноме, поскольку диплоидные клетки содержат дубликатные фиксированные локусы, один на каждую копию хромосом, а донорная ДНК может встраиваться в один или оба фиксированные локусы. Таким образом, каждый из клонов библиотеки может экспрессировать только одну или две молекулы из набора связывающих молекул.

Связывающие молекулы, представленные на поверхности клеток библиотеки, могут быть идентичными (имеют одинаковую аминокислотную последовательность) другим связывающим молекулам, представленным на той же клетке. Библиотека может состоять из клонов клеток, каждая из которых несет одну молекулу из набора связывающих молекул, или клонов, несущих много молекул из набора связывающих молекул на клетку. В качестве альтернативы библиотека может включать некоторые клоны, которые несут одну молекулу из набора связывающих молекул, и некоторые клоны, которые несут много молекул (например, две) набора связывающих молекул. Предпочтительно клоны библиотеки экспрессируют одну или две молекулы из набора связывающих молекул.

В соответствии с этим, библиотека согласно настоящему изобретению может включать клоны, кодирующие более одного представителя набора связывающих молекул, при этом донорная ДНК встраивается в дубликатные фиксированные локусы или множество независимых фиксированных локусов.

Как отмечалось выше, соответствующую кодирующую ДНК для связывающей молекулы легче всего идентифицировать, если соответствующий клон экспрессирует только одну связывающую молекулу. Обычно молекула донорной ДНК кодирует одну связывающую молекулу. Связывающая молекула может представлять собой мультимерную молекулу, поэтому молекула донорной ДНК будет содержать множество генов или открытых рамок считывания, соответствующих различным субъединицам мультимерной связывающей молекулы.

Библиотека согласно настоящему изобретению может кодировать по меньшей мере 100,  $10^3$ ,  $10^4$ ,  $10^5$  или  $10^6$ ,  $10^7$ ,  $10^8$ ,  $10^9$ , или  $10^{10}$  разных связывающих молекул. В случае



мультимерных связывающих молекулы отличия могут заключаться в одной или более субъединицах связывающей молекулы. Одна или более переменных субъединиц могут объединяться с одной или более константными субъединицами с образованием мультимерных связывающих молекул, если константные субъединицы являются одинаковыми (или не отличаются в значительной степени) во всех клонах библиотеки. При создании библиотек мультимерных связывающих молекул возможно возникновение комбинаторных отличий, при этом возможно попарное соединение первого набора субъединиц связывающей молекулы со вторым набором субъединиц связывающих молекул.

#### *Связывающие молекулы*

«Связывающая молекула» в соответствии с настоящим изобретением – связывающая молекула, являющаяся специфическим партнером по связыванию для другой молекулы. К типичным примерам специфических партнеров по связыванию относятся антитело-антиген и рецептор-лиганд.

Набор связывающих молекул, кодируемых библиотекой, обычно имеет общую структуру и один или более участков, различающихся по последовательностям. Следовательно, библиотека позволяет осуществлять отбор представителя требуемого структурного класса молекул, таких как пептид или молекула антитела scFv. Например, связывающие молекулы могут представлять собой полипептиды, имеющие общую структуру и имеющие одну или более областей аминокислотной последовательности, в которых имеются различия.

Эта можно проиллюстрировать, рассмотрев набор молекул антител. Они могут представлять собой молекулы антител, относящиеся к классу с общей структурой, Например, IgG, Fab, scFv-Fc или scFv, различающихся по одной или более областей их последовательности. Молекулы антител обычно характеризуются переменностью последовательностей в областях, определяющих комплементарность (CDR), которые представляют собой области, что главным образом вовлечены в распознавание антигена. Набор связывающих молекул в настоящем изобретении может представлять собой набор молекул антител, которые различаются по одной или более CDR, например, различия могут иметься во всех шести CDR или в одной, или более определенных CDR, таких как CDR3 тяжелой цепи и/или CDR3 легкой цепи.

Молекулы антител и других связывающих молекул более подробно описаны в других разделах в данном документе. Однако потенциал настоящего изобретения не ограничивается дисплеем антител и охватывает дисплей библиотек пептидов или сконструированных белков, включая рецепторы, лиганды, отдельные белковые домены и

альтернативные каркасные белки [58, 59]. Направляемое нуклеазой сайт-специфическое встраивание может быть использованы для создания библиотек других типов связывающих молекул, ранее конструируемых с использованием других систем дисплея. Многие из них включают мономерные связывающие домены, такие как дарпины и липокалины, аффитела и адхироны [58, 59, 152]. Дисплей на клетках эукариот, особенно клетках млекопитающих, также открывает возможность для выделения и конструирования связывающих молекул или мишеней, включая более сложные, мультимерные мишени. Например, Т-клеточные рецепторы (TCR) экспрессированы на Т-клетках и эволюционно распознают пептид, презентируемый в комплексе с молекулами ГКГС на антиген-презентирующих клетках. Могут быть созданы библиотеки, кодирующие и экспрессирующие набор TCR, и может быть произведен их скрининг для идентификации связывания с пептидными комплексами ГКГС, как дополнительно описано в других разделах в данном документе.

В случае мультимерных связывающих молекул, донорная ДНК, кодирующая связывающую молекулу, может быть представлена в виде одной или более молекул ДНК. Например, если необходимо отдельно экспрессировать VH- и VL-домены, их могут кодировать отдельные молекулы донорной ДНК. Донорная ДНК встраивается в клеточную ДНК во многих сайтах встраивания, например, ген связывающей молекулы для VH в одном локусе и ген связывающей молекулы для VL во втором локусе. Способы введения донорной ДНК, кодирующей отдельные субъединицы связывающей молекулы, описаны более подробно в других разделах в данном документе. В качестве альтернативы, обе субъединицы или части мультимерной связывающей молекулы могут кодироваться одной и той же молекулой донорной ДНК, которая встраивается в фиксированный локус.

Связывающая молекула может представлять собой молекулу антитела или белок, отличный от антитела, которая содержит антигенсвязывающий сайт. Антигенсвязывающий сайт может быть получен посредством компоновки пептидных петель на каркасных белках, отличных от антитела, таких как фибронектин или цитохром В и т. д. или путем рандомизации, или мутации аминокислотных остатков петли внутри каркасного белка для обеспечения связывания с требуемой мишенью [60, 61, 62]. Каркасные белки для миметиков антител описаны в источнике WO/0034784, в котором авторы изобретения описывают белки (миметики антител), которые содержат домен фибронектина типа III, имеющий по меньшей мере одну рандомизированную петлю. Подходящий каркасный белок, в который может быть встроена одна или более пептидных петель, Например, набор петель VH CDR антитела, может быть получен из любого домена представителя суперсемейства генов иммуноглобулинов. Каркасный белок может представлять собой человеческий белок или белок, не относящийся человеческому.

Применение антигенсвязывающих сайтов в каркасных белках, отличных от антитела, было ранее рассмотрено [63]. Обычно такие белки имеют стабильный остов и одну или более вариабельных петель, причем в аминокислотную последовательность петли или петель целенаправленно, или случайным образом вносят мутации для получения антигенсвязывающего сайта, обеспечивающего связывание с антигеном-мишенью. Такие белки включают IgG-связывающие домены белка А из *S. aureus*, трансферрин, тетранектин, фибронектин (например, 10-й домен фибронектина типа III) и липокалины. Другие подходы включают применение небольших пептидов с ограниченной конформацией, например, на основе «ноттина» и каркасные белки циклотиды [64]. Учитывая их небольшой размер и сложность, особенно что касается правильного образования дисульфидной связи, преимущества могут заключаться в применении эукариотических клеток для отбора новых связывающих молекул на основе этих каркасных белков. Учитывая общие функции этих пептидов в природе, преимущества могут заключаться в применении библиотек связывающих молекул на основе этих каркасных белков для создания небольших высокоаффинных связывающих молекул, особенно что касается применения для блокировки ионных каналов и протеаз.

Помимо последовательностей антител и/или антигенсвязывающего сайта связывающая молекула может содержать другие аминокислоты, Например, формирующие пептид или полипептид, как, например, домен со складчатой структурой, или для придания молекуле другой функциональной характеристики помимо способности к связыванию антигена. Связывающая молекула может нести детектируемую метку или может быть конъюгирована с токсином, или нацеливающим фрагментом, или ферментом (например, через пептидильную связь или линкер). Например, связывающая молекула может включать каталитический сайт (например, в домене фермента), а также антигенсвязывающий сайт, при этом антигенсвязывающий сайт связывается с антигеном и, таким образом, нацеливает каталитический сайт на антиген. Каталитический сайт может подавлять биологическую функцию антигена, Например, путем расщепления.

#### *Молекулы антител*

Молекулы антител являются предпочтительными связывающими молекулами. Молекулы антител могут представлять собой полноразмерные антитела или иммуноглобулины (Ig), которые имеют четыре полипептидные цепи - две идентичные тяжелые цепи и две идентичные легкие цепи. Тяжелые и легкие цепи формируют пары, каждая из них представляет собой пару VH-VL-доменов, которые содержат антигенсвязывающий сайт. Тяжелые и легкие цепи также содержат константные области: CL легкой цепи и CH1, CH2, CH3 и иногда CH4 тяжелой цепи (пятый домен CH4

содержится в IgM и IgE человека). Две тяжелые цепи соединены дисульфидными мостиками в гибкой шарнирной области. Молекула антитела может содержать VH- и/или VL-домен.

Наиболее распространенным природным форматом молекулы антитела является IgG, которые представляет собой гетеротетрамер, состоящий из двух идентичных тяжелых цепей и две двух идентичных легких цепей. Тяжелые и легкие цепи состоят из модульных доменов с консервативной вторичной структурой, состоящей из четырехслойного антипараллельного бета-складчатого слоя и трехслойного антипараллельного бета-складчатого слоя, стабилизированных одной дисульфидной связью. Каждая из тяжелых цепей антител имеет N-концевой варибельный домен (VH) и 3 относительно консервативных «константных» иммуноглобулиновых домена (CH1, CH2, CH3), тогда как легкая цепь имеет один N-концевой варибельный домен (VL) и один константный домен (CL). Дисульфидные связи стабилизируют отдельные домены и формируют ковалентные связи для соединения четырех цепей в стабильный комплекс. VL и CL легкой цепи ассоциированы с VH и CH1 тяжелой цепи, и эти элементы могут быть экспрессированы отдельно с формированием Fab-фрагмента. CH2- и CH3-домены (также называемые «Fc-доменом») ассоциированы с другой парой CH2:CH3 с образованием тетрамерной Y-образной молекулы, причем варибельные домены из тяжелой и легкой цепей расположены на вершине «Y». CH2- и CH3-домены отвечают за взаимодействия с эффекторными клетками и компонентами системы комплемента в пределах иммунной системы. Рекомбинантные антитела были ранее экспрессированы в формате IgG или в виде Fab-фрагментов (состоящих из димера VH:CH1 и легкой цепи). В дополнение к этому, может быть использована искусственная конструкция, называемая одноцепочечным Fv (scFv), состоящая из ДНК, кодирующей VH- и VL- фрагменты, генетически слитые с ДНК, кодирующей гибкий линкер.

Связывающие молекулы могут представлять собой молекулы антител человека. Таким образом, в случае наличия константных доменов, они предпочтительно представляют собой константные домены человека.

Связывающие молекулы могут представлять собой фрагменты антител или молекулы антител меньшего формата, как, например, одноцепочечные молекулы антител. Например, молекулы антител могут представлять собой молекулы scFv, состоящие из VH-домена и VL-домена, соединенных линкерным пептидом. В молекуле scFv VH- и VL-домены формируют пару VH-VL, в которой области, определяющие комплементарность, VH- и VL-доменов объединяются с образованием антигенсвязывающего сайта.

Другие фрагменты антител, которые содержат антигенсвязывающий сайт антитела, включают, в том числе, (i) Fab-фрагмент, состоящий из VL-, VH-, CL- и CH1-доменов; (ii) Fd-фрагмент, состоящий из VH- и CH1-доменов; (iii) Fv-фрагмент, состоящий из VL- и VH-доменов одного антитела; (iv) dAb-фрагмент [65, 66, 67], который состоит из VH- или VL-домена; (v) выделенные CDR-области; (vi) F(ab')<sub>2</sub>-фрагменты, бивалентный фрагмент, содержащий два связанных Fab-фрагмента (vii) scFv, причем VH-домен и VL-домен связаны пептидным линкером, что обеспечивает ассоциацию двух доменов с образованием антигенсвязывающего сайта [68, 69]; (viii) биспецифические одноцепочечные Fv-димеры (PCT/US92/09965) и (ix) «диатела», мультивалентные или мультиспецифические фрагменты, сконструированные путем слияния генов (WO94/13804; [70]). Fv, scFv или молекулы диател могут быть стабилизированы путем включения дисульфидных мостиков, связывающих VH- и VL-домены [71].

Были сконструированы различные другие молекулы антител, включая один или более антигенсвязывающих сайтов антитела, включая, например, Fab<sub>2</sub>, Fab<sub>3</sub>, диатела, триатела, тетратела и минитела (небольшие иммунные белки). Были описаны молекулы антител и способы их создания и применения [72].

Другими примерами связывающих фрагментов являются Fab', который отличается от Fab-фрагмента вставкой нескольких остатков на карбоксильном конце CH1-домена тяжелой цепи, включая один или более цистеиновых остатков из шарнирной области антитела, и Fab'-SH, который представляет собой Fab'-фрагмент, в котором цистеиновый(е) остаток(ки) константных доменов несут свободную тиоловую группу.

dAb (доменное антитело) представляет собой небольшой мономерный антигенсвязывающий фрагмент антитела, а именно переменную область тяжелой или легкой цепи антитела. VH dAb встречаются у представителей верблюжьих (например, верблюдов, лама) и могут быть получены путем иммунизации верблюжьих антигеном-мишенью, выделения антиген-специфических В-клеток и непосредственного клонирования генов dAb, полученных из отдельных В-клеток. dAb также могут быть получены в культуре клеток. Их небольшие размер, хорошая растворимость и термостабильность делают их особенно пригодными и подходящими с физиологической точки зрения для отбора и созревания аффинности. VH dAb верблюжьих разрабатываются для применения в терапевтических целях под названием «nanobodies™».

Синтетические молекулы антител могут быть созданы путем экспрессии генов, полученных с помощью олигонуклеотидов, синтезированных и собранных в составе подходящих экспрессирующих векторов, например, как было описано Knappik *et al.* [73] или Krebs *et al.* [74].

Биспецифические или бифункциональные антитела формируют второе поколение моноклональных антител, в которых две разные переменные области объединены в одной и той же молекуле [75]. Их применение было продемонстрировано как в области диагностики, так и в области терапии из-за их способности к вовлечению новых эффекторных функций или целенаправленного связывания с несколькими молекулами на поверхности опухолевых клеток. Если необходимо использовать биспецифические антитела, они могут представлять собой стандартные биспецифические антитела, которые могут быть изготовлены многими способами [76], например, могут быть получены химическими методами или из межвидовых гибридом, или могут представлять собой любые из фрагментов биспецифических антител, упомянутые выше. Эти антитела смогут быть получены с помощью химических методов [77, 78] или соматических методов [79, 80], но аналогичным образом и предпочтительно с помощью методов генетической инженерии, которые обеспечивают гетеродимеризацию и, таким образом, облегчают процесс необходимой очистки антител [81]. Примеры биспецифических антител включают антитела, полученные с помощью технологии ViTE™, в которой могут быть использованы связывающие домены двух антител с разной специфичностью и непосредственно связаны через короткие гибкие пептиды. Это позволяет объединить два антитела на одной короткой полипептидной цепи. Диатела и scFv могут быть сконструированы без Fc-области с использованием только переменных доменов, что позволяет потенциально уменьшить влияние антиидиотипических взаимодействий.

Биспецифические антитела могут быть сконструированы в виде полного IgG, в виде биспецифического Fab'2, в виде Fab'PEG, в виде диател или же в виде биспецифического scFv. Кроме того, два биспецифических антитела могут быть связаны с использованием стандартных способов, известных в данной области, с образованием тетравалентных антител.

Биспецифические диатела, в отличие от биспецифических полноразмерных антител также могут оказаться особенно пригодными. Отбор диател (и многих других полипептидов, как, например, фрагментов антител), обладающих соответствующей специфичностью связывания, не представляет никаких трудностей. Если один участок диатела не должно изменяться, например, со специфичностью, направленной против представляющего интерес антигена, то может быть создана библиотека, при этом изменяется другой участок и происходит отбор антитела с соответствующей специфичностью. Биспецифические полноразмерные антитела могут быть созданы с помощью альтернативных методов генной инженерии, которые описаны в источнике Ridgeway *et al.*, 1996 [82].

Библиотека согласно настоящему изобретению может быть использована для отбора молекулы антитела, которая связывает один или более антигенов, представляющих интерес. Отбор из библиотек подробно описан ниже. После отбора затем может быть сконструирована молекула антитела в другом формате и/или с включением дополнительных признаков. Например, отобранная молекула антитела может быть преобразована в другой формат, как. Например, один из форматов антител, описанных выше. Отобранные молекулы антител и молекулы антител, содержащие CDR VH и/или VL отобранных молекул антител, относятся к одному из аспектов настоящего изобретения. Молекулы антител и кодирующая их нуклеиновая кислота могут быть получены в чистом виде.

Могут быть получены фрагменты антител, начиная с молекулы антитела, с помощью способов, таких как расщепление ферментами, например, пепсином или папаином, и/или путем расщепления дисульфидных мостиков с помощью химического восстановления. Другим способом фрагменты антител могут быть получены с помощью методов генетической рекомбинации, хорошо известных специалисту в данной области, или же с помощью пептидного синтеза, например, с помощью автоматизированных синтезаторов пептидов или с помощью синтеза и экспрессии нуклеиновых кислот.

Существует возможность взять моноклональные и другие антитела и применить методы технологии рекомбинантной ДНК для получения других антител или рекомбинантных молекул, которые связывают антиген-мишень. Такие методы могут включать введение ДНК, кодирующей вариабельную область иммуноглобулина, или CDR одного антитела в константные области или константные области плюс каркасные области другого иммуноглобулина. См., например, EP-A-184187, GB 2188638A или EP-A-239400 и многочисленные последующие литературные источники.

Молекулы антител могут быть отобраны из библиотеки и затем модифицированы, например, время полужизни молекулы антитела *in vivo* может быть увеличено за счет химической модификации, например, пегилирования или за счет заключения в липосому.

#### *Источники генов связывающих молекул*

В традиционном методе создания моноклональных антител используется иммунная система лабораторных животных, таких как мыши и кролики, для создания пула высокоаффинных антител, которые затем выделяют путем применения гибридомной технологии. В настоящем изобретении предложен альтернативный метод идентификации антител, вырабатываемых в результате иммунизации. Гены VH- и VL-цепей могут быть амплифицированы после получения из В-клеток иммунизированных животных, и клонированы в соответствующий вектор для введения в библиотеки эукариотических

клеток с последующим отбором из этих библиотек. Фаговый дисплей и рибосомный дисплей дает возможность конструировать очень большие библиотеки ( $>10^9$  клонов), обеспечивая выделение антител человека в отсутствие иммунизации. Настоящее изобретение также может быть использовано в сочетании с такими способами. После раундов отбора с помощью фагового дисплея отобранная популяция связывающих молекул может быть введена в эукариотические клетки с помощью направляемого нуклеазой встраивания, как описано в данном документе. Это дает возможность первоначального применения очень больших библиотек, основанных на других системах (например, фаговый дисплей), для обогащения популяции связывающих молекул, при этом обеспечивая их эффективный скрининг с использованием эукариотических клеток, как описано выше. Таким образом, в настоящем изобретении могут сочетаться наилучшие признаки как фагового дисплея, так и дисплея на поверхности эукариотических клеток для получения высокоэффективной системы, обеспечивающей количественный скрининг и сортировку.

С использованием фагового дисплея и дрожжевого дисплея ранее было продемонстрировано, что также существует возможность создания связывающих молекул, не прибегая к иммунизации, при условии, что используются дисплейные библиотеки достаточного размера. Например, множество связывающих молекул было создано с использованием библиотеки неиммунных антител, состоящей из  $>10^7$  клонов [8 3]. Это в свою очередь дает возможность создания связывающих молекул, специфических к мишеням, которые трудно получить с помощью традиционных методов иммунизации, например, создания антител к «аутоантигенам» или эпитопам, которые консервативны у разных видов. Например, перекрестно-реагирующие связывающие молекулы человека/мыши могут быть обогащены путем последовательного отбора, произведённого на вариантах аналогичной молекулы-мишени человека, а затем мыши. Поскольку специфическая иммунизация людей против представляющих интерес молекул-мишеней не представляется возможной, это свойство является особо важным для создания антител человека, которые являются предпочтительными для реализации терапевтических подходов.

До настоящего времени в примерах дисплея на клетках млекопитающих, в которых размеры и качество библиотеки были ограниченными, связывающие молекулы были только созданы с использованием наборов, которые были предварительно обогащены связывающими молекулами, например, в результате иммунизации или в результате конструирования уже существующих связывающих молекул. Способность получения больших библиотек в эукариотических клетках, а особенно в клетках высших эукариот



создает возможность для выделения связывающих молекул непосредственно из этих библиотек, начиная с неиммунных связывающих молекул или связывающих молекул, которые ранее не были отобраны в другой системе. С настоящим изобретением становится возможным создание связывающих молекул из неиммунных источников. Это в свою очередь открывает возможности для использования генов связывающих молекул из многих источников. Гены связывающих молекул могут быть получены в результате ПЦР с использованием генов природных источников, как, например, генов антител. Гены связывающих молекул также могут быть повторно клонированы на основе существующих библиотек, как, например, библиотек антител, полученных с помощью фагового дисплея, и клонированы в подходящий донорный вектор для направляемого нуклеазой встраивания в клетки-мишени. Связывающие молекулы могут быть полностью или частично синтетическими по происхождению. Кроме того, различные типы связывающих молекул описаны в других разделах в данном документе, например, гены связывающих молекул могут кодировать антитела или могут кодировать альтернативные каркасные белки [58, 59], пептиды или сконструированные белки, или белковые домены.

#### *Дисплей связывающих молекул*

Для получения набора связывающих молекул с целью скрининга мишеней, представляющих интерес, клетки библиотеки могут быть культивированы для экспрессии связывающих молекул либо в растворимой, секретируемой форме, либо в трансмембранной форме. Для дисплея на клеточной поверхности необходимо удержание экспрессированной связывающей молекулы на клеточной поверхности, которая ее кодирует. Связывающие молекулы могут содержать мембранный якорь или быть связанными с ним, как, например, трансмембранный домен, для внеклеточного дисплея связывающей молекулы на клеточной поверхности. Это может включать прямое слияние связывающей молекулы с сигнальной последовательностью, отвечающей за локализацию на мембране, как, например, распознаваемая последовательность GPI, или с трансмембранным доменом, как, например, трансмембранный домен рецептора PDGF [84]. Удержание связывающих молекул на клеточной поверхности также может осуществляться опосредованно путем ассоциации с другой удерживаемой на клеточной поверхности молекулой, экспрессированной в той же самой клетке. Эта ассоциированная молекула как таковая может входить в состав гетеродимерной связывающей молекулы, как, например, закрепленной тяжелой цепи антитела в сочетании с партнером - легкой цепью, которая закреплена опосредовано.

Хотя иммобилизация на клеточной поверхности облегчает отбор связывающих молекул, во многих случаях применения требуется получение секретируемых связывающих молекул в отсутствие иммобилизации. Существует возможность сочетания

закрепления в мембране и секрети в растворимой форме с использованием способа повторного улавливания для прикрепления секретируемых связывающих молекул к рецепторам клеточной поверхности. Один подход заключается в форматировании библиотеки связывающих молекул в качестве секретируемых молекул, которые могут связываться с заякоренной в мембране молекулой, экспрессированной в той же самой клетке, функция которой может состоять в улавливании секретируемых связывающих молекул. Например, в случае антител или связывающих молекул, слитых с Fc-доменами антител, закрепленный на мембране Fc-фрагмент может «отбирать» секретируемые связывающие молекулы, которые экспрессируются в той же самой клетке, что приводит к дисплею мономерной фракции связывающих молекул, которые экспрессируются, тогда как остальные молекулы секретируются в бивалентной форме (US 8551715). Альтернатива заключается в применении закрепленного связывающего домена IgG, как, например, белка А.

Другие способы удержания секретируемых антител с использованием клеток, которые их продуцируют, рассмотрены в источнике Kumar et al. (2012) [8 5] и включают инкапсуляцию клеток внутри микрокапель, улавливание с помощью матрикса, дисплей на поверхности посредством улавливания по степени аффинности (ACSD), технологию секретиции и улавливания (SECANT) и «холодного улавливания» [85]. В примерах, приведенных для описания ACSD и SECANT [85], для облегчения иммобилизации стрептавидина или улавливающего антитела на клеточной поверхности используется биотинилирование. Связанная молекула в свою очередь улавливает секретируемые антитела. В примере с SECANT биотинилирование секретируемой молекулы происходит *in vivo*. С использованием метода «холодного улавливания» секретируемое антитело может быть обнаружено на продуцирующих клетках с использованием антител, специфических к секретируемой молекуле. Было выдвинуто предположение, что это обусловлено связью секретируемого антитела с гликокаликсом клетки [8 6]. В качестве альтернативы было предложено, что секретируемый продукт захватывается окрашенными антителами на клеточной поверхности до эндоцитоза [87]. Приведенные выше способы были использованы для идентификации клонов с высоким уровнем экспрессии в пределах популяции, но потенциально они могут быть адаптированы для идентификации специфичности связывания, при условии, что связь с клеточной поверхностью характеризуется достаточной продолжительностью.

Даже в том случае, когда связывающая молекула непосредственно закреплена на клеточной поверхности, существует возможность получения растворимого продукта. Например, ген, кодирующий отобранную связывающую молекулу, может быть выделен и

клонирован в экспрессирующий вектор, в котором отсутствует заякоренная в мембране последовательность. В качестве альтернативы и как продемонстрировано в Примерах, может быть использована экспрессирующая конструкция, в которой трансмембранный домен кодируется в составе экзона, фланкированного сайтами рекомбинации, Например, сайтами распознавания ROX для рекомбиназы Dre [88]. В этом примере экзон, кодирующий трансмембранный домен, может быть удален путем трансфекции гена, кодирующего рекомбиназу Dre, для переключения экспрессии на секретлируемую форму. Как продемонстрировано в данном документе, секретлируемое антитело было получено по этому способу в отсутствие необходимости в действии рекомбиназы. Предположительно это произошло в результате альтернативного сплайсинга [89].

Любые из приведенных выше способов или другие подходящие подходы могут быть использованы для обеспечения того, что связывающие молекулы, экспрессированные с использованием клонов библиотеки, представлены на поверхности клеток, которые их экспрессируют.

*Скрининг для идентификации связывающих молекул, специфических по отношению к мишени, представляющей интерес*

Как отмечалось выше, библиотека эукариотических клеток может быть использована в способе скрининга связывающих молекул, которые распознают мишень. Такой способ может включать: получение библиотеки, как описано в данном документе, культивирование клеток библиотеки для экспрессии связывающих молекул, экспонирование мишеней со связывающими молекулами, обеспечение распознавания мишени одной или более связывающими молекулами, в случае наличия, и обнаружение распознавания мишени когнатной связывающей молекулой.

Отборы могут быть проведены с использованием целого ряда классов молекул-мишеней, Например, белка, нуклеиновой кислоты, углевода, липида, небольших молекул. Мишень может быть представлена в растворимой форме. Мишень может быть мечена для облегчения обнаружения, Например, она может нести а флуоресцентную метку или она может быть биотинилирована. Клетки, экспрессирующие связывающую молекулу, специфическую к мишени, могут быть выделены с использованием прямо или опосредованно меченой молекулы-мишени, при этом связывающая молекула улавливает меченую молекулу. Например, клетки, которые связываются посредством взаимодействия связывающей молекулы и мишени с флуоресцентно меченой мишенью, могут быть обнаружены и отсортированы с помощью проточной цитометрии или FACS для выделения требуемых клеток. Для проведения отборов, включающих цитометрию, требуются молекулы-мишени, которые непосредственно мечены флуоресцентной меткой или мечены

молекулами, которые могут быть обнаружены с использованием вторичных реагентов, например, к клеткам может быть добавлена биотинилированная молекула-мишень, а связывание с клеточной поверхностью может быть обнаружено с использованием меченого флуоресцентной меткой стрептавидина, как, например, стрептавидина-фикоэритрина. Дополнительная возможность заключается в иммобилизации молекулы-мишени или вторичных реагентов, которые связываются с мишенью, на твердой подложке, как, например, магнитных микрогранулах или агарозных микрогранулах, для обеспечения обогащения клеток, которые связывают мишень. Например, клетки, которые связываются посредством взаимодействия связывающей молекулы и мишени с биотинилированной молекулой-мишенью, могут быть выделены на субстрате, покрытом стрептавидином, например, микрогранулах, покрытых стрептавидином.

При скрининге библиотек предпочтительно делать выборку в избытке, т.е. использовать больше клонов, чем количество независимых клонов, содержащихся в составе библиотеки, для обеспечения эффективного представления библиотеки. Идентификация связывающих молекул из очень больших библиотек, получаемых по настоящему изобретению, может быть осуществлена с помощью проточной сортировки, но может занять несколько дней, особенно в случае избыточной выборки библиотеки. В качестве альтернативы первоначальные отборы могут быть основаны на применении антигена, поддающегося выделению, например, биотинилированного антигена, выделенного на магнитных микрогранулах, покрытых стрептавидином. Таким образом, магнитные микрогранулы, покрытые стрептавидином, могут быть использованы для улавливания клеток, которые связались с биотинилированным антигеном. Отбор с помощью магнитных микрогранул может быть использован в качестве единственного способа отбора или он может осуществляться в сочетании с проточной цитометрией, в случае чего может быть достигнуто лучшее разрешение, например, дифференциация между клоном с более высокими уровнями экспрессии и клоном с более высокой аффинностью [56, 57].

Природа *in vitro* подходов дисплейных технологий делает возможным контроль отбора таким образом, что не представляется возможным за счет иммунизации, например, путем отбора при определенном конформационном состоянии мишени [90, 91]. Мишени могут быть мечены посредством химической модификации (флуоресцеин, биотин) или путем слияния генов (например, белок, слитый с эпитопной меткой, как, например, метка FLAG, или другим доменом белка или целым белком). В качестве метки может выступать нуклеиновая кислота (например, ДНК, РНК или небиологические нуклеиновые кислоты), причем метка представляет собой участок, слитый с целевой нуклеиновой кислотой, или может быть присоединена химическими методами к молекуле другого типа, как, например,

белку. Это может достигаться посредством химической конъюгации или посредством ферментативного присоединения [92]. Нуклеиновая кислота также может быть слита с мишенью в процессе трансляции, как, например, в случае рибосомного дисплея. «Метка» может представлять собой другую модификацию, встречающуюся внутри клетки (например, гликозилирование, фосфорилирование, убиквитинирование, алкилирование, палирование, сумоилирование и другие модификации, описанные в база данных посттрансляционных модификаций (db-PTM) на сайте <http://dbptm.mbc.nctu.edu.tw/statistics.php>), которые могут быть обнаружены с помощью вторичных реагентов. Это позволит получить связывающие молекулы, которые связывают неизвестный белок-мишень на основе определенной модификации.

Мишени могут быть обнаружены с использованием существующих связывающих молекул, которые связываются с такой молекулой-мишенью, Например, антител, специфических к мишеням. Применение существующих связывающих молекул для обнаружения будет давать дополнительное преимущество идентификации связывающих молекул в составе библиотеки связывающих молекул, которые распознают эпитоп, отличный от связывающей молекулы, используемой для обнаружения. Таким способом могут быть идентифицированы пары связывающих молекул для таких применений, как сэндвич-ИФА. По возможности предпочтительно использовать очищенную молекулу-мишень. В качестве альтернативы мишень может быть представлена на поверхности популяции клеток-мишеней, а связывающие молекулы представлены на поверхности клеток библиотеки, причем способ включает экспонирование мишеней со связывающими молекулами путем приведения клеток библиотеки в контакт с клетками-мишенями. Выделение клеток, экспрессирующих мишень (например, с использованием клеток, экспрессирующих биотинилированную мишень), дает возможность обогащения клеток, которые экспрессируют молекулы, связывающиеся с ними. Использование данного подхода будет полезным в случаях, когда вовлечены низкоаффинные взаимодействия, поскольку существует потенциал для сильного эффекта avidности.

Молекула-мишень также может представлять собой неочищенные рекомбинантные или неочищенные природные мишени, при условии, что доступна детектирующая молекула для идентификации связывания с клеткой (как описано выше). В дополнение к этому, связывание молекулы-мишени с клеткой, экспрессирующей связывающую молекулу, может быть обнаружено опосредованно через связь молекулы-мишени с другой молекулой, которую подлежит обнаружению, например, клеточный лизат, содержащий меченую молекулу, может быть инкубирован с библиотекой связывающих молекул для идентификации молекул, не только с меченой молекулой, но также, молекул,

связывающихся с ассоциированными белками-партнерами. В результате этого будет получена панель антител к этим партнерам, которая может быть использована для обнаружения или идентификации партнера (например, с использованием масс-спектрометрии). Фракционирование клеток может быть использовано для обогащения молекул-мишеней, полученных из определенных субклеточных структур. В качестве альтернативы дифференциальное биотинилирование поверхности или цитоплазматических фракций может быть использовано в сочетании с детектирующими реагентами на основе стрептавидина для дисплея на поверхности эукариотических клеток [93,94]. Применение препаратов молекул-мишеней, солюбилизованных с помощью поверхностно-активных веществ, является особенно пригодным подходом для интактных мембранных белков, таких как GPCR и ионные каналы, которые в других случаях сложно получить. Наличие поверхностно-активных веществ может оказывать отрицательное воздействие на эукариотические клетки, несущие связывающие молекулы, что требует выделения генов связывающих молекул без дальнейшего выращивания отобранных клеток.

После обнаружения распознавания мишени когнатной связывающей молекулой, могут быть выделены клетки клона, содержащего ДНК, кодирующую когнатную связывающую молекулу. Затем может быть выделена ДНК, кодирующая связывающую молекулу (например, идентифицированная или амплифицированная) с использованием выделенного клона, в результате чего получают ДНК, кодирующую связывающую молекулу, которая распознает мишень.

Примеры связывающих молекул и мишеней подробно описаны других разделах в данном документе. Классическим примером является библиотека молекул антител, которые могут подвергаться скринингу в отношении связывания с представляющим интерес антигеном-мишенью. Другие примеры включают скрининг библиотеки TCR против целевого комплекса молекул ГКГС и пептида или скрининг библиотеки комплексов пептида и молекул ГКГС против целевого TCR.

#### *Взаимодействие TCR, молекул ГКГС и других рецепторов*

T-клеточные рецепторы (TCR) экспрессированы на T-клетках и эволюционно распознают пептид, презентированный в комплексе с молекулами ГКГС на антиген-презентирующих клетках. TCR представляют собой гетеродимеры, состоящие в 95 % случаев из гетеродимеров альфа и бета и в 5 % случаев из гетеродимеров гамма и дельта. Обе мономерные единицы имеют N-концевой домен иммуноглобулина, который имеет 3 переменные области, определяющие комплементарность (CDR), вовлеченные в осуществление взаимодействия с мишенью. Функциональный TCR содержится в составе комплекса с другими субъединицами, а передача сигнала усиливается за счет коstimуляции

с помощью молекул CD4 и CD8 (молекулы, специфические по отношению к молекулам ГКГС класса I и класса II, соответственно). На антиген-презентирующих клетках белки подвергаются процессингу и презентации на клеточной поверхности в комплексе с молекулами ГКГС, которые сами по себе входят в состав мультимерного белкового комплекса. TCR, распознающие «собственные» пептиды, удаляются на протяжении развития, а сборка системы для распознавания чужеродных пептидов, представленных на антиген-презентирующих клетках, происходит для вызова иммунного ответа. Результат распознавания комплекса пептида и молекул ГКГС зависит от типа T-клеток и аффинности такого взаимодействия.

Большую ценность представляет идентификация генов, кодирующих TCR или комплексы пептида и молекул ГКГС, которые отвечают за взаимодействия, вовлеченные в патологический процесс, Например, который характерен для аутоиммунного заболевания. В случае аутоиммунного заболевания идентификация взаимодействующих партнеров, Например, TCR, отвечающих за патологический процесс, может создать условия либо для специфической блокировки таких взаимодействий, либо для удаления цитотоксических клеток с нарушенной функцией. Было бы желательно сконструировать TCR для изменения свойств связывания, например, за счет повышения аффинности к представляющим интерес мишеням, Например, путем перенацеливания T-клеток в случае злокачественной опухоли или усиления воздействия существующих T-клеток [95]. В качестве альтернативы поведение регуляторных или супрессорных T-клеток может быть изменено в качестве терапевтического воздействия, Например, для ведения или усиления иммунотерапии злокачественной опухоли путем введения специфических TCR в T-клетки или путем использования экспрессированного белка TCR в качестве компонентов терапии [96].

Дисплей библиотек TCR на поверхности дрожжевых клеток и клеток млекопитающих был ранее продемонстрирован. В случае дрожжевых клеток было необходимо сконструировать TCR и представить его в формате одноцепочечной молекулы. Поскольку аффинность взаимодействия между TCR и комплексом пептида, и молекулами ГКГС является низкой, растворимый компонент (например, в данном случае пептид и МНС) обычно представлен в формате мультимерной молекулы. Были сконструированы пептиды со специфичностью TCR в комплексе с молекулами ГКГС класса I [97] и ГКГС класса II [98]. TCR также были экспрессированы на поверхности мутантных T-клеток мыши (в которых отсутствуют альфа- и бета-цепи TCR), и были выделены варианты TCR с улучшенными связывающими свойствами [99]. Например, в источнике Chervin *et al.* описано введение TCR путем заражения ретровирусами, и была создана библиотека с действительным размером  $10^4$  клонов [100]. С использованием направляемого нуклеазой

встраивания связывающих молекул, который предложен в данном документе, может быть использован аналогичный подход для конструирования Т-клеток. Наряду с отбором TCR с измененными свойствами распознавания могут быть использованы дисплейные библиотеки для скрининга библиотек пептидов или вариантов молекул ГКГС для распознавания TCR. Например, комплексы пептида и молекул ГКГС были представлены на поверхности клеток насекомых и использованы для картирования эпитопов TCR, представленных в формате мультимерной молекулы [101].

Как отмечалось выше, способы скрининга могут включать дисплей набора связывающих молекул на клеточной поверхности и исследование с использованием мишени, представленной в виде растворимой молекулы, которая может представлять собой мультимерную мишень. В качестве альтернативы, которая может быть особенно полезной в случае мультимерных мишеней, может быть использован непосредственный скрининг межклеточных взаимодействий, при которых связывающая молекула и мишень представлены на поверхности разных клеток. Например, если активация представляющего интерес TCR приводит к экспрессии репортерного гена, это может быть использовано для идентификации активирующих пептидов или активирующих молекул МНС, содержащихся в составе библиотеки пептидов и МНС. В этом конкретном примере репортерная клетка не кодирует молекулу из библиотеки, но может быть использована для идентификации клетки, которая действительно кодирует ее. Данный подход потенциально может распространяться на подход, основанный на «сравнении одной библиотеки с другой». Например, развивая пример, описанный выше, может быть произведен скрининг библиотеки TCR против библиотеки пептидов и МНС. В более широком смысле пример скрининга библиотеки связывающих молекул, представленных на поверхности одной клетки, с использованием партнера по связыванию на другой клетке, может быть перенесен на другие типы межклеточных взаимодействий, например, для идентификации связывающих молекул, которые подавляют или активируют передачу сигнала в пределах путей Notch или Wnt. Таким образом, настоящее изобретение может быть использовано в альтернативных клеточных системах скрининга, включая системы распознавания на основе межклеточных взаимодействий.

В качестве примера, рекомбинантные рецепторы антигенов («CARS») образуются в результате слияния между связывающим доменом антитела (обычно имеющим формат в scFv) и сигнальным доменом. Они были введены в Т-клеток с целью перенацеливания Т-клеток *in vivo* для атаки опухолевых клеток посредством распознавания антитела и связывания с опухоль-специфическими антигенами. Целый ряд разных факторов может повлиять на успех данной стратегии, включая сочетание специфичности антитела, формата,



аффинности антитела, длину линкера, слитый сигнальный модуль, уровень экспрессии в Т-клетках, подтип Т-клеток и взаимодействие CAR с другими сигнальными молекулами [102,103]. Возможность создания больших библиотек CAR с использованием первичной культуры Т-клеток, включающих отдельные переменные, приведенные выше, и или их комбинации, позволит осуществлять функциональный поиск эффективной и оптимальной конструкции CAR. Этот функциональный «поиск» может осуществляться *in vitro* или *in vivo*. Например, в источнике Alonso-Camino (2009) описано слияние scFv, распознающего СЕА, с  $\zeta$ -цепью комплекса TCR:CD3 и введение этой генетической конструкции в клеточную линию Jurkat человека [104]. При взаимодействии с СЕА, содержащимся либо на клетках HeLa, либо на опухолевых клетках, наблюдалось повышение экспрессии раннего маркера активации Т-клеток CD69. Данный подход может быть использован для идентификации слитых конструкций CAR с соответствующими активирующими или ингибирующими свойствами с использованием культивированных клеток или первичной культуры клеток.

На следующем этапе может быть проведена оценка функциональности конструкций CAR *in vivo*. Например, библиотека CAR, полученных в первичной культуре Т-клеток мыши, может быть введена мышью, несущей опухоль, для идентификации Т-клеточных клонов, которые активно пролиферируют при взаимодействии с опухолью. В случае необходимости может быть проведен предварительный отбор этой библиотеки Т-клеток на основе специфичности связывания с антигеном с использованием способов, описанных выше. В любом случае библиотека вносимых связывающих молекул может быть использована для замены существующей связывающей молекулы (например, ГКГС или TCR, или переменный домен антитела).

### *Скрининг фенотипа*

В данном документе описаны различные способы отбора связывающих молекул, с помощью которых можно модифицировать передачу сигнала в клетках и клеточное поведение.

Библиотека может подвергаться скринингу в отношении изменения клеточного фенотипа, вызванного действием связывающей молекулы на ее мишень.

Антитела, которые способны модифицировать передачу сигнала в клетках путем связывания с лигандами или рецепторами, характеризуются высокой эффективностью при разработке лекарственных препаратов, и потребность в таких терапевтических антителах продолжает расти. Такие антитела и другие классы функциональных связывающих молекул также обладают потенциалом для контроля поведения клеток *in vivo* и *in vitro*. Однако способность контролировать и управлять клеточным поведением зависит от доступности

природных лигандов, которые контролируют специфические сигнальные пути. К сожалению, многие природные лиганды, как, например, те, что контролируют дифференцировку стволовых клеток (например, представители суперсемейства FGF, TGF-бета, Wnt и Notch) часто характеризуются нестабильными взаимодействиями и ограниченной доступностью, обусловленными характером их низкой экспрессии и стабильности. Обладая исключительной специфичностью, антитела обладают большим потенциалом для контроля клеточного поведения.

Идентификация функциональных антител, которые способны модифицировать передачу сигнала в клетках, в историческом плане была относительно трудоемкой, включая выделение клонов, экспрессию антител, характеристику в соответствии с последовательностью и связывающими свойствами, преобразование в системы экспрессии млекопитающих и оценку с использованием функциональных анализов в клеточных системах. Подход, основанный на дисплее на поверхности эукариотических клеток, описанный в данном документе, способствует уменьшению таких затрат, но все еще существует потребность в получении антитела и добавлении его в отдельную репортерную культуру клеток. Следовательно, предпочтительной альтернативой может быть непосредственный скрининг библиотек связывающих молекул, экспрессированных в эукариотических клетках, в отношении влияния связывания на передачу сигнала в клетках или поведение клеток путем использования продуцирующей клетки как таковой в качестве репортерной клетки. После введения генов антител могут быть идентифицированы клоны в составе полученной в результате популяции клеток, характеризующихся изменением экспрессии репортерного гена или измененными фенотипами.

В целом ряде недавних публикаций было описано создание библиотек антител путем клонирования наборов генов антител в репортерные клетки [47, 105, 106]. Эти системы характеризуются сочетанием экспрессии и репортеринга в пределах одной клетки и обычно введением популяции антител, отобранных против заранее определенной мишени (например, с использованием фагового дисплея).

Популяция генов антител может быть введена в репортерные клетки для получения библиотеки с помощью способов, описанных в данном документе, а также могут быть идентифицированы клоны в пределах популяции с фенотипом, измененным с помощью антител (например, измененной экспрессией генов или жизнеспособностью). Для того, чтобы этот отбор, основанный на фенотипе, функционировал, требуется сохранение связи между геном антитела, содержащемся в экспрессирующей клетке (генотип), и следствием экспрессии антитела (фенотип). Ранее это было достигнуто либо посредством закрепления антител к клеточной поверхности [47], как описано в случае дисплея антител, или

посредством применения полужидкой среды для удержания секретируемых антител вблизи продуцирующих их клеток [105]. В качестве альтернативы антитела и другие связывающие молекулы могут удерживаться внутри клетки [107]. Связывающие молекулы, удержанные на клеточной поверхности или в окружающей среде, могут взаимодействовать с эндогенным или экзогенным рецептором на клеточной поверхности, вызывая активацию рецептора. Это в свою очередь может вызвать изменение экспрессии репортерного гена или изменение фенотипа клетки. В качестве альтернативы антитело может блокировать рецептор или лиганд для уменьшения активации рецептора. Ген, кодирующий связывающую молекулу, которая обуславливает изменение поведения клеток, затем может быть выделен для получения или дальнейшего конструирования.

Альтернативой этому «подходу, основанному на применении мишени» является возможность введения «наивной» популяции антител, которая не подвергалась предварительному отбору для выявления определенной мишени [108]. Для идентификации отдельных клеток популяции с измененным поведением используется репортерная система. Поскольку предварительная информация о мишени отсутствует, для реализации этого подхода, который не основан на применении мишени, в особенности потребуется использование большого набора антител, поскольку предварительное обогащение популяции антител к мишени не представляется возможным. Эффективность данного подхода будет повышена за счет использования направляемого нуклеазой встраивания трансгенов, как описано в настоящем изобретении.

Подход «функциональный отбор» может быть использован для других целей, включающих применение библиотек, полученных в эукариотических клетках, в частности, высших эукариот, как, например, клетках млекопитающих. Может быть произведено слияние антитела с сигнальным доменом таким образом, что связывание с мишенью будет вызывать активацию рецептора. Kawahara *et al.* сконструировали рекомбинантные рецепторы, в которых флуоресцеин, целенаправленно связывающийся с внеклеточным scFv, был слит со спейсерным доменом (домен D2 рецептора Epo) и различными внутриклеточными доменами рецепторов к цитокинам, включая рецептор тромбопоэтина (Tpo), рецептор эритропоэтина (Epo), gp130, рецептор ИЛ-2 и рецептор EGF [109, 110, 111]. Они были введены в ИЛ-3-зависимую линию про-В-клеток (BaF3) [27], при этом было показано, что рекомбинантные рецепторы осуществляют антиген-зависимую активацию рекомбинантного рецептора, что приводит к ИЛ-3-независимому росту. Этот же подход был использован в модельных экспериментах, чтобы продемонстрировать опосредованную антигеном хемотаксию клеток BaF3 [110]. Область применения данного подхода была расширена со стабильных культур клеток на клетки первичной культуры, примерами чего

являются жизнеспособность и рост гемопоэтических стволовых клеток, отвечающих на стимуляцию Т<sub>ро</sub> [112], или ИЛ-2 зависимой первичной культуры Т-клеток, при этом вместо стимуляции природными Т<sub>ро</sub> и ИЛ-2, соответственно, была использована стимуляция направляемых флуоресцеином рекомбинантных рецепторов scFv. Таким образом, система на основе рекомбинантных антител и рецепторов может быть использована для активации зависимой от мишени экспрессии генов или фенотипических изменений в клетках первичной культуры или стабильных репортерных клетках. Эта способность может быть использована для идентификации слитых связывающих молекул, которые активируют передачу сигналов, или связывающих молекул, которые ее подавляют.

При модификации вышеприведенного подхода отдельные VH- и VL-домены из антитела к лизоциму были слиты с внутриклеточным доменом [113]. В ответ на добавление лизоцима отмечался клеточный рост, свидетельствуя о том, что антиген индуцировал димеризацию или стабилизацию отдельных партнеров VH и VL по слиянию. Таким образом, для оптимального ответа в этой системе необходимо взаимодействие трех взаимодействующих компонентов.

Несмотря на то, что вышеприведенные способы описаны в данном документе применительно к молекулам антител, они также могут быть адаптированы и осуществлены с использованием библиотек других связывающих молекул.

Комплементация фрагментов белка представляет собой альтернативную систему для изучения и отбора белок-белковых взаимодействий на клетках млекопитающих [114, 115]. Сюда относится восстановление функции дополнительных репортерных белков посредством белок-белковых взаимодействий. Репортерные белки, которые были использованы, включают убиквитин, DNAE-интеин, бета-галактозидазу, дигидрофолатредуктазу, GFP, люциферазу светлячка, бета-лактамазу, протеазу TEV. Например, недавним примером данного подхода является использование мембранного гибрида 2 (MaMTH) млекопитающих, когда ассоциация белка-приманки с дополнительным убиквитином и слитым белком - фактором транскрипции с белком-партнером и дополнительным убиквитином приводит к восстановлению распознавания убиквитина и высвобождению фактора транскрипции для обеспечения экспрессии репортерного гена [116]. Опять-таки вследствие нарушения передачи сигнала могут быть выявлены связывающие молекулы, которые препятствуют такому взаимодействию или усиливают его.

#### *Выделение и изменение формы связывающих молекул и кодирующей ДНК*

После отбора представляющей интерес связывающей молекулы или клона из библиотеки обычно на следующем этапе выделяют (например, идентифицируют или

амплифицируют) ДНК, кодирующую связывающую молекулу. Необязательно может понадобиться модификация нуклеиновой кислоты, кодирующей связывающую молекулу, например, для изменения структуры связывающей молекулы и/или вставки кодирующей последовательности в другой вектор.

Если в качестве связывающей молекулы используется молекула антитела, способ может включать выделение ДНК, кодирующей молекулу антитела из клеток клона, амплификацию ДНК, кодирующей по меньшей мере одну переменную область антитела, предпочтительно и VH-, и VL-доменов, и вставку ДНК в вектор для получения вектора, кодирующего молекулу антитела. Молекула мультимерного антитела, несущая константный домен, может быть преобразована в молекулу одноцепочечного антитела для экспрессии в растворимой секретируемой форме.

Антитела могут быть представлены в разных форматах, но какой бы ни был выбран формат антитела, после выделения гена антитела возникает возможность изменения его конфигурации на целый ряд разных форматов. После выделения VH- или VL-доменов их можно повторно клонировать в экспрессирующие векторы, содержащие требуемые домены-партнеры (см. Пример 1, в котором показана экспрессирующая кассета с двойным промотором IgG).

Этап изменения формата может включать изменение формы связывающих молекул, состоящих из пары субъединиц (например, молекулы scFv), на другой молекулярный формат связывающей молекулы (например, Ig или Fab), в котором сохраняется изначальное попарное объединение субъединиц. Такие способы описаны более подробно в других разделах в данном документе и могут быть применены к клонам, продуцирующим моно-, олиго- и поликлональные антитела. Способ может быть использован для преобразования всей получаемой на выходе популяции «*en masse*» с помощью любых широко используемых технологий дисплея, включая фаговый, дрожжевой или рибосомный дисплей.

*Дисплей молекул scFv, слитых с Fc-доменами, на поверхности клеток млекопитающих.*

Хотя многие библиотеки антител, полученные с помощью фагового дисплея, подвергаются форматированию для дисплея scFv, эукариотические системы дисплея обеспечивают презентацию в формате Fab или IgG. Чтобы полностью использовать потенциал для экспрессии IgG/Fab, особенно при использовании scFv из других систем дисплея, понадобится использование отобранных, связанных VH- и VL-доменов в бактериальной экспрессирующей системе и их экспрессия в эукариотической системе, слитой с соответствующими константными доменами. В данном документе описан способ

преобразования популяций scFv в иммуноглобулин (Ig) или фрагмент, антигенсвязывающий фрагмент (Fab), таким образом, чтобы сохранялось изначальное попарное объединение VH- и VL-цепей. В настоящем изобретении преобразование становится возможным вследствие использования отдельных клонов, смесей олигоклональных антител или целых популяций, получивших формат scFv, с сохранением изначального попарного объединения VH- и VL-цепей. Способ осуществляется с использованием промежуточной «миникольцевой ДНК», полученной нерепликативным путем, с которой вносится новый «спейсерный» фрагмент ДНК. Кольцевая ДНК подвергается линейаризации (например, путем расщепления рестриктазами или ПЦР), что обуславливает изменение относительного положения изначальных VH- и VL- фрагментов и размещение «спейсерной» ДНК между ними. После линейаризации продукт может быть клонирован в предпочтительный вектор, например, вектор для экспрессии на клетках млекопитающих. Таким способом могут быть заменены все элементы, кроме VH и VL. Элементы для экспрессии в клетках бактерий могут быть заменены на элементы для экспрессии на клетках млекопитающих и слияния с альтернативными партнерами. Для завершения процесса преобразование потребуются только один этап трансформации бактерий *E. coli* для создания популяций бактериальных колоний, каждая из которых несет плазмиду, кодирующую уникальное рекомбинантное антитело с форматом Ig или Fab. Выходя за рамки преобразования scFv в IgG/Fab, способ может быть использован для изменения формы любого из двух связанных элементов ДНК для клонирования в вектор, таким образом, что после изменения формы каждый ДНК элемент попадает в окружение разных контрольных элементов ДНК с сохранением изначального попарного объединения. Был описан ранее разработанный способ, в котором используются 2 последующих этапа клонирования [117] для замены этих элементов в отличие от настоящего способа, который осуществляется с использованием промежуточного кольцевого промежуточного продукта, полученного нерепликативным путем.

Способ изменения структуры связывающей молекулы или популяции связывающих молекул может включать преобразование scFv в Ig или его фрагмент, например, Fab. Способ может включать преобразование нуклеиновой кислоты, кодирующей scFv в ДНК, кодирующую иммуноглобулин (Ig) или его фрагмент, как, например, формат Fab, таким образом, чтобы сохранить изначальное попарное объединение тяжелой (VH) и легкой (VL) цепей переменного домена. Предпочтительно преобразование осуществляется с использованием кольцевой промежуточной ДНК, которая может представлять собой «миникольцевую» ДНК, полученную нерепликативным путем. Для реализации способа

требуется одна трансформация *E. coli* для непосредственного получения бактериальных трансформантов, несущих плазмиды, кодирующие ДНК Ig или Fab.

Способ может быть применен к клонам, продуцирующим моно-, олиго- и поликлональные антитела. Способ может быть использован для преобразования всей получаемой на выходе популяции «*en masse*» с помощью любых широко используемых технологий дисплея, включая фаговый, дрожжевой или рибосомный дисплей.

В более широком смысле этот аспект настоящего изобретения относится к способу, который обеспечивает изменение формы любых двух связанных элементов ДНК на вектор, в котором элементы ДНК клонированы под контролем отдельных промоторов или разделены альтернативными контрольными элементами, но сохраняют изначальное попарное объединение цепей ДНК.

В примерах 7 и 14 такие способы описаны более подробно.

За выделением и опциональным изменением структуры ДНК, кодирующей связывающие молекулы, может следовать введение этой ДНК в другие клетки для создания производной библиотеки, как описано в других разделах в данном документе, или ДНК, кодирующая одну или более определенных представляющих интерес связывающих молекул, может быть введена в клетку-хозяина для экспрессии. Тип клеток-хозяев может отличаться от клеток библиотеки, из которой они были получены. В целом, ДНК будет получена в виде вектора. ДНК, введенная в клетку-хозяина, может встраиваться в клеточную ДНК клетки-хозяина. Затем могут быть отобраны клетки-хозяева, экспрессирующие молекулу секретируемого растворимого антитела.

Клетки-хозяева, кодирующие одну или более связывающих молекул, могут быть получены в культуральной среде и культивированы для экспрессии одной или более связывающих молекул.

#### *Производные библиотеки*

После получения библиотеки с использованием способа по настоящему изобретению один или более клонов библиотеки могут быть отобраны и использованы для получения последующей библиотеки второго поколения. Если библиотека была составлена путем введения ДНК в эукариотические клетки, как описано в данном документе, клетки библиотеки могут быть культивированы для экспрессии связывающих молекул, и один или более клонов, экспрессирующих представляющие интерес связывающие молекулы, могут быть выделены, например, путем отбора связывающих молекул против мишени, как описано в других разделах в данном документе. Эти клоны впоследствии могут быть использованы для создания производной библиотеки, содержащей ДНК, кодирующую второй набор связывающих молекул.

Для создания производной библиотеки в донорную ДНК из одного или более выделенных клонов вносят мутации для получения второго набора связывающих молекул. Мутации могут представлять собой вставку, замену или делецию одного или более нуклеотидов. Если связывающая молекула представляет собой полипептид, мутация будет заключаться в изменения последовательности кодируемой связывающей молекулы путем вставки, замены или делеции одной или более аминокислот. Мутация может затрагивать одну или более областей, как, например, одну или более CDR молекулы антитела, обеспечивая получение набора связывающих молекул, формирующих класс молекул с общими структурными характеристиками, которые отличаются по одному или более участкам, определяющим разнообразие, как описано в других разделах в данном документе.

Создание производной библиотеки может включать выделение донорной ДНК из одного или более выделенных клонов, введение мутации в ДНК для получения производной популяции донорных молекул ДНК, кодирующих второй набор связывающих молекул, и введение производной популяции донорных молекул ДНК в клетки для создания производной библиотеки клеток, содержащих ДНК, кодирующую второй набор связывающих молекул.

Выделение донорной ДНК может включать получение и/или идентификацию ДНК клона. Такие способы могут охватывать амплификацию ДНК, кодирующей связывающую молекулу из выделенного клона, например, путем ПЦР и введение мутаций. ДНК может быть секвенирована, а мутантная ДНК синтезирована.

В качестве альтернативы в донорную ДНК может быть введена мутация в одном или более выделенных клонов путем индукции мутации ДНК в этих клонах. Таким образом, может быть создана производная библиотека из одного или более клонов, не требующая выделения ДНК, например, посредством эндогенной мутации в клетках DT40 птиц.

Дисплей антител особенно хорошо подходит для создания производных библиотек. Как только гены антител выделены, появляется возможность применения множества подходов к мутагенезу (например, ПЦР сниженной точности, олигонуклеотид-направленный мутагенез, перетасовка цепей) с образованием дисплейных библиотек родственных клонов, из которых могут быть отобраны варианты с улучшенными свойствами. Например, при использовании перетасовки цепей ДНК, кодирующих популяцию выбранного клона VH, смесь или популяция олигоклональных иммуноглобулинов могут быть субклонированы в вектор, кодирующий антитело подходящей формы и кодирующий набор VL-цепей, получивших подходящий формат [118]. В качестве альтернативы и опять-таки с использованием примера VH-цепей, клон,



продуцирующий VH, смесь или популяция олигоклональных иммуноглобулинов могут быть введены в популяцию эукариотических клеток, которые кодируют и экспрессируют популяцию партнеров легкой цепи, получивших соответствующую форму (например, VL-CL-цепь для связи с тяжелой цепью с измененной формой на IgG или Fab). Популяция клонов, продуцирующих VH-цепи, может быть получена из любого из источников, которые обсуждались выше, включая В-клетки иммунизированных животных или гены scFv из отобранных популяций фагов. В последнем примере клонирование отобранных генов VH-цепей в вектор, содержащий набор генов легких цепей, позволяет объединить перетасовку цепей с изменением формата (например, на формат IgG) за один этап.

Особым преимуществом дисплея на эукариотических клетках является возможность контроля точности этапа отбора/скрининга. Путем уменьшения концентрации антигена клетки, экспрессирующие связывающие молекулы с наибольшей аффинностью, можно отличить от клонов с меньшей аффинностью в пределах популяции. Визуализация и количественная оценка процесса созревания с использованием проточной цитометрии является основным преимуществом дисплея на поверхности эукариотических клеток, поскольку она дает представление о доле позитивных клонов в библиотеке наивных клеток и дает возможность прямого сравнения аффинности отобранных клонов и родительской популяции на протяжении сортировки. После сортировки аффинность отдельных клонов можно определить путем предварительной инкубации с целым рядом концентраций антигенов и анализа с помощью проточной цитометрии или теста с помощью гомогенной флуоресценции с временным разрешением (TRF), или же с использованием поверхностного плазмонного резонанса (SPR) (Biacore).

#### *Характеристики и форма библиотеки*

Настоящее изобретение позволяет создавать библиотеки эукариотических клеток, обладающих многими предпочтительными характеристиками. Настоящее изобретение относится к библиотекам, обладающим любым одним или более из следующих признаков:

**Разнообразие.** Библиотека может кодировать и/или экспрессировать по меньшей мере  $10^2$ ,  $10^3$ ,  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ ,  $10^8$  или  $10^9$  разных связывающих молекул.

**Равномерное встраивание.** Библиотека может состоять из клонов, содержащих донорную ДНК, встроенную в фиксированный локус, или в ограниченное количество фиксированных локусов в клеточной ДНК. Следовательно, каждый клон в библиотеке содержит донорную ДНК в фиксированном локусе или по меньшей мере в одном из фиксированных локусов. Предпочтительно, чтобы клоны содержали донорную ДНК, встроенную в один или два фиксированных локуса в клеточной ДНК. Согласно пояснениям, приведенным в других разделах в данном документе, в последовательности,

распознаваемой сайт-специфической нуклеазой, находится сайт встраивания. Встраивание донорной ДНК для получения рекомбинантной ДНК подробно описано в других разделах в данном документе и может приводить к разным результатам, в зависимости от количества сайтов встраивания. Если в клетках, используемых для создания библиотеки, имеется один потенциальный сайт встраивания, библиотека будет представлять собой библиотеку клонов, содержащих донорную ДНК, встроенную в один фиксированный локус. Следовательно, все клоны библиотеки будут содержать гены связывающих молекул в одинаковом положении в клеточной ДНК. В качестве альтернативы, если имеется много потенциальных сайтов встраивания, библиотека может представлять собой библиотеку клонов, содержащих донорную ДНК, встроенную во многие или разные фиксированные локусы. Предпочтительно, чтобы каждый клон библиотеки содержал донорную ДНК, встроенную в первый и/или второй фиксированный локус. Например, библиотека может включать клоны, в которых донорная ДНК встраивается в первый фиксированный локус, клоны, в которых донорная ДНК встраивается во второй фиксированный локус, и клоны, в которых донорная ДНК встраивается как в первый, так и второй фиксированный локусы. В предпочтительных вариантах реализации имеется только один или два фиксированных локуса в клонах в библиотеке, хотя существует возможность встраивания донорной ДНК во многие локусы, если это потребуется для применения в определенных целях. Следовательно, в некоторых библиотеках каждый клон может содержать донорную ДНК, встроенную в любой один или более из нескольких фиксированных локусов, например, три, четыре, пять или шесть фиксированных локусов.

В случае библиотек, содержащих субъединицы связывающих молекул, встроенные в отдельные сайты, клоны библиотеки могут содержать ДНК, кодирующую первую субъединицу связывающей молекулы, встроенную в первый фиксированный локус, и ДНК, кодирующую вторую субъединицу связывающей молекулы, встроенную во второй фиксированный локус, при этом клоны экспрессируют мультимерные связывающие молекулы, содержащие первую и вторую субъединицы.

Равномерная транскрипция. Относительные уровни транскрипции связывающих молекул в разных клонах библиотеки поддерживаются в ограниченных пределах, поскольку донорная ДНК встраивается в ограниченное количество локусов и в одинаковый локус в разных клонах (фиксированный локус). Относительно равномерная транскрипция генов связывающих молекул обуславливает сравнимые уровни экспрессии связывающих молекул клонами из библиотек. Связывающие молекулы, представленные на поверхности клеток библиотеки, могут быть идентичными (имеют одинаковую аминокислотную последовательность) другим связывающим молекулам, представленным на той же клетке.

Библиотека может состоять из клонов клеток, каждая из которых несет одну молекулу из набора связывающих молекул, или клонов, несущих много молекул из набора связывающих молекул на клетку. В качестве альтернативы библиотека может включать некоторые клоны, которые несут одну молекулу из набора связывающих молекул, и некоторые клоны, которые несут много молекул (например, две) набора связывающих молекул. Предпочтительно клоны библиотеки экспрессируют одну или две молекулы из набора связывающих молекул.

Например, библиотека клонов эукариотических клеток согласно настоящему изобретению может экспрессировать набор, состоящий по меньшей мере из  $10^3$ ,  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ ,  $10^8$  или  $10^9$  разных связывающих молекул, например, IgG, Fab, scFv или scFv-Fc-фрагменты антител, причем каждая клетка содержит донорную ДНК, встроенную в фиксированный локус в клеточной ДНК. Донорная ДНК кодирует связывающую молекулу и может дополнительно содержать генетический элемент для отбора клеток, в которые донорная ДНК встраивается в фиксированном локусе. Клетки библиотеки могут содержать ДНК, кодирующую экзогенную сайт-специфическую нуклеазу.

Эти и другие признаки библиотек согласно настоящему изобретению дополнительно описаны в других разделах в данном документе.

Настоящее изобретение касается библиотеки либо в чистом виде, в виде популяции клонов библиотеки в отсутствие других эукариотических клеток, либо в смеси с другими эукариотическими клетками. К другим клеткам могут относиться эукариотические клетки того же самого типа (например, той же самой клеточной линии) или разные клетки. Дополнительные преимущества могут быть получены путем комбинации двух или более библиотек согласно настоящему изобретению или комбинации библиотеки согласно настоящему изобретению со второй библиотекой, или второй популяцией клеток, либо для облегчения или расширения скрининга, или для применения в других целях, как описано в данном документе, или очевидно для специалиста.

Библиотека согласно настоящему изобретению, один или более клонов, полученных из библиотеки, или клетки-хозяева, в которые была введена ДНК, кодирующая связывающую молекулу из библиотеки, могут быть получены в культуральной среде. Клетки могут быть культивированы и затем концентрированы с целью получения клеточного осадка для удобства транспортировки и хранения.

Библиотеки обычно получают *in vitro*. Библиотека может содержаться в емкости, такой как культуральный флакон, содержащий клетки библиотеки, суспендированные в культуральной среде, или емкости, содержащей осадок или концентрированную суспензию

эукариотических клеток, составляющих библиотеку. Библиотека может составлять по меньшей мере 75 %, 80 %, 85 % или 90 % эукариотических клеток в указанной емкости.

## КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

Более подробное описание вариантов реализации настоящего изобретения представлено ниже со ссылкой на прилагаемые графические материалы, которые описаны ниже:

Фигура 1. Вектор для экспрессии антител, получивших формат IgG

- a. pDUAL D1.3, экспрессирующий вектор с двойным промотором для секреции IgG (в остове вектора «pCMV/мус/ER»)
- b. pINT3-D1.3, экспрессирующий вектор с двойным промотором для секреции IgG (в остове вектора «pSF-CMV-f1-Pac1)
- c. Остов вектора pCMV/мус/ER. Сайт ECoR1 предшествует промотору CMV. Сайты BstB1 и BstZ171 фланкируют сайты поли(A)-последовательности SV40.
- d. Остов вектора pSF-CMV-f1-Pac1 (Oxford Genetics)
- e. Синтетическими ген с экзоном, кодирующим трансмембранный участок (TM) PDGFR и экзон, отвечающий за секрецию (sec). Сплошными стрелками представлены сайты рекомбинации Rox.

Фигура 2. Последовательность pD1 (SEQ ID NO: 1, 2, 3, 4 и 5): кассета с двойным промотором для экспрессии антител на поверхности.

Признаки:

Промотор pEF 13-1180

Лидерная последовательность BM40 1193-1249

Гуманизированный D1.3 VL 1250-1578

С-каппа человека 1577-1891BGH с поли(A)-последовательностью 1916-2130

Промотор CMV 2146-2734

Лидерная последовательность VH мыши с интроном 32832-3414

Гуманизированный D1.3 VH 3419-3769

Оптимизированные CH1-CH3 IgG2 человека

Фигура 3. Конструирование донорной плазмиды AAVS (pD2)

- a. Изображение локуса AAVS человека. Экзон 1 и Экзон 2 локуса AAVS (кодирующего протеинфосфатазу 1, регуляторную субъединицу 12C, PPP1R12C) разделены интроном длиной 4428 п.о. Сплайсинг осуществляется в рамке считывания «0», т.е. сплайсинг происходит между 2 интактными кодонами от каждого экзона. Для

расщепления данного интрона доступны конструкции TALEN и CRISPR/Cas9. Заштрихованными блоками ниже линии представлены области, расположенные с левой и правой стороны этого сайт расщепления, которые используются в составе векторных конструкций для осуществления гомологичной рекомбинации в этом локусе (левое гомологичный участок AAVS «левое HA» и правое гомологичный участок AAVS («правое HA»)).

б. Изображение антитела, кодирующего донорную плазмиду pD2. Левые и правые гомологичные участки показаны на концах изображения конструкции. Синтетическому гену устойчивости к бластицидину предшествует акцепторный сайт сплайсинга, с помощью которого обеспечивается слияние «в рамке считывания» с экзоном 1 AAVS. Также показана кассета для экспрессии антител, состоящая из генов легкой цепи D1.3 и тяжелой цепи D1.3 IgG2, находящихся под контролем промоторов pEF и CMV, соответственно.

с. Последовательность донорной конструкция pD1-huD1.3 (SEQ ID NO: 6, 7 и 8). Гомологичные участки AAVS показаны подчеркнуты и выделены полужирным шрифтом. Для краткости кассета с генами антител (ранее показанная на Фигуре 2) не показана в деталях. Сайты рестрикции, использованные при клонировании, показаны полужирным шрифтом. Последовательность остова плазмиды показана включительно до гена устойчивости к ампициллину.

#### Фигура 4. Экспрессия IgG на клеточной поверхности

а, с. Для выявления жизнеспособных клеток использовали малоугловое рассеяние, а канал FL3 использовали для анализа окрашивания в (а, с). Положительно окрашенные клетки, обнаруженные с использованием канала FL3 (представляющие нежизнеспособные клетки, которые поглотили 7-AAD), были исключены из анализа. Все клетки были трансфицированы pD2-D1.3 в отсутствие (а, b) или присутствие (с, d) AAVS TALEN и были окрашены антителом Fc-фрагменту.

Фигура 5. Связывание антигена антителом на клеточной поверхности. Жизнеспособные клетки были отобраны, исходя из данных малоуглового рассеяния и поглощения 7AAD (а). Клетки инкубировали с меченым флуоресцентной меткой лизоцимом куриного яйца (b).

Фигура 6. Влияние осуществляемого TALEN расщепления геномной ДНК на встраивание гена устойчивости к бластицидину. На Фигуре показано количество колоний, полученных в описанных условиях.

Фигура 7. Анализ встраивания донорной плазмиды pD2-D1.3 в локус AAVS. После трансфекции осуществляли селекцию клеток с использованием бластицидина и получали

геномную ДНК. В образцах 1-9 отмечался положительный результат, обусловленный вставкой, осуществляемой нуклеазой TALE, направляемой AAVS, образцы 10-11 были выделены из клонов, полученных из клетки, трансфицированной в отсутствие нуклеазы TALE. Анализ геномной ДНК осуществляли с помощью ПЦР, как описано в тексте. а. Верификация от 5'-конца сайта встраивания. В. Верификация от 3'-конца сайта встраивания.

Фигура 8. Конструкция вектора pD6, экспрессирующего scFv-Fc

а. Антитела, получившие форму scFv, клонируют в сайты Nco1/Not1 для обеспечения слияния «в рамке считывания» с Fc-областью IgG2 человека.

б. Последовательность pD6 от сайта Nco1 до Pme1 (SEQ ID NO: 9, 10 и 11).

Фигура 9. Отбор связывающих молекул из библиотеки scFv-Fc на клеточной поверхности (из отобранных популяций фагов). Показан анализ клеток с помощью проточной цитометрии;

а-с. Объединенная популяция молекул sFv-Fc, специфичных к CD229, полученных в результате 2 раундов отбора с помощью фагового дисплея с использованием CD229

д, ф. Объединенная популяция молекул sFv-Fc, специфичных к  $\beta$ -галактозидазе, полученных в результате 1 раунда отбора с помощью фагового дисплея с использованием  $\beta$ -галактозидазы (клетки  $\beta$ -galR1)

е. Объединенная популяция молекул sFv-Fc, специфичных к  $\beta$ -галактозидазе, полученных в результате 2 раундов отбора с помощью фагового дисплея с использованием  $\beta$ -галактозидазы

Образец (а) представляет неокрашенные клетки, а остальные клетки были окрашены меченым фикоэритрином антителом к Fc-фрагменту иммуноглобулина человека (в FL2) и 100 нМ соответствующего биотинилированного антигена/меченого FITC стрептавидина (в FL1). Клетки анализировали через 13 суток (а, б, д, е). В примерах с и ф показаны клетки, окрашенные через 20 суток, а в выделенной области показаны клетки, отобранные с помощью проточной цитометрии

г. Клетки линии  $\beta$ -galR1, отобранные с помощью проточной цитометрии (Фиг. 6 ф), выращивали в течение 22 суток и повторно анализировали в отношении экспрессии scFv-Fc, а также связывания антигенов (с использованием 100 нМ антигена). На Фиг. 6 г показан неокрашенный эквивалент.

На Фиг. 6 ж показаны неотсортированные клетки линии  $\beta$ -galR1 из первоначальной популяции (как на д), которые выращивали в течение 42 суток после трансфекции (j). Немеченные клетки каждой популяции показаны для сравнения (г, и)

Фигура 10. Дисплей на клетках млекопитающих и сортировка библиотеки антител, получивших формат IgG.

Проводили селекцию популяции антител с использованием  $\beta$ -галактозидазы и 1 или 2 раундов фагового дисплея, изменяли их форму на IgG и подвергали целенаправленному встраиванию с помощью нуклеазы в локус AAVS клеток НЕК293. На панелях a, b показаны клетки, полученные из популяции фагов раунда 1 либо a, неотсортированные (после выращивания в течение 38 суток), либо b отсортированные с помощью проточной цитометрии и выращенные в течение 19 суток. На панелях c, d показаны клетки, полученные популяции фагов раунда 2 либо c, неотсортированные (после выращивания в течение 38 суток, либо d, отсортированные и выращенные в течение 19 суток.

Фигура 11. Создание больших библиотек наивных scFv-Fc и отбор связывающих молекул.

Клетки из библиотеки наивных scFv-Fc были окрашены 500 нМ биотинилированного антигена и стрептавидинп-FITC наряду с меченым фикоэритрином антителом к Fc-фрагменту, как и раньше. В выделенных областях показаны клетки, которые были отобраны с помощью проточной сортировки. Образцы были мечены биотинилированными:

- a. CD28
- b.  $\beta$ -галактозидазой
- c. Тиреоглобулином
- d. EphB4

Фигура 12. нацеливающий вектор для введения интрона, содержащего «сайт множественной посадки» для сравнения способов встраивания.

- a. Промежуточная GFP-экспрессирующая плаزمид (pD3)
- b. Направляемый AAVS1 нацеливающий вектор (pD4). «Сайт посадки» содержит элементы, контролирующие встраивание, а к которым относятся сайты FRT, lox 2272, мегануклеазы I-Sce1 и GFP TALEN

После встраивания pD4 в локус AAVS в составе геномной ДНК появляется множество сайтов рекомбинации или расщепления нуклеазой. Вносимая плазмид pD5 (Фиг. 15) содержит левые и правые гомологичные участки, эквивалентные последовательности, содержащейся с каждой стороны «сайта посадки» для осуществления вставки гена антитела путем гомологичной рекомбинации.

Фигура 13. Последовательность pD4 (SEQ ID NO: 12, 13, 14, 15, 16, 17 и 18.). Компоненты последовательности включают: левый гомологичный участок AAVS 19-822 сайт FRT 832-879, сайт Lox 2272 884-917, сайт мегануклеазы I-Sce1 933-950 левый сайт связывания GFP-TALEN 954-968, GFP правый сайт связывания GFP-TALEN 984-997

T2A 1041-1103

GFP 1104-1949

Промотор PGK 2178-2691

Пурамицин-дельта-тимидинкиназа 2706-4307

LoxP 4634-4667

Правый гомологичный участок AAVS 4692-5528

Фигура 14. Верификация встраивания интрона с «сайтами множественной посадки» в клон 6F. После трансфекции р пурамицина и получали геномную ДНК. Образцы под номером 1 представляют собой суммарную популяцию отобранных молекул. Образец 2 представляет собой клон 6F, образец 3 – клон, трансфецированный в отсутствие TALEN, и образец 4 – клетки НЕК293 дикого типа. Праймеры и условия, описанные в тексте, были использованы для верификации путем ПЦР точности встраивания на 5'- и 3'- концах геномной ДНК. Основная полоса (надлежащего размера) наблюдается для отобранного клона (6F), а также отобранной популяции.

Фигура 15. Последовательность донорной плазмиды для встраивания в сайты Flp/GFP TALEN (pD5) (SEQ ID NO: 19, 20 и 21). Компоненты последовательности включают:

AAVS HA 13-233

сайт FRT 243-290

Lox 2272 295-328

I-SceI 344-361, ген устойчивости к бластицидину 417-818, поли(A)-последовательность 832-1070

Фигура 16. Последовательность конструкции мегануклеазы I-SceI (SEQ ID NO: 22 и 23)

Фигура 17. Сравнительный анализ с помощью проточной цитометрии направляемого нуклеазой встраивания с использованием мегануклеазы I-SceI и рекомбиназ.

Клетки клона 6F контрансфецировали pD5-D1.3 и плазмидами, кодирующими указанную нуклеазу/рекомбиназу. Проводили селекцию клеток с использованием бластицидина и анализировали 3 суток после трансфекции с использованием биотинилированных антител к Fc-фрагменту иммуноглобулина человека и стрептавидина-фикоэритрина. Указано процентное содержание положительно окрашенных клеток (также эти результаты представлены в краткой форме в Таблице 5)

а. Нетрансфецированные, b только донорная плазида с I-SceI, d, eGFP TALEN, e. Cre, f. рекомбиназа Flp (кодируемая плазмидой pOG44).



Фигура 18. Направляемое нуклеазой встраивание запускает гомологичную рекомбинацию и «негомологичное соединение концов» (NHEJ).

a. Изображение структуры плазмиды pD5, использованной для целенаправленного встраивания «сайта множественной посадки» в составе интрона клеток клона 6F, на котором показано положение праймера J48. «Сайт посадки» в этой плазмиде включает в себя сайт FRT, сайт lox2272 и сайт мегануклеазы I-Sce1 (сайт GFP TALEN отсутствует).

b. Изображение сайта встраивания в составе клона 6F (получен из pD4), на котором показано положение праймера J44. «Сайт посадки» включает в себя элементы для направления встраивания, а именно FRT, lox 2272, мегануклеазу I-Sce1 и GFP TALEN.

c. Изображение сайт встраивания клона 6F после гомологичной рекомбинации pD5, на котором показано положение праймеров J44 и J48.

d. Изображение сайта встраивания клона 6F после NHEJ или F1p-рекомбинации pD5, на котором показано положение праймеров J44, J46 и обратного праймера J44. Двусторонней стрелкой обозначена «дополнительная» плазмидная ДНК, включенная в состав в результате NHEJ или F1p-направляемого встраивания. Следует отметить, что в этом примере вносимая плаزمидная ДНК (pD5) содержит гомологичные участки (которые направляют гомологичную рекомбинацию, но не нужны для NHEJ). Эти последовательности сохраняются после встраивания с помощью NHEJ, что обуславливает дубликацию последовательности, подставленной в составе гомологичных участков, при этом одна пара в данном случае происходит из плазмиды, а другая пара представляет собой эндогенные геномные последовательности. Для простоты гомологичные участки, кодируемые плазмидой, не показаны, показаны только эквивалентные им последовательности, содержащиеся в геноме.

e. Праймеры 44 и 48 были использованы в качестве ПЦР-праймеров для образцов i-iv, когда была использована геномная ДНК из клеток, трансфицированных генами следующих нуклеаз/интеграз:

i. Sce1, ii. TALEN (GFP), iii. F1p (pOG44), iv. Только донорная плазмидная. В качестве маркеров молекулярной массы были использованы GeneRuler 1kb ladder (New England Biolabs). С помощью праймеров J44 и J48 было установлено, что в расщепленных нуклеазой образцах i и ii i и ii произошла гомологичная рекомбинация с образованием полосы длиной 1928 п.о. (указана стрелкой).

Праймеры 44 и 46 были использованы для образцов v-viii, когда была использована геномная ДНК из следующих образцов. v. Sce1, vi. TALEN (GFP), vii. F1p (pOG44), viii. Только донорная плазмидная. С помощью праймеров J44 и J46 было установлено, что

расщепление донорной и геномной ДНК мегануклеазой I-SceI приводило к NHEJ (образец v.) с образованием полосы длиной 1800 п.о. (указана стрелкой). Как и предполагалось, в результате встраивания, опосредованного Flp, была получена полоса аналогичного размера (vii). NHEJ не происходило при использовании GFP TALEN, поскольку во вносимой плазмиде отсутствовал сайт расщепления.

Фигура 19. Секреция антитела IgG в культуральный супернатант.

а. Окрашенный Кумасси гель с IgG, очищенным с использованием белка А, который был получен из культуральных супернатантов.

i. IgG, выделенный из супернатанта трансфицированных pD2-D1.3 клеток в отсутствие трансфекции геном рекомбиназы Dre.

ii. IgG, выделенный из супернатанта трансфицированных pD2-D1.3 клеток, трансфицированных геном рекомбиназы Dre. ИФА секретируемых поликлональных антител. Отсортированные клетки из эксперимента, показанного на Фигуре 9Н (первоначально из клеток, продуцирующих популяцию антител, отобранных с помощью 1 раунда фагового дисплея), выращивали в течение 7 суток после сортировки, и собирали культуральный супернатант. Планшеты для ИФА покрывали либо  $\beta$ -галактозидазой (10 мкг/мл), либо БСА (10 мкг/мл) на протяжении ночи. Культуральный супернатант разводили до концентрации 66 % добавлением 33 % конечного объема 6 % раствора обезжиренного сухого молока (Marvel) в PBS (описан выше как «неразведенный образец»). Супернатант также разводили в соотношении 1/10 в PBS и смешивали с 6 % MPBS аналогичным образом. Обнаружение связавшегося слитого белка scFv-Fc осуществляли с использованием меченого Eu антитела мыши к IgG человека (Perkin Elmer номер по каталогу 1244-330).

Фигура 20. Получение фрагментов ДНК для придания отобранной популяции молекул scFv формы IgG, (а) создание ДНК-вставки CL-pA-CMV-Sigp (как представлено на Фигуре 21b). ПЦР-амплификация плазмиды pD2 с использованием праймеров 2595 и 2597 и очистка после разделения в агарозном геле. Дорожка m, маркер длин фрагментов ДНК Generuler 1 kb ladder (Thermo, SM031D), дорожка 1 CL-pA-CMV-Sigp ДНК-вставка, (b) создание ДНК-вставки scFv осуществлялось, как описано в Примере 6. Дорожка m, маркер длин фрагментов ДНК Generuler 1 kb ladder (Thermo, SM031D), дорожка 1 контрольный образец, дорожка 2 очищенный scFv, дорожка 3 популяция фрагментов ДНК, кодирующих scFv, полученная в результате проведения первого раунда селекции с использованием  $\beta$ -галактозидазы, дорожка 4 популяция фрагментов ДНК, кодирующих scFv, полученная в результате проведения первого второго раунда селекции с использованием  $\beta$ -галактозидазы, дорожка 5 популяция фрагментов ДНК, кодирующих scFv, полученная в результате проведения первого второго раунда селекции с использованием CD229, (c)

очистка «миникольцевой» ДНК, расщепленной с помощью NheI и XhoI. Миникольцевые ДНК (Фигура 21 с), образующиеся в результате лигирования расщепленной NcoI/NotI ДНК, кодирующей scFv (Фигура 21, вставка а), и ДНК, кодирующей константную область легкой цепи ( $C_L$ ), поли(А)-последовательности (pA), промотора CMV и сигнального пептида (Фигура 21, вставка b), очищали с использованием микроцентрифужных колонок и после расщепления рестрикционными эндонуклеазами NheI и XhoI разделяли в 1 % агарозном геле. На дорожках представлены: m, маркер длин фрагментов ДНК Generuler 1 kb ladder (Thermo, SM031D), дорожка 1 продукт, полученный в результате одного раунда селекции с использованием  $\beta$ -галактозидазы, продукт, полученный в результате двух раундов селекции с использованием  $\beta$ -галактозидазы, дорожка 3 продукт, полученный в результате двух раундов селекции с использованием CD229. Линеаризованный продукт размером 2,6 т.п.о., указанный стрелкой, был вырезан и очищен.

Фигура 21. Схематическое изображение процесса преобразования формата антитела из scFv в IgG. ДНК-вставку (а), кодирующую  $V_H$ - и  $V_L$ -домены антител, лигируют с ДНК-фрагментом (b), кодирующим константную область легкой цепи ( $C_L$ ), последовательностью полиаденилирования (pA), промотором цитомегаловируса (CMV) и сигнальным пептидом (SigP). Соединению молекул ДНК а и b для получения «миникольцевой» ДНК нерепликативным путем способствует лигирование по «липким концам». После лигирования «миникольцо» подвергается линеаризации с использованием рестриктаз NheI и XhoI. Линеаризованный продукт d затем очищают и лигируют с расщепленным вектором e. Вектор e включает промотор pEF и последовательность SigP, расположенные выше сайта NheI и кодирует константные домены 1-3 тяжелой цепи (CH) антитела, расположенные ниже сайта XhoI. Продукт лигирования вставки d с вектором e дает в результате плазмиду f, которая может быть использована для трансформации бактерий, а выращивание с использованием подходящего селективного маркера обеспечит получение и очистку плазмидной ДНК с помощью стандартных способов. Очищенная плазида f может быть введена в клетки млекопитающих [134] для гетерологичной экспрессии антитела Ig. В качестве альтернативы ДНК, кодирующая CH1-3 в векторе e, может быть заменена на ДНК, кодирующую один CH1-домен для экспрессии Fab-фрагмента.  $V_H$  и  $V_L$  представляют собой переменные домены тяжелой и легкой цепей антитела, соответственно. На изображении представлены: ДНК, кодирующая промотор фактора элонгации (pEF), константный домен легкой цепи ( $C_L$ ) и константные домены 1-3 тяжелой цепи (CH1-3) антитела, последовательность полиаденилирования (pA), промотор цитомегаловируса (CMV) и сигнальный пептид (SigP).

Фигура 22. Дополнительный пример получения фрагментов ДНК, требуемых для преобразования scFv в IgG (а) вставки scFv, созданные, как описано в Примере 14, разделяли в 1 % агарозном геле, приготовленном на TBE. Дорожки 1 и 14 соответствуют маркеру длин фрагментов ДНК 500 п.о., начинающегося с 500 п.о. Дорожки 2-13 соответствуют ПЦР-продуктам scFv. (б) Очистка линейаризованного «миникольца» d (Фигура 21) осуществляли путем разделения в 1 % агарозном геле, приготовленном на TBE. Слева направо первая дорожка соответствует маркеру длин фрагментов ДНК (маркер длиной 1 т.п.о., Lifetech, 15615-024), а остальные дорожки соответствуют линейаризованному «миникольцу» d. (с) Поскольку (б) за исключением промотора CMV заменен на последовательность P2A и использован маркер длин фрагментов ДНК DNA Generuler 1 kb ladder (Thermo, SM031D).

Фигура 23. Направляемое нуклеазой встраивание генов связывающих молекул с использованием систем проточной электропорации. Смесь в соотношении 50:50 плазмид pD6, кодирующих либо антитело к FGFR1, либо антитело к FGFR2 антитело подвергали электропорации с использованием системы проточной электропорации. Через 13 суток отбора в присутствии бластицидина клетки были мечены FGFR1-Fc, меченым Dyelight-633 (FGFR1-Dy633), или FGFR2-Fc, меченым Dyelight 488 (FGFR2-Dy488). В виде точечных диаграмм представлены:

- a. Одиночное окрашивание FGFR2-488 (образец 1b из Таблицы 7)
- b. одиночное окрашивание FGFR1-633 (образец 1 b из Таблицы 7)
- c. Двойное окрашивание FGFR1-633/ FGFR2-488 (образец 1b из Таблицы 7)
- d. Двойное окрашивание FGFR1-633/ FGFR2-488 (образец 3 из Таблицы 7)

Фигура 24. Выделение генов антител после проточной сортировки.

Популяция молекул антител была отобрана с использованием одного раунда фагового дисплея, и библиотека для дисплея на клетках млекопитающих была получена с помощью проточной электропорации с использованием системы Maxcyte. Сортировку клеток осуществляли с использованием либо 1 нМ, либо 10 нМ антигена, и сразу после этого выделяли мРНК. Гены антител были получены с помощью ПЦР, клонированы в бактериальный экспрессирующий вектор, и была установлена доля положительных клеток по результатам ИФА («выход при использовании 1 нМ», «выход при использовании 10 нМ»). Этот выход сравнивали с первоначальным выходом из раунда 1 («выход при использовании одного раунда фагового дисплея»). На графике показан профиль ИФА-сигналов, полученных для каждой популяции.

Фигура 25. Вектор pINT20 для экспрессии T-клеточных рецепторов.

a. Изображение плазмиды pINT20 с двойным промотором, на котором показаны гомологичные участки AAVS, селективный ген устойчивости к пурамицину (при этом участок возле акцепторного сайта сплайсинга показан ниже). Альфа-цепь (включающая в себя переменный домен альфа-цепи, при этом константный домен CD3 $\zeta$  альфа-цепи мыши фланкирован сайтами рестрикции Nhe1, Not1 и Acc651 и находится под контролем промотора pEF. Бета-цепь (включающая в себя переменный домен бета-цепи, при этом константный домен CD3 $\zeta$  бета-цепи фланкирован сайтами Nco1, Xho1 и hind3 и находится под контролем промотора CMV).

b. Последовательность у акцепторного сайта сплайсинга и начало гена устойчивости к пурамицину (SEQ ID NO: 24 и 25)

c. Последовательность конструкции альфа-цепи Т-клеточного рецептора клона c12/c2, в которой показаны сайты рестрикции Nhe1, Not1 и Acc651 (SEQ ID NO: 26 и 27).

d. Последовательность конструкции бета-цепи Т-клеточного рецептора клона c12/c2, в которой показаны сайты рестрикции Nco1, Xho1 и Hind 3 (SEQ ID NO: 28 и 29).

e. Последовательность конструкции альфа-цепи Т-клеточного рецептора клона 4JFH, в которой показаны сайты рестрикции Nhe1/Not1 (SEQ ID NO: 30 и 31).

f. Последовательность конструкции бета-цепи Т-клеточного рецептора клона 4JFH, в которой показаны сайты рестрикции Nco1/Xho1 (SEQ ID NO: 32 и 33).

g. Стратегия и праймер использованные для мутации CDR3 c12/c2 TCR альфа-цепь (SEQ ID NO: 34, 35, 36 и 37).

h. Стратегия и праймер, использованные для мутации CDR3 бета-цепи TCR c12/c2 (SEQ ID NO: 38, 39, 40 и 41).

(N= A, C, G, T; S= C или G; W= A или T)

Фигура 26. Распознавание комплексов пептида и молекул МНС Т-клеточными рецепторами, введенными в клетки млекопитающих с помощью направляемого нуклеазой встраивания.

TCR1 представляет собой пептид 1 (SLLMWITQV), распознающий c12/c2 TCR в форме комплекса с меченым фикоэритрином HLA-A2 (пептид 1). TCR2 представляет собой пептид 2 (ELAGIGILTV), распознающий 4JFH TCR, в форме комплекса с меченым фикоэритрином HLA-A2 (пептид 2). Образцы a и c показывают клетки, экспрессирующие TCR1, которые были подвергнуты воздействию пептида 1 (a) или пептида 2 (c). Образцы b и d показывают клетки, экспрессирующие TCR2, которые были подвергнуты воздействию пептида 1 (a) или пептида 2 (c).

Образцы e и f показывают нетрансфицированные клетки HEK293, которые были мечены с использованием пептида 1 (e) или пептида 2 (f).

g. Плазмиду, кодирующую TCR1, смешивали с 100-кратным избыточным количеством плазмиды, кодирующей TCR2, которая была введена с помощью направляемого нуклеазой встраивания в клетки НЕК. 1,15 % клеток были мечены с использованием пептида 1. 1,15 % клеток были положительно окрашенными.

h. Плазмиду, кодирующую TCR2, смешивали с 100-кратным избыточным количеством плазмиды, кодирующей TCR1, которая была введена с помощью направляемого нуклеазой встраивания, и была мечены с использованием пептида 2. 0,62 % клеток были мечены. Положительно окрашенные клетки собирали с помощью проточной сортировки и выделяли мРНК для анализа в отношении специфического обогащения TCR.

Образцы i-l иллюстрируют экспрессию библиотеки Т-клеточных рецепторов в клетках НЕК293. Библиотека TCR была введена путем электропорации Maxcyte и подвергалась отбору в течение 11 суток с использованием пурамицина. Образец i. показывает клетки, меченные, APC-меченым антителом к TCR (ось y), в примере j. показаны клетки, меченные фикоэритрин-меченым пептидом 1 и молекулами МНС (ось x) Образец k. показывает нетрансфицированные клетки, меченные как антителом к TCR, так и пептидом 1:МНС. Образец l. показывает трансфицированные библиотекой TCR1 клетки, меченные как антителом к TCR, так и пептидом 1 и молекулами МНС. Образцы m-n иллюстрируют экспрессию TCR в клетках Jurkat. TCR1 был введен путем электропорации Атаха и подвергался отбору в течение 25 суток с использованием пурамицина. Плазмиду трансфицировали в присутствии (m) или в отсутствие (n) нуклеазы TALE и инкубировали с меченым APC антителом к  $\beta$ -цепи TCR. Образцы o-g иллюстрируют активацию Т-клеточного рецептора в тех же самых TCR1-трансфицированных клетках Jurkat. Все клетки мечены антителом к CD69 (ось y). Образец o не подвергался стимуляции, а образец p подвергался стимуляции антителом к CD3 в течение 24 часов. Образцы q и r инкубировали в течение 24 часов с 2 мкл и 6 мкл, соответственно, меченых PE молекул МНС и пептида 1. Все клетки также подвергались воздействию антитела к CD28.

Фигура 27. Векторы pINT21 CAR1 и pINT21 CAR2 для введения библиотек химерных антигенных рецепторов (CAR) в клетки человека.

Изображение плазмиды pINT21 с одним промотором, на котором показаны гомологичные участки AAVS, селективный ген устойчивости к пурамицину, промотор CMV, контролирующей слияние связывающей молекулы с сигнальным доменом CD3 $\zeta$ . Сайты NcoI и Not I подходят для клонирования связывающей молекулы.

a. pINT21 CAR1 обеспечивает слияние связывающей молекулы с околочелюстным, трансмембранным и сигнальным доменом CD3 $\zeta$

- b. pINT21 CAR2 обеспечивает слияние связывающей молекулы с шарнирной областью CD8 и трансмембранным доменом, активирующими доменами 4-1BB и CD3 $\zeta$
- c. Последовательность CD3 $\zeta$  в pINT 21\_CAR1 (SEQ ID NO: 42, 43 и 44)
- d. Последовательность CD8, 4-1BB и CD3 $\zeta$  в pINT 21\_CAR2 (SEQ ID NO: 45 и 46)
- e. Последовательность FMC63 H-L (антитело к CD19) (SEQ ID NO: 47 и 48)

Фигура 28. Экспрессия scFv и альтернативного каркасного белка в составе конструкции химерного антигенного рецептора, введенной в клетки человека с помощью опосредованного нуклеазой встраивания.

Клетки НЕК были трансфецированы антителами к FGFR1 (b) или lox1 adhiрон (c) и мечены меченными FGFR1 и lox1 соответственно. В качестве контроля те же антигены инкубировали с нетрансфецированными клетками НЕК293 (a и c, соответственно).

Популяции из библиотек фагового дисплея, отобранных на мезотелине и CD229, были введены в клетки НЕК с помощью опосредованного нуклеазой встраивания (f и h, соответственно) и были отобраны с использованием пурамицина в течение 11 суток. Эти клетки или нетрансфецированные клетки НЕК293 инкубировали с меченым мезотелином (e, f) или CD229 (g, h). На изображениях e и h представлены нетрансфецированные клетки НЕК293.

Фигура 29. Последовательность альтернативных связывающих каркасных белков для дисплея на клетках млекопитающих библиотек связывающих молекул разных форматов без труда может быть введена с помощью направляемого нуклеазой встраивания с использованием вектора, описанного в данном документе. В качестве примера были получены конструкции Adhiрон с фланкирующими сайтами Nco I и Not I для введения в конструкции CAR или слитые конструкции с Fc-фрагментом.

- a. Последовательность связывающегося с lox1 Adhiрон\_lox1A (SEQ ID NO: 49 и 50)
- b. Последовательность связывающегося с lox1 Adhiрон\_lox1B (SEQ ID NO: 51 и 52) (Вариабельные петли показаны полужирным шрифтом и подчеркиванием на последовательности белка).
- c. Потенциальные мутагенные праймеры для создания библиотеки связывающих молекул в составе петли 1 (adhiрон mut1) (SEQ ID NO: 53, 55 и 56) или в составе петли 2 (adhiрон mut2) (SEQ ID NO: 54, 57 и 58).

Ниже представлено изображение участка, покрытого праймерами (нижняя цепь), на котором показана трансляция белка, n соответствует разному количеству кодонов NNS, что что приводит к образованию петель разной длин.

d. Последовательность связывающегося с трипсином ноттина MCoTI-II с фланкирующими сайтами NcoI и NotI, обеспечивающая экспрессию ноттина в составе векторов, описанных в данном документе. Последовательность первой петли выделена подчеркиванием SEQ ID NO: 59 и 60).

e. Стратегия для создания библиотеки мутированных форм ноттинов. В этом примере петлю 1 была заменена 10 случайным образом расположенных аминокислот. В этом примере были введены кодоны VNS (V=A, C или G), что приводит к получению 24 кодонов, кодирующих 17 аминокислот. Эта последовательность может быть введена в клон, кодирующий MCoTI-II с использованием стандартных способов (SEQ ID NO: 61, 62 и 63).

Фигура 30. Пример последовательности для опосредованной нуклеазой вставки генов антител посредством лигирования или соединения концов, опосредованное микрогомологией (MMEJ).

a. Последовательность pD7-SceI (нуклеотиды 1-120) (SEQ ID NO: 64, 65). Последовательность аналогична pD6 (см. Фигуру 3с, 8), за исключением того, что левый участок AAVS между ECoRI и NsiI было заменено на последовательность, распознаваемую мегануклеазой I-SceI (жирный шрифт). Кроме того, правый участок AAVS между AscI и MluI было заменено на вставку, кодируемую праймерами 2723 и 2734 (не показано).

b. Последовательность pD7-ObLiGaRe (нуклеотиды 1-120) (SEQ ID NO: 66, 67). Последовательность аналогична pD6 (см. Фигуру 3с, 8), за исключением того, что левый участок AAVS между ECoRI и NsiI было заменено на сайты распознавания AAVS TALE правого и левого участка. Кроме того, правый участок AAVS между AscI и MluI было заменено на вставку, кодируемую праймерами 2723 и 2734 (не показано).

Фигура 31. направляемое нуклеазой встраивание связывающихся молекул в локусе ROSA 26.

a. Показана последовательность левого гомологичного участка включительно до начала гена устойчивости к пурамицину, в которой показаны праймеры и сайты рестрикции, упомянутые в примере 22 (SEQ ID NO: 68).

b. Последовательность правого гомологичного участка для направляемого нуклеазой встраивания в локус ROSA 26, в которой показаны праймеры и сайты рестрикции, упомянутые в примере 22 (SEQ ID NO: 69).

## **Примеры**

Пример 1. конструирование векторов для экспрессии антител, получивших формат IgG



Для осуществления отбора связующих молекул (например, антитело, белок или пептид) на генетическом уровне требуется введение гена, кодирующего эту связующую молекулу, и контроль экспрессии этого гена с использованием экзогенного промотора, или путем прямого встраивания трансгена ниже промотора, предсуществующего в клеточной ДНК, например, эндогенного промотора. Антитела являются наиболее часто используемым видом связывающих молекул, и они могут подвергаться форматированию для экспрессии в разных формах. В приведенных ниже примерах описана продукция антител формата scFv-Fc, у которых scFv-фрагмент слит с Fc-доменом, путем экспрессии одного гена. Авторы также приводят пример экспрессии антител, получивших формат молекул IgG2 человека. Чтобы экспрессировать антитела, получившие формат IgG или FAb, в продуцирующих клетках, таких как клетки высших эукариот, требуется отдельная экспрессия тяжелых и легких цепей. Это можно осуществить путем введения отдельных плазмид, кодирующих каждую цепь, или путем их введения на одной плазмиде. В составе одной плазмиды могут быть экспрессированы 2 цепи в составе одной мультицистронной мРНК. Экспрессия различных белков в составе одной матричной РНК требует наличия элементов, как, например, последовательности внутренней посадки рибосомы (IRE), которые обеспечивают инициацию трансляции во вторичном положении в нисходящем направлении. В качестве альтернативы можно использовать элементы последовательности, способствующие остановке/повторной инициации трансляции, как, например, последовательности 2A вирусов [119].

В качестве альтернативы, множество различных белков может быть экспрессировано в составе одной плазмиды с использованием множества промоторов. На Фигуре 1a и 1b показана структура 2 похожих экспрессирующих кассет в составе разных каркасов векторов (pDUAL и pINT3), которые были разработаны для экспрессии секретлируемых антител, получивших формат IgG. Эти экспрессирующие кассеты были созданы с использованием комбинации генного синтеза и амплификации с помощью полимеразной цепной реакции стандартных элементов, таких как промоторы и поли(A)-последовательности. Первые отдельные плазмиды были созданы в составе pCMV/myc/ER (Фигура 1c, Life Technologies) для экспрессии тяжелой цепи (pBIOCAM1-NewNot) и легкой цепи (pBIOCAM2-pEF) антител. Элементы из pBIOCAM2-pEF (включая промотор pEF, ген легкой цепи и поли(A)-сайт) были клонированы в pBIOCAM1-NewNot) для получения pDUAL. Показанные примеры включают VH- и VL-домены от гуманизированного антитела к лизоциму под названием D1.3 [12 0] и называются pDUAL-D1.3 и pINT3-D1.3. Элементы pDUAL D1.3, представленные на Фигуре 1a, содержатся между сайтом EcoR1 и поли(A)-

сайтом BGH остова плазмиды из pCMV/мус/ER (Life Technologies номер по каталогу V82320, Фигура 1с).

Аналогичным образом отдельные кассеты с генами легкой цепи и тяжелой цепи вводили в pSF-pEF (Oxford Genetics OG43) и pSF-CMV-F1-Pac1 (Oxford Genetics OG111), соответственно, для создания p1NT1 и p1NT2. Они были объединены путем клонирования кассеты с геном легкой цепи (включая промотор pEF, ген легкой цепи и поли(А)-сайт) выше промотора CMV в p1NT2, для получения p1NT3. Элементы p1NT3-D1.3, представленные на Фигуре 1b, клонированы между первым Bgl2 и SBF1, которые представлены в составе плазмиды pSF-CMV-F1-Pac1 (Фигура 1d, Oxford Genetics OG111).

Немедленно-ранний промотор цитомегаловируса (промотор CMV) является сильным промотором, и он был использован для контроля экспрессии тяжелых цепей. pDUAL D1.3 также включает в себя состоящую из трех частей лидерную последовательность аденовируса 2 (TPL) и усиленный основной поздний промотор (enh MLP), следующий непосредственно ниже промотора CMV [121]. Фактор элонгации 1 альфа вырабатывается/синтезируется в большом количестве практически во всех типах клеток эукариот, поэтому промотор (pEF) гена этого белка обычно используют для контроля экспрессии трансгенов[122]. В pDUAL-D1.3 и p1NT3-D1.3 промотор pEF используется для контроля экспрессии легкой цепи антитела. Сайты полиаденилирования бычьего гормона роста (полиА BGH) содержатся в конце каждой экспрессирующей кассеты.

Секреция отдельных тяжелых и легких цепей в эндоплазматическом ретикулуме (и, в конечном счете, в культуральном супернатанте) находится под контролем 2 различных лидерных последовательностей. Секреция легкой цепи контролируется лидерной последовательностью VM40 [123]. За ней следуют сайты клонирования Nhe1 и Not1, которые обеспечивают клонирование генов VL в рамке считывания, которые, в свою очередь, слиты с геном С каппа-цепи человека. Контроль секреция тяжелой цепи осуществляется за счет разделения лидерной последовательности с помощью интрона гена VH мыши (который встречается в pCMV/мус/ER). За лидерной последовательностью следуют сайты Nco1 и Xho1, которые обеспечивают клонирование генов VH в рамке считывания, за которыми следует ген IgG2 с оптимизированными кодонами. Гены VL и VH гуманизированного антитела D1.3 [120] были клонированы в сайты Nhe1/Not1 и Nco1/Xho1, соответственно, в составе pDUAL-D1.3 и p1NT3-D1.3.

Варианты этих плазмид, кодирующие заякоренные в мембране варианты белков, были созданы для дисплея на клетках млекопитающих. Плазмида pD1 была создана путем расщепления pDUAL-D1.3 с помощью Bsu36I (которая разрезает CH3-домен гена тяжелой цепи IgG2) и с помощью BstZ171, которая разрезает остов за поли-А областью SV40 кассеты

с геном устойчивости к неомицину (Фигура 1с). Следовательно, в результате этого удаляется большая часть СНЗ-домена и вся экспрессирующая кассета с геном устойчивости к неомицину. СНЗ-домен заменяется на синтетическую вставку с совместимыми концами Bsu36I и BstZ171 (представлены на Фигуре 1е). Была сконструирована синтетическая вставка для замены стоп-кодона в конце СНЗ-домена антитела на донорный сайт сплайсинга и интрон, что обеспечивает сплайсинг СНЗ-конца с экзоном, кодирующим первые 5 внутриклеточных остатков трансмембранного домена рецептора PDGF человека [84], стоп-кодом и дополнительным донорным сайтом сплайсинга. За ним следует дополнительный интрон и акцепторный сайт сплайсинга, а затем следует кодон для одной аминокислоты и стоп-кодон (Фигура 1е). Эти 2 синтетические интрона, которые фланкируют экзон, кодирующий трансмембранный домен, были сконструированы с сайтами узнавания ROX, расположенными в их составе. Сайты ROX распознаются рекомбиназой Dge, вызывая рекомбинацию между ДНК, содержащей эти сайты [88]. Включение 2 сайтов ROX, фланкирующих экзон, кодирующий трансмембранный домен, создает возможность для удаления этого экзона путем трансфекции гена, кодирующего рекомбиназу Rox. Предположительно в результате этого образуется секретлируемый продукт – антитело.

На Фигуре 2 показана последовательность полученной в результате плазмиды с двойным промотором для экспрессии антител, экспрессирующей гуманизованное антитело D1.3 к лизоциму (в настоящем документе называется pD1-D1.3 (SEQ ID NO: 1). Специфичность связывания с лизоцимом обеспечивается посредством включения доменов последовательностей VH и VL из D1.3 [120] между сайтами рестрикции Nco/XhoI и NheI/NotI, соответственно. Последовательность показана от сайта EcoRI до BstZ171. Последовательности за пределами сайтов EcoRI и BstZ171 получены из остова вектора, который представлен на Фигуре 1с.

Пример 2. Конструирование вектора (pD2) для направленного встраивания кассеты с генами антител в локус AAVS

Расщепление геномной ДНК с использованием сайт-специфических нуклеаз облегчает вставку гетерологичной ДНК путем гомологичной рекомбинации или негомологичного соединения концов (NHEJ). ДНК клеток НЕК293 расщепляли нуклеазами, которые целенаправленно воздействуют на первый интрон протеинфосфатазы 1, ген регуляторной субъединицы 12С (PPP1R12C). Было установлено, что этот локус является основным сайтом встраивания аденоассоциированного вируса, и он называется сайтом AAVS (Фигура 3а). Сайт AAVS считается локусом, который представляет собой

«безопасную гавань» для вставки и экспрессии гетерологичных генов в клетках человека [12 4].

После сайт-специфического расщепления геномной ДНК появляется возможность встраивания кассеты, экспрессирующей белок, с использованием гомологичной рекомбинации. Для этого необходимо фланкировать экспрессирующую кассету участками, гомологичными последовательностям, встречающимся на любой стороне сайта расщепления генома. Для направленного встраивания в локус AAVS фрагмент локуса AAVS длиной 804 п. о., расположенный в 5'-направлении по отношению к предполагаемому сайту расщепления, амплифицировали с помощью ПЦР для создания сайтов EcoR1 и Mfe 1 на 5'- и 3'-конце, соответственно. Этот продукт, представляющий собой левое гомологичный участок для целенаправленного встраивания кассеты с генами антител, клонировали в сайт EcoR1 pD1, восстанавливая сайт EcoR1 на 5'-конце. В случае правого гомологичного участка фрагмент локуса AAVS длиной 836 п.о., расположенный в 3'-направлении по отношению к сайту расщепления, амплифицировали с помощью ПЦР для создания сайтов Bstz171 на каждом конце, и его клонировали в Bstz171 pD1. Конструкция представлена на Фигуре 3b, а последовательность полученной в результате конструкции (pD2) показана на Фигуре 3с.

Во время клонирования левого гомологичного участка AAVS на 3'-конце также была произведена вставка сайтов рестрикции Nsi1 и Pad1. Эти сайты были впоследствии использованы для клонирования синтетического интрона, за которым следует ген устойчивости к бластицидину и поли(А)-сайт. В гене устойчивости к бластицидину отсутствует промотер, но ему предшествует акцепторный сайт сплайсинга, который обеспечивает слияние в рамке считывания с вышерасположенным экзоном из локуса AAVS (Фигура 3а, b). Встраивание в локус AAVS вызывает экспрессию беспромоторного гена устойчивости к бластицидину. Последовательность этой итоговой конструкции, называемой pD2, показана на Фигуре 3с.

Последовательность кассеты с геном антитела, включающей в себя промотер pEF, ген легкой цепи D1.3, поли(А)-участок, промотер CMV, ген тяжелой цепи D1.3, альтернативные сайты сплайсинга и поли(А)-сайт показана на Фигуре 2. Во избежание повторений эта последовательность представлена на Фигуре 3с в виде блока с надписью: «КАССЕТА ДЛЯ ЭКСПРЕССИИ ГЕНА АНТИТЕЛА D1.3».

Пример 3. Направляемое AAVS TALEN встраивание конструкции IgG для экспрессии антитела на клеточной поверхности и связывания антигена

Клетки HEK293F (Life Technologies), выращиваемые в среде Freestyle, трансфецировали ДНК pD2-D1.3 в присутствии или в отсутствии пары векторов, со

вставкой AAVS для TALEN. Пара участков AAVS для TALEN («исходная плаزمида с последовательностями AAVS») была описана ранее [125], а TALEN распознает последовательность:

Левый участок для TALEN: 5' (T)CCCCTCCACCCCACAGT (SEQ ID NO: 70)

Спейсер 5' GGGGCCACTAGGGAC (SEQ ID NO: 71)

Правый участок для TALEN: последовательность, комплементарная к 5' AGGATTGGTGACAGAAAA (SEQ ID NO: 72)

(т. е.) 5' TTTTCTGTCACCAATCCT (SEQ ID NO: 73)

Альтернативная, более эффективная, направляемая последовательностями AAVS пара доменов TALEN, была выявлена и использована в последующих экспериментах (pZT-AAVS1 L1 TALE-N и pZT-AAVS1 R1 TALE, номер по каталогу No GE601A-1 System Biosciences). Эта пара, которая распознает один и тот же сайт (но не первый остаток «Т», показана в скобках выше), называется парой «AAVS-SBI» TALEN.

Клетки высевали в количестве  $0,5 \times 10^6$  клеток/мл и трансфецировали на следующий день в количестве  $10^6$  клеток/мл с использованием смеси ДНК:полиэтиленимина (PoIyPIus), которую добавляли в соотношении 1:2 (мас./мас.). Клетки трансфецировали 0,6 мкг/мл pD2 и котрансфецировали либо pDNA3.0 в качестве контроля (0,6 мкг/мл), либо комбинированными плазмидами, содержащими левые и правые «исходные последовательности AAVS» TALEN (0,3 мкг каждого/мл). pD3, которая экспрессирует EGFP под контролем промотора CMV (см. ниже), была включена в эксперимент в качестве контроля трансфекции, при этом эффективность трансфекции составляла 35 %. Селекцию клеток осуществляли в суспензионной культуре с использованием среды Freestyle (Life Technology) с добавлением 5 мкг/мл бластицидина.

Для того, чтобы определить экспрессию антител на клеточной поверхности, клетки окрашивали антителом к Fc-фрагменту иммуноглобулина человека в соответствии со следующим протоколом:

1. Через 16 дней после трансфекции  $0,5-1 \times 10^6$  клеток из популяций, отобранных с использованием бластицидина, центрифугировали в течение 2 минут (200-300 xg) при 4°C.
2. Клетки промывали с использованием 1 мл промывочного буфера (0,1 % БСА в PBS, Gibco № 10010) и осаждали в течение 2 минут (200-300 xg) при 4 °C.
3. Клетки ресуспендировали в 100 мкл окрашивающего буфера (1 % БСА в PBS) и добавляли 5-1 мкл из конъюгированных флуорохром антител. В качестве антител использовали меченные фикоэритрином антитела к Fc-фрагменту IgG человека (клон HP6017, номер по каталогу 409304, Biolegend) или меченные фикоэритрином антитела к

IgG2a мыши, изотипический контроль к-цепи (номер по каталогу 400214, Biolegend). Инкубацию проводили в течение > 30 мин при 4 °С в темноте.

4. Промывали дважды с использованием 1 мл промывочного буфера и ресуспендировали в 500 мкл промывочного буфера.

5. Добавляли 5 мкл раствора для оценки жизнеспособности клеток (№ 00-6993-50 eBioscience), содержащий 50 мкг/мл 7-амино-актиномицина D (7-AAD), чтобы идентифицировать мертвые клетки.

6. Клетки анализировали на проточном цитометре (Becton Dickinson FACS II).

На Фигуре 4 показано, что здесь популяция клеток, экспрессирующих антитела, была значительно большей при трансфекции pD2-D1.3 в присутствии AAVS, направленного на TALEN, при этом количество положительно окрашенных клеток составляло 86 % по сравнению использованным отдельно pD2-D1.3 – 1,5 % положительно окрашенных клеток.

Функциональные свойства антител к лизоциму, экспрессированных на поверхности, определяли путем оценки связывания с меченым антигеном. Лизоцим куриного яйца (Sigma: L6876) метили с использованием системы для конъюгации Lightning-Link Rapid (DyLight 488, Innova Biosciences: 322-0010) следующим образом:

1. Добавляли 10 мкл реагента LL-Rapid Modifier к 100 мкл лизоцима (200 мкг растворили в 100 мкл PBS) и аккуратно перемешивали.

2. Полученную добавляли смесь к смеси Lightning-Link® Rapid, перемешивали и аккуратно ресуспендировали с использованием возвратно-поступательного перемешивания пипеткой.

3. Инкубировали смесь в течение 15-30 мин в темноте при комнатной температуре.

4. Добавляли 10 мкл реагента LL-Rapid Quencher к реакционной смеси и аккуратно перемешивали.

5. Хранили при 4°C. Конечная концентрация лизоцима-Dy488 составляет 1,6 мкг/мкл.

6. Использовали 6 мкл лизоцима-Dy488 (~ 10 мкг) на одно окрашивание.

7. Окрашивание, отмывку и проточную цитометрию проводили, как описано выше.

Анализ показывает, что 86 % клеток, трансфицированных pD2-HUD 1,3, связывали меченый HEL (судя по гейту M1) по сравнению с 0,29 % в случае нетрансфицированных клеток (Фигура 5).

Пример 4. Сайт-специфические нуклеазы (TALEN, направляемые AAVS) усиливают встраивание донорной ДНК

Также производили посев трансфицированных клеток и селекцию с использованием бластицидина, чтобы определить количество клеток, в которых была активирована экспрессия беспромоторного гена устойчивости к бластицидину. Через 24 ч после трансфекции клетки высевали в количестве  $0,25 \times 10^6$  клеток/чашку Петри диаметром 10 см (из культурального пластика), и выращивали в минимальной питательной среде, содержащей 10 % фетальной бычьей сыворотки (10270-10<sup>6</sup>, Gibco) и 1 % заменимых аминокислот (MEM\_NEAA № 11140-035 Life Technologies). Еще через 24 ч добавляли 5 мкг/мл бластицидина и производили замену среды каждые 2 дня. Через 9 дней все клетки, которые не включили плазмиду pD2, погибли. Через 12 дней планшеты окрашивали 2 % метиленовым синим (в 50 % метанола). Плотность колоний была слишком высокой для точной количественной оценки, но наблюдалось повышение количества устойчивых к бластицидину колоний при наличии AAVS TALEN, что указывает на целенаправленное встраивание в локус AAVS. Меньшее количество ДНК вводили для более точного количественного определения.

Трансфекции проводили, как описано ранее, с использованием либо 50, 200 либо 400 нг pD2-D1.3/10<sup>6</sup> клеток в присутствии или в отсутствие AAVS TALEN (0,3 мкг/мл каждой нуклеазы TALEN, в случае наличия, Фигура 1А). Общее количество вводимой ДНК довели до 1,2 мкг ДНК на 10<sup>6</sup> клеток с использованием контрольной плазмиды pсDNA3.0. Через 24 часа после трансфекции  $0,25 \times 10^6$  клеток высевали в чашку Петри диаметром 10 см и через 24 часа после посева добавляли 7,5 нг/мл бластицидина. Через 10 дней после селекции с бластицидином колонии окрашивали 2 % метиленовым синим (в 50 % метанола). Результаты показаны на Фигуре 6 и представлены в краткой форме в Таблице 1А. Показано, что котрансфекция ДНК, кодирующей TALEN, направленные на AAVS, приводит к увеличению количества колоний устойчивых к бластицидину, приблизительно в 10 раз.

Сравнение проводилось между такими плазмидами как «исходная плазида с последовательностями AAVS» и «AAVS SBI», направляющих субъединицы нуклеазы TALEN, которая целенаправленно воздействует на локус AAVS. В Таблице 1В показано увеличение числа колоний, устойчивых к бластицидину, при использовании пары TALEN «AAVS SBI».

Таблица 1. Определение количества колоний, устойчивых к бластицидину, полученных в результате трансфекции pD2-D1.3

А.

Плазмида, несущая ген фермента	Донорная плазмида pD2-D1.3 (нг/10 <sup>6</sup> клеток)	С AAVS TALEN	Без AAVS TALEN
Исходная плазмида с последовательностями и AAVS	50	319	32
Исходная плазмида с последовательностями и AAVS	200	526	41
Исходная плазмида с последовательностями и AAVS	400	686	75

## В.

Плазмида, несущая ген фермента	Донорная плазмида pD2-D1.3 (нг/10 <sup>6</sup> клеток)	С AAVS TALEN	Без AAVS TALEN (Контрольная плазмида pсDNA3.0)
Исходная плазмида с последовательностями и AAVS	300	1420	111
Исходная плазмида с последовательностями и AAVS	1000	1080	127
Исходная плазмида с последовательностями и AAVS	3000	560	70
AAVS-SBI	300	2800	111
AAVS-SBI	1000	1630	127
AAVS-SBI	3000	870	70

В данном примере авторы сравнили эффект встраивания с помощью нуклеазы TALEN, используя либо экспрессию антитела на клеточной поверхности (Пример 3), либо активацию беспромоторного гена устойчивости к бластицидину (Пример 4). Преимущество направляемого нуклеазами встраивания является более очевидным при оценке экспрессии антител по сравнению с влиянием, оказываемым на колонии, устойчивые к бластицидину.



Одним из возможных объяснений этого эффекта является то, что уровни экспрессии, необходимые для выживания в присутствии бластицидина, могут быть значительно меньшими, чем уровни экспрессии, необходимые для выявления экспрессии IgG2 на поверхности. Таким образом, ошибочное встраивание/сплайсинг беспромоторного гена устойчивости к бластицидину может приводить к снижению уровня экспрессии гена устойчивости к бластицидину на фоне повышения числа колоний, устойчивых к бластицидину в отсутствие значительной экспрессии антител.

#### Пример 5. Определение точности встраивания с использованием AAVS TALEN.

Для изучения точности встраивания были выделены колонии, полученные в эксперименте из Примера 4/Таблицы 1А (из дублированных, неокрашенных планшетов), размножены, а геномная ДНК из этих клеток была использована в качестве матрицы в ПЦР. Для получения геномной ДНК клетки собирали и ресуспендировали в 700 мкл лизирующего буфера (10 mM Tris.Cl, pH = 8,0, 50 mM ЭДТА, 200 mM NaCl, 0,5 % SDS, дополненного 0,5 мг/мл протеиназы К (добавляли непосредственно перед лизисом)). Затем клетки, ресуспендированные в лизирующем буфере, переносили в микроцентрифужную пробирку и инкубировали при 60 °С в течение приблизительно 18 часов. На следующий день к лизату добавляли 700 мкл изопропанола, чтобы осадить геномную ДНК. Микроцентрифужную пробирку осаждали при 13 000 об/мин в течение 20 минут. Осадок геномной ДНК промывали 70 %-ным этанолом и осаждали при 13 000 об/мин в течение 10 минут. После центрифугирования аккуратно отбирали супернатант, стараясь не касаться осадка геномной ДНК. Осадок геномной ДНК затем повторно ресуспендировали в 100 мкл буфера, содержащего 10 mM Трис (pH 8,0), 1 mM ЭДТА и инкубировали при 60°С в течение 30 минут с открытой крышкой для того, чтобы избавиться от остатков этанола. К 100 мкл этого раствора добавляли РНКазу-А (конечная концентрация 20 мкг/мл) и инкубировали при 60 °С в течение примерно одного часа. Концентрацию геномной ДНК измеряли с помощью спектрофотометра (NanoDrop).

Для идентификации точности встраивания были сконструированы ПЦР-праймеры, которые гибридизуются в геномном локусе AAVS за пределами левых и правых гомологичных участков. Они были попарно соединены с помощью встраивания специфических праймеров. Праймеры на 5'-конце:

Граница левого участка AAVS-ПЦР-прямой (9625) 5'  
 CCGGAACTCTGCCCTCTAAC (SEQ ID NO: 74) граница BSD\_ - ПЦР-обратный (9626): 5'  
 TAGCCACAGAATAGTCTTCGGAG (SEQ ID NO: 75)

Если происходит правильное встраивание, получают продукт длиной 1,1 т.п.о. В случае 8 из 9 клонов, полученных в результате направленного AAVS встраивания, наблюдалась полоса ожидаемого размера (Фигура 7а, b). В случае 2 клонов, устойчивых к бластицидину, полученных в отсутствие TALEN, образование продукта не наблюдалось (Фигура 11а), что указывает на случайное встраивание. Праймеры на 3'-конце:

Последовательность донорной плазмиды PDGFRTM-2, прямой праймер 5' ACACGCAGGAGGCCATCGTGG (SEQ ID NO: 76), граница правого участка AAVS1\_ ПЦР-обратный 5' TCCTGGGATACCCCGAAGAG (SEQ ID NO: 77)

Таким образом получают продукт длиной 1,5 т.п.о. с правильным встраиванием. Для 7 из 9 клонов, полученных в результате направленного встраивания, наблюдалась полоса ожидаемого размера. В случае 2 клонов, устойчивых к бластицидину и полученных в отсутствие TALEN, образования продукта не наблюдалось (рисунок 11b). Таким образом, большинство устойчивых к бластицидину клеток получают в результате правильного встраивания в локус AAVS, в то время как устойчивые к бластицидину колонии, полученные в отсутствие TALEN, были встроены неправильно.

Пример 6. Создание дисплейной библиотеки scFv из отобранной популяции, полученной с помощью фагового дисплея, и отбор посредством дисплея на клетках млекопитающих

Растворимые формы антител с форматом, измененным на scFv, были ранее экспрессированы в составе вектора pBIOCAM5-3F, в котором экспрессия осуществлялась под контролем промотора CMV, а часть вектора, в состав которой входили человеческий Fe, His6 и 3xFLAG, представляла собой С-концевой участок экспрессируемого слитого белка [105, 12 6]. Его модифицировали для создания вектора pBIOCAM5newNot, в котором сайт Not1 был встроен в Fc-область антитела (как показано на Фигуре 8). Его использовали в качестве исходного вектора для создания вектора pD6 (Фигура 8) для экспрессии слитых белков scFv-Fc, прикрепленных к клеточной поверхности. Были сконструированы праймеры (2598 и 2619) для амплификации экспрессионной кассеты, содержащей кодирующую scFv-Fc последовательность под контролем промотора CMV, с использованием в качестве матрицы плазмиды pBIOCAM5newNot. Праймер 2598 гибридизуется перед промотором CMV и вносит в последовательность амплифицируемого фрагмента сайт Pac1 (подчеркнут).

2598: TTTTTTTТААТТАА GATTATTGACTAGTTATТААТАGТААТСААТТАСGGGGTС (SEQ ID NO: 78)

Праймер 2619 гибридизуется возле конца Fc-домена и вносит в начало интрона донорный сайт сплайсинга и сайт Pme1 (подчеркнут).

2619: TTTTTTGTTTAAACTTACCTTGGATCCCTTGCCGGGGCTCAGGCTCAGGGAC  
(SEQ ID NO: 79)

Полученный ПЦР-продукт совместим с сайтами Pac1 и Pme1 из pD2 (Фигура 3).

Расщепление pD2 с помощью Pac1 и Pme1 приводит к удалению последовательности: промотор рEF-лидерная последовательность-легкая цепь-промотор CMV- лидерная последовательность- тяжелая цепь.

При клонировании продукта ПЦР, расщепленного с помощью Pac1/Pme1, в плазмиду помещается последовательность, содержащая: промотор CMV-лидерная последовательность-сайты Nco1/Not1- Fc-фрагмент иммуноглобулина человека.

Клонирование таким способом обеспечивает размещение кассеты, кодирующей scFv-Fc, соответствующим образом для сплайсинга с расположенным ниже трансмембранным доменом, который был описан ранее, для презентации IgG на клеточной поверхности при использовании pD2. Последний вектор pD6 показан на Фигуре 8, последовательность D6 показана в направлении от сайта Nco1 к сайту Pme1.

Отбор методом фагового дисплея был проведен с использованием библиотеки фагового дисплея McCafferty [7] с помощью бета-галактозидазы (Rockland, номер по каталогу B000-17) и CD229 (R и D Systems, Cat-898 CD-050) в качестве антигенов. По сути, были использованы способы отбора и субклонирования, которые были описаны ранее [6, 7, 118, 127]. Гены scFv из популяций, вытекающих из одного или двух циклов отбора с использованием бета-галактозидазы и двух циклов отбора с использованием CD229 были получены с помощью ПЦР. Гибридизация праймера M13Leadseq осуществляется в бактериальной лидерной последовательности, находящейся перед геном scFv, а гибридизация Notmycseq осуществляется в теге mус, находящемся за геном scFv в векторе для фагового дисплея [127].

M13Leadseq (SEQ ID NO: 80)

AAA TTA TTC GCA ATT CCT TTG GTT CCT

Notmycseq (SEQ ID NO: 81)

GGC CCC ATT CAG ATC CTC TTC TGA GAT GAG

ПЦР-продукт расщепляли с помощью Nco1 и Not1, полученную вставку очищали после расщепления и разделения в агарозном геле. Расщепленный продукт был лигирован в сайты Nco1 и Not1 бактериальной экспрессирующей плазмиды pSANG10-3F, а экспрессию и скрининг антител осуществляли, как было ранее описано [127]. После 2

раундов отбора с использованием бета-галактозидазы и CD229, 40/190 (21 %) и 35/190 (18 %) клонов оказались положительными по результатам ИФА.

550 нг вставки, вырезанной с помощью Nco/Not, лигировали с плазмидой pD6 (2,4 пг), гидролизованной Nco1 и Not1, для получения конструкции, экспрессирующей белок, полученный в результате слияния scFv и Fc-области IgG2 человека. Лигированную ДНК трансформировали в электрокомпетентные клетки NEB5alpha (New England Biolabs, номер по каталогу C2989), в результате чего была получена библиотека размером 2-3 x10<sup>7</sup> клонов для каждой популяции. Получали ДНК и котрансфецировали в клетки HEK293 100mls, выращенных в среде Freestyle, как описано выше, с использованием 0,3 мкг донорной ДНК (библиотека pD6) на 10<sup>6</sup> клеток. Клетки котрансфецировали с использованием 0,5 мкг каждой из плазмид «AAVS-SBI» TALEN (pZT-AAVS1 L1 TALE-N и pZT-AAVS1 R1 TALE, номер по каталогу GE601A-1 System Biosciences).

Через 24 ч после трансфекции общий объем культуры увеличился в два раза и через 24 часа добавляли бластицидин (10 мкг/мл). Каждые 3-4 дня вносили новую среду и через 6 дней концентрацию бластицидина увеличили до 20 мкг/мл.

Для того, чтобы определить размер библиотеки, 20 000 клеток высевали в чашки Петри диаметром 10 см (из культурального пластика) через 24 часа после трансфекции и выращивали в минимальной питательной среде, содержащей 10 % фетальной бычьей сыворотки (10270-106, Gibco) и 1 % заменимых аминокислот (MEM\_NEAA № 11140-035 Life Technologies). Еще через 24 ч добавляли 10 мкг/мл бластицидина и производили замену среды каждые 2 дня. Через 8 дней планшеты окрашивали 2 % метиленовым синим (в 50 % метанола). Результаты показаны в Таблице 2. Показано, что были получены библиотеки размером около 3 x 10<sup>6</sup> клонов (что составляет 3 % от трансфецированных клеток) для 3-х популяций.

Таблица 2. Определение размера библиотеки scFv-Fc.

Образец	Отсутствие колоний/20000 клеток	Отсутствие колоний/10 <sup>6</sup> клеток	Размер библиотеки
β-галактозидаза, раунд 1	546	27300	2,7 x 10 <sup>6</sup>
β-галактозидаза, раунда 2	654	32700	3,2 x 10 <sup>6</sup>
CD229, 2 раунда	556	27800	2,8 x 10 <sup>6</sup>

Протокол для мечения и поточной сортировки 10-20 x 10<sup>6</sup> клеток приведен ниже. Первоначальный анализ проводили через 13 дней после трансфекции с использованием

всего лишь  $10^6$  клеток/образец и с меньшими инкубационными объемами (объемы реагентов составляют 1/10 от показанных).

На Фигуре 9 показано, что на 13-й день после трансфекции по меньшей мере 43-46 % клеток экспрессируют слитый белок scFv-Fc на клеточной поверхности и его можно обнаружить либо с помощью FITC-, либо фикоэритрин-меченых антител к Fc-фрагменту. Связывание биотинилированной бета-галактозидазы также обнаруживается в пределах этой популяции с использованием либо FITC-, либо фикоэритрин-меченого стрептавидина. Используя стрептавидин-FITC была выявлено 11,8 % и 39 % положительных клеток в отношении экспрессия антител и связывания антигена с использованием библиотек, полученных из выходных популяций, полученных из 1 или 2-х раундов отбора методом фагового дисплея, соответственно. В случае CD229, 66 % клеток, полученных в результате 2-х раундов фагового дисплея, были положительными по scFv-Fc, а 24 % из них были положительными в отношении связывания CD229 (15 % от общей численности популяции).

На 20-е сутки после трансфекции клетки метили в соответствии с протоколом, который представлен ниже (с использованием биотинилированного антигена/фикоэритрин-меченого стрептавидина и FITC-меченого антитела к Fc-фрагменту иммуноглобулина человека).

1. Собрать, промыть и довести концентрацию клеток до  $15-20 \times 10^6$  клеток на пробу. Осадить клетки при 250 g в течение 4 минут при комнатной температуре, промыть клетки 1 мл PBS + 0,1 % БСА (4 °C), осадить клетки при 250 g в течение 4 минут при комнатной температуре, ресуспендировать в 1 мл PBS + 1 % БСА

2. Добавить биотинилированный антиген до конечной концентрации 100 нМ и инкубировать 30 мин при 4°C

3. Промыть клетки 2 раза по 1 мл 0,1 % БСА путем центрифугирования при 1500 об/мин в течение 5 минут.

4. Добавить либо:

10 мкл FITC-меченого стрептавидина (1 мкг/мл, Sigma, номер по каталогу S3762) и 20 мкл фикоэритрин-меченых антител к Fc-фрагменту иммуноглобулина человека (200 мкг/мл, Biolegend, номер по каталогу 405203), либо:

20 мкл фикоэритрин-меченого стрептавидина (200 мкг/мл, Biolegend, номер по каталогу 405203) и 20 мкл FITC-меченых антител к Fc-фрагменту иммуноглобулина человека (200 мкг/мл, Biolegend, номер по каталогу 409310) в PBS + 1 % БСА в течение 15 мин при 4°C в темноте.

5. Промыть клетки 2 раза по 1 мл 0,1 % БСА путем центрифугирования при 1500 об/мин в течение 5 минут.

6. Ресуспендировать их в 500 мкл охлажденного на льду PBS + 1 % БСА
7. Добавить 20 мкл 7 AAD/флакон для оценки жизнеспособности.

Для сортировки проводили гейтинг клеток на основе размера клетки, зернистости, длительности импульса и жизнеспособности (посредством окрашивания 7-AAD, малоугловое светорассеяние и боковое рассеяние. Результаты показаны на Фигуре 9c и f. В общей сложности проводили сортировку 10 миллионов клеток, и 3,1 % и 7 % двойных положительных клеток были отобраны для библиотек, полученных из выходных популяций, полученных в результате 2-х раундов отбора (CD229 R2) и 1-ого раунда отбора с использованием  $\beta$ -галактозидазы ( $\beta$ -galR1), соответственно.

Клетки, отобранные из клеток  $\beta$ -galR1, выращивали еще в течение 20 дней и повторно анализировали (Фигура 9h). Показно, что в результате этого большинство клеток экспрессируют scFv-Fc и связывают  $\beta$ -галактозидазу. На этой фигуре также показано, что доля двойных положительных клеток в составе не подвергавшейся отбору популяции не уменьшалась через 42 дней после трансфекции (Фигура 9k).

Геномную ДНК получали из  $150\ 000\text{-}10^6$  отсортированных клеток. Геномную ДНК получали с использованием способа, описанного ранее, или с использованием набора «GenElute mammalian genomic DNA miniprep kit» (Sigma G1N10).

Гены scFv амплифицировали методом ПЦР с использованием геномной ДНК с использованием следующих праймеров:

2623 (SEQ ID NO: 82)

TAAAGTAGGCGGTCTTGAGACG

2624 (SEQ ID NO: 83)

GAAG GTG CTGTTG AACTGTTCC

Реакции ПЦР проводили с использованием полимеразы Phusion (NEB, номер по каталогу M0532S) в буфере от изготовителя, содержащем 0,3 мкМ каждого праймера и 3 % ДМСО. 100-1000 нг геномной ДНК использовали в качестве матрицы в 50 мкл реакционной смеси. Осуществляли 30 циклов при 98 °С в течение 10 секунд, 55°С в течение 25 секунд, 72°С в течение 45 секунд. В результате этого был получен продукт размером 1,4 т.п.о., который расщепляли с помощью Nco1 и Not1. Была получена полоса приблизительно 750-800 п.о. и очищена после разделения в агарозном геле до клонирования в pSANG10. Лигированную ДНК трансформировали в клетки BL21 (компетентные клетки BL21 (DE3), Edge Bio Ultra, номер по каталогу 45363). Таким образом, scFv-фрагменты, полученные из отсортированной популяции, могут быть экспрессированы в бактериях, как было описано ранее [7, 127].

В качестве альтернативы выделению гена антитела и экспрессии в альтернативной комбинации вектор/хозяин можно получить секретируемые антитела непосредственно из отобранных клеток либо после клонирования одной клетки, либо с использованием отсортированной популяции для получения смеси поликлональных антител. Для примера, культуральный супернатант отбирали из отсортированных клеток (из клеток  $\beta$ -galR1) через 7 дней в культуре. С использованием планшетов, покрытых  $\beta$ -галактозидазой, были получены положительные результаты по ИФА (см. Пример 13 и Фигуру 19b).

Пример 7. Создание и отбор из дисплейной библиотеки IgG, состоящей из отобранной популяции, полученной с помощью фагового дисплея

Фрагменты ДНК, кодирующие scFv, представляющие собой продукты, которые получают в результате 1 и 2 раундов отбора с помощью фагового дисплея антител с использованием  $\beta$ -галактозидазы и CD229, были получены, как описано в Примере 6. Популяциям молекул scFv придавали формат IgG в соответствии с Примером 14 и подробным описанием способа, которое приведено ниже.

ДНК-вставку, кодирующую константный домен легкой каппа-цепи человека ( $C_L$ ), последовательность полиаденилирования (pA), промотор CMV и сигнальный пептид из  $V_H$ -цепи иммуноглобулина мыши (представленный между сайтами NotI и NcoI pD2, которые показаны на Фигуре 21b), амплифицировали методом ПЦР в составе плазмиды pD2 с использованием праймеров 2595 (GAGGGCTCTGGCAGCTAGC) (SEQ ID NO: 84) и 2597 (TCGAGACTGTGACGAGGCTG) (SEQ ID NO: 85). ПЦР-реакции проводили с использованием полимеразы KOD hot start polymerase (Novagen, номер по каталогу 71086-4) в буфере от производителя, содержащем 0,25 мкМ каждого праймера. 10 нг ДНК плазмиды pD2 использовали в качестве матрицы в 50 мкл реакционной смеси. Осуществляли 25 циклов при 98 °C в течение 10 секунд, 55 °C в течение 25 секунд, 72 °C в течение 40 секунд. В результате этого был получен продукт длиной 1,8 т.п.о., который расщепляли с использованием NcoI и NotI и очищали после разделения в агарозном геле (Фигуре 20a представлена в виде вставки  $C_L$ -pA-CMV-SigP на Фигуре 21b).

Фрагменты ДНК, кодирующие scFv, представляющие собой продукты, которые получают в результате 1 и 2 раундов отбора с помощью фагового дисплея антител с использованием  $\beta$ -галактозидазы и CD229, были получены, как описано в Примере 6. На Фигуре 20b показаны популяции scFv, отобранных с использованием  $\beta$ -галактозидазы и CD229, разделенных с помощью электрофореза в 1 % агарозном геле.

Реакции лигирования между вставкой scFv и вставкой  $C_L$ -pA-CMV-SigP осуществляли путем инкубации расщепленной NcoI/NotI вставки scFv (1 мкг) с расщепленной NcoI/NotI вставкой  $C_L$ -pA-CMV-SigP (1 мкг) с ДНК-лигазой T4 (1,5 мкл,

Roche, 10-481-220-001) в буфере от производителя в общем объеме 40 мкл с образованием «мини-кольца», изображенного на Фигуре 21с. После лигирования образцы инкубировали при 16 °С в течение 16 часов, очищали с использованием микроцентрифужной колонки, подвергали расщеплению с помощью NheI и XhoI, а продукт длиной 2,6 т.п.о. (изображен на Фигуре 21d) очищали с помощью электрофоретического разделения в 1 % агарозном геле (Фигура 20с).

Вставку ДНК, изображенная на Фигуре 21d, кодирующая V<sub>L</sub>-C<sub>L</sub>-pA-CMV-SigP-V<sub>H</sub> (0,5 мкг), лигировали с расщепленным NheI/XhoI, очищенным после разделения в агарозном геле вектором pD2 (0,7 мкг) (Фигура 21e), с ДНК-лигазой T4 (1,5 мкл, Roche, 10-481-220-001) в буфере от производителя в общем объеме 40 мкл, для получения направленного вектора, изображенного на Фигуре 21f. Он кодирует популяции антител, получивших формат молекул IgG, образующихся в результате первого или второго раунда фагового дисплея антител с использованием β-галактозидазы или CD229. После лигирования образцы инкубировали при 16°C в течение 16 часов, очищали с использованием микроцентрифужной колонки и элюировали с использованием воды для ВЭЖХ.

Лигированную ДНК трансформировали в электрокомпетентные клетки NEB5-альфа (New England Biolabs, номер по каталогу C2989), в результате чего была получена библиотека размером 1-4 x10<sup>5</sup> клонов для каждой популяции. Получали ДНК и котрансфецировали в клетки HEK293 100 мл, выращенные на среде Freestyle, как описано выше, с использованием 0,3 мкг донорной ДНК (библиотека pD6- на 10<sup>6</sup> клеток. Клетки котрансфецировали с использованием 0,5 мкг каждой из плазмид «AAVS-SBI» TALEN (pZT-AAVS1 L1 TALE-N и pZT-AAVS1 R1 TALE, номер по каталогу GE601A-1 System Biosciences).

Через 24 ч после трансфекции объем культуры увеличился в два раза и еще через 24 часа добавляли бластицидин (10 мкг/мл). Каждые 3-4 дня вносили новую среду и через 6 дней концентрацию бластицидина увеличили до 20 мкг/мл.

Для того, чтобы определить размер библиотеки, 250000 клеток высевали в чашки Петри диаметром 10 см (из культурального пластика) через 24 часа после трансфекции и выращивали в минимальной питательной среде, содержащей 10 % фетальной бычьей сыворотки (10270-106, Gibco) и 1 % заменимых аминокислот (MEM\_NEAA № 11140-035 Life Technologies). Еще через 24 ч добавляли 10 мкг/мл бластицидина и производили замену среды каждые 2 дня. Через 8 дней планшеты окрашивали 2 % метиленовым синим (в 50 % метанола). Результаты показаны в Таблице 3. Показано, что были получены библиотеки



размером около  $5-9 \times 10^5$  клонов (что составляет от 0,5 % до 0,9 % трансфицированных клеток) для 3-х популяций.

Таблица 3. Определение размера дисплейных библиотек молекул на клетках млекопитающих, получивших формат IgG

Образец	Количество колоний/ $0,25 \times 10^6$ клеток	Количество колоний/ $10^6$ клеток	Размер библиотеки ( $\times 10^5$ )
β-галактозидаза, раунд 1	1337	5348	5,3
β-галактозидаза, раунд 2	1972	7888	7,9
CD229, раунд 2	2175	8700	8,7

Основную часть популяции клеток, трансфицированных с использованием продуктов 1-го или 2-го раундов отбора с использованием β-галактозидазы, подвергали отбору в бластицидин-содержащей среде, как описано выше. Через 19 дней метили  $10-20 \times 10^6$  клеток, а также проводили проточную сортировку, как описано в Примере 6. Отсортированные клетки выращивали в течение 17 дней и повторно анализировали с помощью проточной цитометрии (Фигура 10). Это показало, что большинство клеток в данном эксперименте характеризовались двойным положительным окрашиванием в отношении экспрессии IgG и связывания с β-галактозидазой.

Геномную ДНК получали из отсортированных клеток, а ДНК, кодирующую вставку IgG, выделяли с помощью ПЦР. Вставку с IgG-кодированием амплифицировали с использованием полимеразы KOD (Merck, номер по каталогу 71086-3) при температуре отжига  $60^\circ\text{C}$  и с применением 30 циклов. Был использован буфер от производителя с 5 % ДМСО с использованием  $0,3 \text{ мкМ}$  праймеров 2597 (SEQ ID NO: 54) и 2598 (SEQ ID NO: 47). Продукт требуемого размера подвергали очистке после разделения в агарозном геле. Продукт, очищенный после разделения в агарозном геле, затем использовали для ПЦР с использованием полимеразы KOD (Merck, номер по каталогу 71086-3) в буфере от производителя с 5 % ДМСО с использованием  $0,3 \text{ мкМ}$  праймера 2625 (SEQ ID NO: 55) в комбинации либо с праймером 1999 (SEQ ID NO: 56) (для образца R1), либо 2595 (SEQ ID NO: 53) (для 4R1 и 5R1) с использованием температуры отжига  $60^\circ\text{C}$  и 30 циклов. Продукты гнездовой ПЦР подвергали очистке после разделения в агарозном геле и подвергали двойному расщеплению с использованием NheI-HF (NEB, номер по каталогу R3131S) и XhoI (NEB, номер по каталогу R0146S), чтобы лигировать их с pINT3 (Фигура 1), который

подвергался двойному расщеплению аналогичным образом, для экспрессии растворимых форм связывающих молекул, получивших формат IgG. Последовательности праймеров:

2597: AGGGGTTTTATGCGATGGAGTT (SEQ ID NO: 85)

2598: GTTACAGGTGTAGGTCTGGGTG (SEQ ID NO: 78)

2625: CCTTGGTGCTGGCACTCGA (SEQ ID NO: 86)

1999: AAAAAGCAGGCTACCATGAGGGCCTGGATCTTCTTTCTCC (SEQ ID NO: 87)

2595: GAGGGCTCTGGCAGCTAGC (SEQ ID NO: 84)

#### Пример 8. Создание и отбор из библиотеки наивных scFv

В источнике Schofield и др. [7] описано создание библиотеки фагового дисплея (библиотека McCafferty), в которой гены антител из В-лимфоцитов целого ряда доноров были сначала клонированы с образованием «промежуточной библиотеки», а затем повторно клонированы с образованием окончательной функциональной библиотеки фагового дисплея. Эта та же самая промежуточная библиотека и методика были использованы для создания новой библиотеки библиотека (IONTAS) из  $4 \times 10^{10}$  клонов. Из этой библиотеки получали плазмидную ДНК, стараясь обеспечить достаточное количество молекул библиотеки в составе бактериального инокулята. Был установлен целый ряд ПЦР-реакций с использованием, в целом, 2 мкг матричной ДНК. ПЦР-продукт подвергали расщеплению с помощью NcoI и NotI, очищали после разделения в агарозном геле и лигировали, как описано в Примере 6. 9,3 мкг pD6 и 0,93 мкг полученной методом ПЦР вставки подвергали лигированию на протяжении ночи, продукты лигирования очищали с использованием экстракции фенолом и хлороформом, и клетки DH5-альфа подвергали электропорации ДНК, как было описано ранее [7]. В результате дисплея с использованием вектора, содержащего слитые гены scFv-Fc, была создана библиотека, состоящая из  $2,4 \times 10^8$  клонов. Из этой «библиотеки наивных клеток» получали ДНК, клонировали в pD6 и трансфецировали в клетки HEK293F (Life Technologies) в объеме 1 л, выращенные в среде Freestyle (как описано выше). 0,3 мкг ДНК библиотеки pD6, 0,5 мкг каждой пары TALEN «AAVS-SBI». Через 24 ч после трансфекции объем культуры увеличивался в два раза, и через 48 часов после трансфекции начинали проводить селекцию с использованием бластицидина, как описано выше. Размер библиотеки определяли путем осуществления посева аликвот культуры через 24 часа после трансформации и селекции с использованием бластицидина, как описано выше. Была создана библиотека из  $0,9 \times 10^7$  клонов.

Целый ряд антигенов подвергали биотинилированию с помощью набора EZ-Link SuIfo-NHS-LC-Biotin (Pierce номер по каталогу 21327) в соответствии с инструкциями изготовителя. В качестве антигенов использовали бычий тиреоглобулин (Calbiochem,

номер по каталогу 609310), рекомбинантный CD28-Fc человека (R and D Systems, номер по каталогу 342-CD-200) и рекомбинантный EphB4-Fc мыши (R and D Systems, номер по каталогу 446-B4-200). Также использовали биотинилированную  $\beta$ -галактозидазу (Rockland, номер по каталогу B000-17).

Селекцию трансфецированных клеток производили в суспензионной культуре с использованием бластицидина в течение 17 дней, как описано выше. Клетки собирали, промывали и доводили концентрацию до  $15\text{-}20 \times 10^6$  клеток на пробу. Клетки были получены, как описано, и биотинилированный антиген добавлен в концентрации 500 нМ. Мечение и проточную сортировку проводили, как описано выше. Используя контрольные клетки, инкубированные только с меченым фикоэритрином антителом к Fc-фрагменту, был получен «гейт», который включал 0,05 % этих клеток. Используя тот же гейт для меченых клеток были включены от 0,28 до 0,51 % клеток (Фигура 11). Клетки собирали и выращивали для использования в дополнительных раундах сортировки и амплификации генов scFv из библиотеки наивных молекул.

Пример 9. Создание клеточной линии с «сайтами множественной посадки» для сравнения подходов, основанных на использовании нуклеаз и рекомбиназ, к встраиванию в геном

Для осуществления сравнения способов встраивания на основе либо расщепления геномной ДНК, либо встраивания, опосредованного рекомбиназой, был сконструирован направляемый AAVS вектор (pD4), который обеспечивает введение интрона с множественными «сайтами посадки» (Пример 3). К ним относятся сайт FRT, распознаваемый рекомбиназой Flp, и пара сайтов lox2272/loxP, распознаваемых рекомбиназой Cre. Для обеспечения целенаправленного расщепления в состав pD4 также включили последовательность из GFP, для которой была сконструирована пара TALEN [128], и сайт мегануклеазы I-Sce 1 для обеспечения встраивания, направляемого эндонуклеазой. Была сконструирована совместимая вносимая донорная плаزمиды (pD5) с соответствующими сайтами узнавания таким образом, что встраивание, направляемое нуклеазой или рекомбиназой, приводит к активации беспромоторного гена устойчивости к бластицидину и встраиванию кассеты, экспрессирующей антитела.

Структура плазмиды и последовательности pD4 показаны на Фигуре 12b. Вначале была создана промежуточная плаزمиды pD3, которая включает ген GFP под контролем CMV-промотора, за которым следуют слитые гены устойчивости к пуромицину/тимидинкиназы под контролем промотора PGK (Фигура 12a). Она был создана путем расщепления pBIOCAM1-newNot с использованием SacI (в конце CMV-промотора) и BstBI (между геном Neo поли(A)-последовательностью, Фигура 1a). Это позволило

удалить экспрессирующую кассету с геном устойчивости к неомицину и позволяет производить замену на синтетическую вставку, содержащую ген усиленного зеленого флуоресцентного белка (EGFP) под контролем CMV-промотора. Эта конструкция с геном GFP была слита на С-конце с остатками 422-461 мутантной формы последовательности PEST орнитин-декарбоксилазы мыши. Эта последовательность PEST была включена в плазмиду pZsGreen1-DR (Clontech), и, как было показано, обуславливала уменьшение периода полужизни слитого GFP до 1 часа. Кассету, кодирующую промотор PGK, слитые гены устойчивости к пурамицину/тимидинкиназы (Puго deltaTK) и кассету с поли(A)-последовательностью вырезали из плазмиды pFLEXIBLE [129] с использованием Xmn1 и FSE1, а также клонировали в сайты Sma1 и FSE1, содержащиеся в исходной синтетической вставке. Полученная в результате плазида (называемая pD3) кодирует ген GFP под контролем CMV-промотора и ген устойчивости к пурамицину под контролем промотора PGK.

Для создания окончательного варианта нацеливающего вектора pD4, удаляли CMV-промотор и осуществляли вставку левого гомологичного участка AAVS. Фрагмент локуса AAVS длиной 850 п.о. амплифицировали методом ПЦР для получения левого гомологичного участка AAVS, фланкированного EcoR1 на 5'-конце и Mre1 на 3'-конце. Его клонировали в сайт EcoR1/Mre1 из pD3, что приводило к удалению CMV-промотора. Сайт Nsi1 был также включен на 3'-конце этого левого гомологичного участка AAVS. Соседние участки Mre1 и Nsi1 использовали для введения синтетического фрагмента, путем слияния интрона с геном EGFP, как показано на Фигуре 13. Синтетический интрон, предшествующий гену EGFP, включает в себя:

- сайт распознавания для рекомбиназы FRT
- сайт рекомбинации Iox 2272
- сайт мегануклеазы I-Sce1
- сайт распознавания GFP TALEN
- последовательность T2A для остановки рибосом [130]

Правый гомологичный участок AAVS был получен с помощью ПЦР для создания сайтов Hpa1 и BstZ171 на 5'- и 3'-концах. Этот фрагмент клонировали в сайты Hpa1 и BstZ171 из pD3. Полученная в результате плазида pD4 кодирует кассету с геном устойчивости к пурамицину («Puго deltaTK») и может использоваться для введения «сайтов посадки» в локус AAVS, включая для сравнения различные сайты нуклеаз и рекомбиназ. Последовательность pD4 показана на Фигуре 13 (SEQ ID NO: 5, SEQ ID NO: 6 и SEQ ID NO: 7). Левый и правый участки для целенаправленного встраивания AAVS, подробно показаны на Фигуре 3, а соответствующие сокращения на Фигуре 13.

Инtron, введенный в pD4, содержит сайты распознавания TALEN, полученные из GFP [128]. Пара субъединиц TALEN, направленно связывающихся с eGFP (eGFP-TALEN-18-левый и eGFP-TALEN-18-правый), распознает последовательность, показанную ниже (все последовательности и праймеры представлены в 5'-3' направлении), где заглавные буквы соответствуют сайту узнавания левого и правого участков TALEN, спейсерная последовательность показана с использованием нижнего регистра. Правый участок TALEN распознает последовательность, комплементарную показанной последовательности. Первая пара оснований эквивалентна положению -16 в данной последовательности по отношению к последовательности иницирующего кодона ATG из GFP (подчеркнута в спейсерном участке).

TCCACCGGTTCGCCAccatggtgagcaagggCGAGGAGCTGTTCA (SEQ ID NO: 88)

Плазмида pD4 также включает в себя сайт мегануклеазы I-SceI и сайт FRT, распознаваемый рекомбиназой Flp. В конечном итоге, pD4 включает Iox 2272 и IoxP (которые несовместимы друг с другом), фланкирующие GFP и кассеты, экспрессирующие гены устойчивости к пуромицину. Включение таких же 2 сайтов IoxP, фланкирующих донорную плазмиду (pD5, показана ниже), позволяет произвести замену встроенной кассеты (включая кассету PGK риго delta TK), путем замены ее вносимой кассетой, контролирующей экспрессию гена устойчивости к бластицидину и антител с помощью опосредованного рекомбиназой кассетного обмена.

#### *Создание клеточной линии путем трансфекции pD4.*

Клетки HEK293F ресуспендировали в количестве  $10^6$  клеток/мл, добавляли ДНК: полиэтиленимин (PolyPlus) в соотношении 1:2 (мас/мас). Клетки трансфецировали 0,6 мкг/мл pD4 и котрансфецировали либо с использованием пары «исходная плазмида с последовательностями AAVS» для TALEN, либо с использованием pCDNA3.0, в качестве контроля (0,6 мкг/мл). pD3, который экспрессирует EGFP в составе CMV-промотора был включен в эксперимент в качестве контроля трансфекции, и эффективность трансфекции с его использованием составляла 35 %. Через 24 часа трансфецированные клетки высевали в количестве  $0,5 \times 10^6$  клеток/10см чашку Петри диаметром (тканевую культуру обрабатывали) и выращивали в минимальной питательной среде, содержащей 10 % фетальной бычьей сыворотки (10270-106, Gibco) и 1 % заменимых аминокислот (MEM\_NEAA № 11140-035 Life Technologies). 5 мкг/мл пуромицина добавляли еще через 24 ч и производили замену среды каждые 2 дня. Через 5 дней нетрансфецированные клетки или клетки, трансфецированные только с pD3, были мертвы. Через 12 дней было приблизительно 200 колоний на клетках, трансфецированных только с pD4 и приблизительно 400 колоний на клетках, трансфецированных с pD4 и парой AAVS TALEN.

Устойчивую популяцию к пурамицину, возникающую из трансфекции с рD4 и парой AAVS TALEN, анализировали для точного встраивания. Кроме того, одну колонию собирали с этой популяции (клон 6F) и сравнивали с колонией из устойчивой популяции к пурамицину, возникающей из рD4 трансфекции при отсутствии пары AAVS TALEN. Для правильной идентификации правильного встраивания, были созданы ПЦР-праймеры, которые гибридизуются в геномном локусе AAVS за пределами левого и правого гомологичных участков. Они были в паре со вставкой специфических праймеров. Праймеры на 5'-конце:

AAVS1\_HA-L\_Nested\_Forw1 GTGCCCTTGCTGTGCCGCCGGAАCTCTGCCCTC

(SEQ ID NO: 89)

EGFP\_Synthetic\_gene\_Rev\_Assembly TTCACGTCGCCGTCCAGCTCGAC (SEQ ID NO: 90)

Purotk\_seq\_fow2 TCCATACCGACGATCTGCGAC (SEQ ID NO: 91)

AAVS 1\_Right\_arm\_Junction\_PCR\_Rev TCCTGGGATACCCCGAAGAG (SEQ ID NO: 77)

На Фигуре 14 показано, что клон 6F и популяция, исправляют оба левых и правых участков, но клон отобранный из AAVS не направленной популяции является отрицательным. Таким образом, ПЦР-анализ указывает на то, что точность встраивания донорной кассеты больше, когда направляемый AAVS TALEN расщепляет геномную ДНК.

рD4 внедряет беспромоторный ген в рамке считывания гена GFP, приводимый от промотора AAVS. Проточная цитометрия устойчивых популяций к пурамицину, показала отсутствие экспрессии GFP.

Эта неспособность к экспрессии может быть связана с сочетанием короткого периода полужизни (определяемого элементом последовательности PEST орнитин-декарбоксилазы мыши) в сочетании с пониженной экспрессией в результате применения промотора T2A. На самом деле было обнаружено, что добавление элемента T2A перед беспромоторным элементом гена устойчивости к бластицидину (как это было описано для рD2) приводит к уменьшению количества колоний, устойчивых к бластицидину, в 4 раза. Несмотря на отсутствие экспрессии GFP, тем не менее, встраивание многочисленных сайтов посадки дает возможность для сравнения встраивания в геном, обеспечиваемого рекомбиназой и расщеплением ДНК.

Пример 10. Создание вектора для введения кассеты антитела (рD5) в сайт "множественной посадки"

После введения интрона из «сайта множественной посадки» в локус AAVS, существует возможность ввести кассету антитела через средства, направленного нуклеазой или направленного рекомбиназой. Для этого была создана донорная плаزمиды рD5, где произведена экспрессионная кассета, фланкированные левые и правые гомологичные

участки, которые равнозначны фланкирующим последовательностям сайта расщепления GFP TALEN, введенными в pD4. pD5 сам по себе не включает целый участок сайта распознавания GFP TALEN и введение происходит с помощью гомологичной рекомбинации. Встраивание донорной плазмиды на основе гомологичной рекомбинации приводит к введению гена устойчивости к бластицидину, в котором отсутствует промотор, но которому предшествует акцепторный сайт сплайсинга, что приводит к слиянию в рамке считывания с расположенным выше экзона из локуса AAVS, как описано ранее. Встраивание в локусе AAVS вызывает экспрессию беспромоторного гена устойчивости к бластицидину. Введенная кассета также кодирует антитело IgG, с измененным форматом, тяжелые и легкие цепи под контролем PEF и CMV промоторов, соответственно, как описано выше. pD5 включает в себя сайт мегануклеазы I-Sce1, который может привести к расщеплению входящего донора, дающий возможность для NHEJ (см. Пример 12). Сайт FRT также включен в донорную плазмиду pD5, позволяет включение, направленное рекомбиназой, к беспромоторному гену устойчивому к бластицидину и антителу экспрессирующих кассет в том же самом локусе. Как уже говорилось выше, Cre-рекомбиназа будет действовать на сайтах IoxP в доноре и геномной ДНК опосредованной рекомбиназой кассетного обмена.

Последовательность pD5 показана на Фигуре 15. Последовательность на 5'-конце сайта GFP TALEN из локуса AAVS. Секция 267 п. о. локуса AAVS выше от сайта расщепления TALEN образована с помощью ПЦР. Использовали праймеры, которые создали EcoR1 и Mfe 1 сайт на 5 'и 3' концах, и продукт клонировали в сайт EcoR1 из pD1-D1.3. Сайт EcoR1 воссоздавали на 5 'конце. Во время клонирования левое гомологичный участок Nsi1 и Pac1 также вставляли в 3 'конце. Правый гомологичный участок, включающий примерно 700 п.о., равнозначной последовательности 3' GFP TALEN, создали путем сборки ПЦР. ПЦР-праймеры вводили на 5 'и на 3'-конец сайта BstZ171 к собранному фрагменту и его клонировали в сайт BstZ171 из pD1-D1.3. ПЦР-праймеры также вводили на 5 'конец сайта Hpa1.

ПЦР-фрагмент, охватывающий интрон (который включает в себя сайты узнавания для GFP TALEN, эндонуклеазы I-Sce1, рекомбиназы Flp и рекомбиназы Cre), акцепторную область сплайсинга, устойчивый ген к бластицидину и поли(А)-сайт (описанный выше), был создан на 5 'конце сайта Nsi1 и на 3'-конце сайта Pac1. Его клонировали в сайт Nsi1 и Pac1 из плазмиды, описанной выше, для создания pD5-D1.3 (последовательность показана на Фигуре 15, а плазмидная конструкция показана на Фигуре 18а).

Пример 11 Сравнение направляемого нуклеазой и направляемого Flp встраивания кассеты, экспрессирующей антитела

Система Flp-In, которую ранее использовали для встраивания, опосредованного рекомбиназой, антитела экспрессирующих кассет [18] использует мутантную Flp-рекомбиназу (в плазмиде pOG44), которая обладает только 10 % активности при 37 °C в нативной Flp-рекомбиназе [19]. Был определен вариант Flp-рекомбиназы (Flpe) с более выраженной термостабильностью и активностью при 37°C, чем дикий тип [19, 20]. В дальнейшем это усовершенствовали за счет оптимизации кодона для создания Flpo [131], закодированного в составе плазмиды cCAGGS-Flpo (Genebridges номер по каталогу A203). Сравнивали влияние обоих вариантов Flp-рекомбиназы (закодированной в pOG44 и cCAGGS- Flpo). Рекомбинацию направленной Cre-рекомбиназы, также проверяли с помощью трансфекции клеток с плазмиды, которая кодирует Cre-рекомбиназу [132] (pCAGGS-Cre, Genebridges номер по каталогу A204). В каждом векторе рекомбиназа экспрессируется под контролем промотора  $\beta$ -актина курицы непосредственно раннего энхансера CMV. Большая T-последовательность внутриядерной локализации SV40 используется для внутриядерной локализации [20]. В исходных векторах (cCAGGS-Flpo и pCAGGS-Cre) экспрессия рекомбиназы связана с геном устойчивости к пуромицину при помощи внутренним сайтом рибосомного входа (IRES), который был удален с использованием стандартных методов молекулярной биологии.

Эксперимент проводили для сравнения эффективности геномного направленного расщепления по сравнению с встраиванием, направленным рекомбиназой, антитела кассеты. Результаты оценивали 2 способами:

1. Измерение количества устойчивых колоний к гену устойчивости к бластицидину, возникающих при введении беспромоторного гена устойчивости к бластицидину
2. Оценка степени экспрессии антитела осуществляется за счет различных подходов.

Как описано в Примере 9, сайты узнавания для Cre-рекомбиназы (lox2272 и loxP) и Flp-рекомбиназы (FRT) были ранее встроены в локус AAVS в клоне 6F. Кроме того, сайты распознавания для пары GFP TALEN и для сайта мегануклеазы I-Sce1 также присутствуют в пределах одного интрона. Донорная плаزمида pD5-D1.3 несет те же сайты распознавания (кроме GFP TALEN) внутри интрона выше от беспромоторного гена устойчивости к бластицидину. Правильное встраивание приводит к активации гена устойчивости к бластицидину. pD5-D1.3 также отвечает за экспрессируемый на клеточной поверхности ген антитела D1.3, получившего формат IgG.

Котрансфекция из pD5-D1.3 с pOG44 или pCAGGS- Flpo (2 варианта кодирования Flp-рекомбиназой) должна привести к встраиванию всей плазмиды pD5 в сайт FRT из клона



6F. Донорная плаزمида pD5-D1.3 также имеет сайт lox2272 внутри синтетического интрона выше от гена устойчивости к бластицидину и сайт loxP в конце антителиа экспрессирующей кассеты. Под действием экспрессии, Cre-рекомбиназа из pCAGGS-Cre кассетного обмена, опосредованного рекомбиназой, должна привести к встраиванию гена устойчивости к бластицидину и экспрессии кассеты антителиа в сайты lox2722 и loxP в пределах клона 6F.

Эффективность от встраивания вектора с использованием методов, направленных рекомбиназой, с методами, направленными геномным расщеплением, сравнивали с использованием пары TALEN (EGFP-TALEN-18-Левый и EGFP-TALEN-18-Правый) относительно последовательности из GFP (Reyon и *et al.*, 2012). В случае элемента GFP TALEN, между левым и правым гомологичными участками будет встраиваться следующее геномное расщепления TALEN.

Для того, чтобы провести сравнение с сайтом мегануклеазы I-Sce1, был построен кодон оптимизированного гена, кодирующий сайт I-Sce1 (Фигура 16). Этот ген имеет N-конец эпитопа HA, tag/SV40, ядерную локализацию сигнала (NLS) на N-конце и фланкирующие сайты Nco1 и Xba1, на 3' и 5' концами. Ген клонировали в вектор pSF-CMV-F1-Pac1 (Oxford Genetics OG111), где запускается экспрессия от промотора CMV.

Трансфекцию проводили с использованием клона 6F с правильно встроенным «сайтом множественной посадки». Клетки суспендировали в количестве  $10^6$  мл и трансфецировали с 50 нг pD5-D1.3 донорной плазмиды/ $10^6$  клеток наряду с ферментами, кодирующими плазмид (Таблица 4А).

Через 24 часа трансфецированные клетки высевали в чашки Петри в количестве  $0.5 \times 10^6$  клеток/10 см (тканевую культуру обрабатывали) и выращивали в минимальной питательной среде, содержащей 10 % фетальной бычьей сыворотки (10270-106, Gibco) и 1 % заменимых аминокислот (MEM\_NEAA № 11140-035 Life Technologies). 5 мкг/мл бластицидина добавляли еще через 24 ч и производили замену среды каждые 2 дня. Через 12 дней планшеты окрашивали 2 % метиленовым синим (в 50 % метанола) и подсчитывали численность колоний (Таблица 2). При непосредственном сравнении между Flp-рекомбиназой, Cre-рекомбиназой и TALEN, наибольшее количество колоний получены с использования GFP TALEN, где произошло увеличение в 9 раз по сравнению с контролем, содержащим «только донорную ДНК» (Таблица 4А). Кроме того, обнаруживается, что использование оптимизированного гена Flp фактически привело к сокращению количества колоний, устойчивых к бластицидину, по сравнению с контролем, содержащим «только донорную ДНК», возможно за счет токсичности повышенной активности Flp-рекомбиназы. Было также увеличение количества колоний с использованием pCAGGS-Cre по сравнению с только донорным контролем.

Второй эксперимент проводили в сравнении GFP TALEN с повышенной Flp (от cCAGGS-Flp), а также с низкой активностью Flp-фермента, кодируемый в pOG44 из системы Flp-In (как использует Zhou и др. [17, 18, US7, 884054]. Сравнивали GFP TALEN и Cre-рекомбиназу (Таблица 4В). Клетки трансфицированы с количеством ДНК, показаны на миллион клеток.  $0,25 \times 10^6$  клеток высевали и определяли количество колоний, устойчивых к бластицидину, как описано выше. Клетки также отбирали для устойчивости к бластицидину в жидкую культуру в течение 30 дней, прежде чем определить долю клеток, которые экспрессируют IgG на поверхности (как описано выше). В Таблице 4В показано, что TALEN превосходил другие подходы в исчислении количества устойчивых колоний. Опять же, использование оптимизированных Flp внутри cCAGGS-Flp фактически привело к сокращению количества колоний, устойчивых к бластицидину, по сравнению контролями, содержащими «только донорную ДНК». Использование Cre-рекомбиназы снова приводило к увеличению количества колоний, устойчивых к бластицидину, по сравнению с контролем, тогда как при использовании гена Flp в составе pOG44 было показано лишь незначительное увеличение по сравнению с контролем.

Таблица 4. Сравнение подходов, основанных на встраивании, направляемом нуклеазой TALE и направляемом рекомбиназой.

А.

Плазмида, несущая ген фермента	Количество	Количество донорной плазмиды (нг/ $10^6$ клеток)	Количество дозорной рD5-D1.3 (нг/ $10^6$ клеток)	Колонии, устойчивые к бластицидину (колонии/ $10^6$ клеток)
Пара GFP TALEN	0,575 мкг каждой	50	50	152 (304)
pCAGGS-Flp	1,15 мкг	50	50	1 (2)
pCAGGS-Cre	1,15 мкг	50	50	57(114)
контроль (pCDNA3.0)	1,15 мкг	50	50	17(34)

В.

Каждый образец в количестве 2 мкг на $10^6$ клеток	Количество донорной плазмиды рD5-D1.3 мкг/ $10^6$ клеток	Колонии, устойчивые к бластицидину,	Колонии, устойчивые к бластицидину, на $10^6$ клеток (процент колоний, устойчивых к бластицидину)	Процент положительных клеток
Пара GFP TALEN	0,6	270	1080 (1,1 %)	95,6

cCAGGS- Flp <sub>o</sub>	0,6	35	140 (0,14 %)	Слишком мало клеток
pOG44	0,6	96	384 (0,38 %)	6,4
pCAGGS-Cre	0,6	180	720 (0,72 %)	4
Контроль (PCDNA3.0)	0,6	81	324 (0,32 %)	37,3

с.

Каждый образец в количестве 2 мкг на 10 <sup>6</sup> клеток	Количество донорной плазмиды pD5-D1.3 пг/10 <sup>6</sup> клеток	Колонии, устойчивые к бластицидину,	Процент положительных клеток
GFP TALEN	2	210	95,3
GFP TALEN	6	120	73,6
cCAGGS- Flp <sub>o</sub>	2	62	Слишком мало клеток
cCAGGS- Flp <sub>o</sub>	6	41	Слишком мало клеток
pOG44	2	178	35,6
pOG44	6	63	58,7
pCAGGS-Cre	2	84	40,9
pCAGGS-Cre	6	52	Слишком мало клеток
Контроль (PCDNA3.0)	2	340	65,6
Контроль (PCDNA3.0)	6	82	55,5

С добавлением большего количества ДНК-донора (Таблица 4С) было отмечено увеличение численности колоний на промежуточных уровнях (2 мкг/миллион клеток) и снижение на более высоких уровнях по всем направлениям (6 мкг/млн клеток). Ни один из других образцов не достиг уровня дисплея антитела со встраиванием, направленным GFP TALEN.

Клетки были также отобраны для устойчивости к бластицидину в жидкой культуре и клетки окрашивали для экспрессии антитела, как описано выше. Встраивание направленного TALEN дало значительно более высокую долю антител-положительных клеток, по сравнению с другими подходами. Клетки, трансфицированные cCAGGS- Flp<sub>o</sub> и с высокими концентрациями pCAGGS-Cre, не были здоровы и были в недостаточном количестве для проведения проточной цитометрии.

Сравнение распространялось на включенный сайт эндонуклеазы I-SceI. Синтетический ген, кодирующий I-SceI, синтезировали (Фигура 16) и клонировали в 1 сайт NcoI/Xba из pSF-CMV- f1-PacI (Oxford Genetics). Клетки суспендировали в  $10^6$ /мл и трансфецировали на каждый мл клеток ( $10^6$  клеток/мл) в присутствии 300 нг донорной плазмиды pD5-D1.3 вместе с плазмидами, кодирующими ферменты (1 мкг /  $10^6$  клеток). На следующий день 0,05 мл клеток высевали и отбирали в бластицидин, и окрашивали через 14 дней, как было описано ранее. Таблица 5 показывает, что наибольшее количество колоний устойчивых к бластицидину происходили из сайта мегануклеазы I-SceI с последующей парой EGFP TALEN. Обе Cre- и Flp-рекомбиназы (закодированы внутри pOG44) дали количество немного выше, чем контроль, содержащий «только донорную ДНК». По-прежнему трансфекция Flpe, кодирующая плазмиды, фактически уменьшила количество колонии по сравнению с контролем, содержащим «только донорную ДНК».

Таблица 5. Сравнение подходов, основанных на встраивании, направляемом мегануклеазой и направляемом рекомбиназой

Каждый образец в количестве 20 мкг на $10^6$ клеток	Количество донорной плазмиды pD5-D1.3 мкг/ $10^6$ клеток	Колонии, устойчивые к бластицидину,	Колонии, устойчивые к бластицидину, $10^6$ клеток (процент колоний, устойчивых к бластицидину)	Процент положительных клеток, день 7	Процент положительных клеток, день 13
Пара GFP TALEN	0,6	90	1800	27,6	55 %
I-SceI мегануклеаза		150	3000	29,9	47 %
pOG44	0,6	60	1200	2,79	6,5 %
pCAGGS-Cre	0,6	56	1120	2,95	6,6 %
cCAGGS- Flp <sub>o</sub>	0,6	4	80	5,9 (низкое количество клеток)	Слишком мало клеток
Контроль (PCDNA3.0)		40	800	3,3	4,9 %

(SBI AAVS в клетках HEK дикого типа		251 (x2)	10040(1 %)	ND	ND
---	--	----------	------------	----	----

После трансфекции большую часть клеток отбирали для устойчивости к бластицидину в жидкой культуре, а после 7 и 13 дней окрашивали с Fc-фикоэритрин меченым антителом, как описано выше. На Фигуре 17 (суммировано в Таблице 5) показана наиболее высокая экспрессия антител, полученная для клеток, трансфецированных I-Sce1 эндонуклеазой и eGFP TALEN (47 % и 55 % соответственно) по сравнению с "только донорными" (4,9 %). В отличие от этого, процент позитивных клеток антител, когда клетки котрансфецировали плазмидами, кодирующими Flp-рекомбиназу (pOG44) или Cre-рекомбиназу, был 6,6 % и 6,5 % соответственно. Доля позитивных клеток антител продолжает увеличиваться при непрерывном отборе в бластицидин и достигает 85-90 % положительных антител, в случае I-Sce1 и EGFP TALEN образцы трансфецировали на 19 день, когда производили анализ. Таким образом, мегануклеазы обеспечивают альтернативный подход к эффективному встраиванию, направленного нуклеазой, из антитело-кодирующих трансгенов.

Пример 12. Направляемое нуклеазой встраивание кассеты с генами антител может происходить как путем гомологичной и, так и NHEJ

Эффективность встраивания трансгенов в клеточную ДНК может быть повышена за счет введения двухцепочечных разрывов (DSB). Эндогенные механизмы репарации ДНК в эукариотических клетках, включают гомологичную рекомбинацию негомологичного конца соединения (NHEJ) и ее варианты. Все обеспечивают средства для введения генов, кодирующие связывающие молекулы внутри библиотеки. Гомологичная рекомбинация обеспечивает точное соединение между гомологичными участками и вставленными трансгенами, но требует предоставления гомологичных участков в донорную плазмиду. ДНК для гомологичной рекомбинации может предоставляться в виде линейной или кольцевой ДНК. При использовании NHEJ происходит повторное лигирование концов ДНК в отсутствие необходимости в наличии гомологичной матрицы. Данный подход к репарации ДНК является менее точным и может привести к вставлению или удалению. NHEJ, тем не менее, обеспечивает простой способ встраивания в рамке считывания экзонов в интрон или допускает встраивание промотора: кассеты гена в геном. Использование негомологичных способов позволяет использовать донорные векторы, которым недостаточно гомологичных участков, тем самым упрощая конструкцию донорной ДНК.

Клон 6F имеет GFP TALEN и сайты узнавания нуклеазы I-Sce1, встроенные в геном, и они будут расщепляться, когда эти нуклеазы предоставляться. Донорный вектор pD5 не имеет сайта распознавания нуклеазы GFP TALE, но имеет гомологичные участки фланкирующего сайта расщепления и таким образом, ожидается встраивание путем только гомологичной рекомбинации. Расщепление геномной ДНК в соседнем сайте мегануклеазы I-Sce1, также приведет к встраиванию элементов pD5 путем гомологичной рекомбинации. Однако pD5 также имеет сайт мегануклеазы I-Sce1, который может быть расщеплен *in vivo*, когда обеспечивается наличие I-Sce1. Это создаст линейный продукт ДНК, который потенциально может встраиваться с NHEJ. Как описано ранее, это может быть даже эффективными преимуществами при использовании *in vivo* расщеплении донорной ДНК, когда используется NHEJ.

На Фигуре 18a представлена донорная ДНК в pD5-D1.3, а на Фигуре 18b представлен геномный локус клеток клона 6F, включающий «сайт множественной посадки». На Фигуре 18c представлены результаты гомологичной рекомбинации между pD5-D1.3 (Фигура 18a) и сайтом множественной посадки клона 6F (Фигура 18b). В отличие от этого на Фигуре 18d представлены результаты NHEJ. В этом случае включена дополнительная ДНК из остова вносимой плазмиды (представлена двойной стрелкой). Flp-опосредованная рекомбинация в сайте «множественной посадки» приводит к образованию аналогичного продукта. Для того, чтобы определить по какому пути используется с образцами, описано в Примере 11 (показано на Фигуре 17), геномную ДНК получали из отобранной популяции с бластицидином, как было описано выше. Разработали обратный ПЦР-праймер (J44), который гибридизуется со встроенным РГК-промотором. Это использовали в сочетании с любой J48, которая гибридизуется в конце белка IgG. Праймеры J44 и J48 разработаны, чтобы выявить гомологичную рекомбинацию, создающую полосу 1928 п.о, когда I-Sce1 отвечает за встраивание (обозначено стрелкой на Фигуре 18e). (Потенциальная полоса 5131 п.о. может быть создана с помощью этой пары праймеров, когда осуществился NHEJ, но этот долгий продукт не виден в геноме методом ПЦР этого эксперимента.).

Праймер J46 разработан для гибридизации внутри гена  $\beta$ -лактамазы в пределах основного вектора. Праймеры J44 и J46 ускоряют создание полосы из 1800 п. о., когда осуществился NHEJ. Предполагаемая полоса аналогичного размера, где Flp-рекомбиназа приводит к встраиванию опосредованного рекомбиназой.

J44: AAAAGCGCCTCCCCTACCCGGTAGAAT (SEQ ID NO: 92) J46:  
GGCGACACGGAAATGTTGAATACTCAT (SEQ ID NO: 93) J48:  
CACTACACCCAGAAGTCCCTGAGCCTG (SEQ ID NO: 94)

На Фигуре 18е ясно показано, что гомологичная рекомбинация осуществляется только с образцами, обработанными с GFP TALEN и мегануклеазой I-Sce1 ((I и II по сравнению с III и IV). В противоположность этому, NHEJ осуществляется только тогда, когда расщепление совершается с помощью мегануклеазы I-Sce1 (Фигура 18е V), но не с GFP TALEN (Фигура 18е VI). Как и следовало ожидать, подобный размер полосы обнаруживается в образце обработанным с Flp-рекомбиназой (Фигура 18е VII). Таким образом, этот эксперимент показывает, встраивание, направленного нуклеазой, кассеты антитела может осуществляться одновременно гомологичной рекомбинацией и NHEJ.

Пример 13. Получение секретлируемых и мембраносвязанных фрагментов антител с использованием одной и той же клетки

Как было описано выше, дисплей векторов млекопитающих pD2 и pD5 сконструирован с экзоном, кодирующий трансмембранный домен фланкирующих между двумя сайтами узнавания ROX, признаваемых Dre-рекомбиназой [88]. Для того чтобы определить, можно ли преобразовать из мембраносвязанной формы в секретлируемую форму популяции устойчивые к бластицидину, вытекающие из трансфекции пары pD2-D1.3/AAVS TALEN, повторно трансфецировали плазмидой, кодирующей Dre-рекомбиназу (pCAGGs-Dre). Это было основано на использовании плазмиды pCAGGs-Dre-IRES puго [88], которая контролирует ген Dre-рекомбиназы в составе промотора CAGGs (GeneBridges A205). Ген устойчивости к пурамицину удалили с использованием стандартных методов молекулярной биологии. Через 22 дня после отбора с бластицидином, клетки раскладывали на расстоянии  $0,5 \times 10^6$  клеток/мл и трансфецировали, как описано ранее, с  $0,5 \mu\text{кг}$  pCAGGs-Dre на  $10^6$  клеток. После 6 дней собрали супернатанты, антитела очищали с использованием белка-A и образцы подвергали электрофорезу на геле SDS-PAGE и окрашивали Кумасси синим. На Фигуре 19а показано, что секретлируемое антитело обнаружено в супернатанте даже без трансфекции с геном Dre-рекомбиназы. Это может возникнуть из альтернативного сплайсинга, где экзон, кодирующий трансмембранный домен, пропускается. В качестве альтернативы антитело в культуре супернатанта может возникнуть в результате расщепления мембраносвязанного антитела. Трансфекция из Dre-рекомбиназы повысила уровень секретлируемого антитела (Фигура 19а).

Производство секретлируемого слияния scFv-Fc также продемонстрировали в эксперименте, описанного в Примере 7 (Фигура 9h). Популяции антител scFv, выбранные с помощью фагового дисплея из 1 раунда на  $\beta$ -галактозидазе, ввели в вектор pD6 и встраивали в локус AAVS клеток HEK293, используя AAVS TALEN. Антигенсвязующие клетки сортировали по размеру и выбранные клетки выращивали в течение 7 дней после отбора без замены среды, чтобы обеспечить увеличение количества антител. Планшеты

ИФА покрывали либо  $\beta$ -галактозидазой (10 мкг/мл) или БСА (10 мкг/мл) на протяжении ночи. Культуральные супернатанты, из 7 дневных культур, смешивали с 50 % объемом 6 % MargeI-PBS и тестируемого образца в трех экземплярах. 1/10 разбавление также тестировали. Проводили выявление связанного слияния scFv-Fc с использованием антитела IgG-EU человека (Perkin Elmer номер по каталогу 1244-330). На Фигуре 19b показано, что связывание антитела можно выявить непосредственно из любой культуры супернатантов в чистом виде или с разбавлением 1/10. Это показывает, что обе поверхности дисплея и секрета антитела может достигаться в пределах одних и тех же клеток без дополнительных стадий. Это дает возможность получить секретируемое антитело непосредственно из выбранных клеток любого из следующего клонирования одной клетки или используя выбранную популяцию, как показано здесь, создавая поликлональную смесь антитела.

#### Пример 14. Простой способ придания молекулам scFv формата IgG или Fab

Новый способ изобретен для эффективного превращения антител получившими форму, как scFv с IgG форматом, как описано в Примере 7. Это преобразование является необходимым процессом на протяжении проектов, направленных на разработку новых антител, в которых используются библиотеки фагового дисплея scFv, где в качестве окончательного формата требуется антитело, получившее формат IgG или Fab. Современные способы являются трудоемкими и вовлекают отдельное клонирование переменного участка тяжелых ( $V_H$ ) и переменного участка легких ( $V_L$ ) цепей в соответствующие векторы экспрессии. Более того, превращение популяции из scFvs «*en masse*» (целиком) не возможно, так как теряется связь между цепями  $V_H$  и  $V_L$ . Это является проблемой, потому как  $V_H$  и  $V_L$  цепи способствуют специфичности связывания с антигеном. В настоящее время неспособность легко превращать популяции из scFv в Ig или Fab формат, ограничивает возможность скрининга большого количества антител, в окончательном формате они будут использоваться в клинике. Способность скрининга рекомбинантных антител в Ig или Fab формате для связывания мишени, скрининга сборки клетки и биофизических свойств и функций, включая агрегатное состояние, является необходимым шагом для выбора кандидата средств антитела в качестве клинических кандидатов. Чем больше количество антител, протестированных на данном этапе, в IgG или Fab-формате, тем больше вероятность выбора лучшего кандидата средства антитела.

Описываемый здесь способ для превращения популяций одноцепочечного антитела (scFv) в иммуноглобулин (Ig) или Fab формат таким образом, чтобы сохранить изначальное попарное объединение тяжелой ( $V_H$ ) и легкой ( $V_L$ ) цепей переменного домена. Способ позволяет превращения моноклональные, олигоклональные или поликлональные scFv



одновременно с Ig или Fab-форматом. Предпочтительно, способ включает получение «миникольцевой ДНК» нерепликативным путем. Предпочтительно, полный процесс превращения влечет за собой единую трансформацию бактерий, таких как *E.coli*, для создания популяции из бактериальных колоний, каждая из которых несет плазмиду, кодирующую уникальное рекомбинантное антитело, получившее формат Ig или Fab. Это отличается от альтернативных способов, требующих двух отдельных шагов клонирования и трансформации [117].

В более широком плане этот аспект изобретения относится к способу превращения генетической конструкции с 3-связанными генетическими элементами А, В и С (представленные с помощью  $V_H$ , линкером и соответственно  $V_L$ , в случае с scFv) в формате, где порядок фланкирующих элементов (А и С) поменяли местами в одной стадии клонирования. Промежуточный элемент может быть сохранен, но с наибольшей пользой способ позволяет замену этого промежуточного элемента новым элементом D (чтобы дать C-D-A). В примере превращения scFv в IgG или Fab, когда С представляет собой антитело  $V_L$  домена и А представляет собой домен  $V_H$ . В этом примере элемент D инкапсулирует константный домен легкой цепи, поли-(А) сайт, промотор и лидерную последовательность, слитую с доменом  $V_H$  (элемент А). В процессе продукт (C-D-A) клонировали повторно, что приводило к изменению фланкирующих последовательностей. В примере преобразования scFv в IgG,  $V_L$ -элемент предшествует промотору и лидерной последовательности, и  $V_H$ -элемент следует домену  $C_{H1}$ - $C_{H2}$ - $C_{H3}$  в случае антител, получивших форму IgG, и домену  $C_{H1}$ , в случае антител, получивших форму Fab. Способ можно применять в более широком смысле, где элементы А и С (с использованием вышеописанной номенклатуре) могут представлять другие генетические элементы, например, в строительстве белков с циркулярной перестановкой, где слит оригинал N- и С- конца и неизвестный внутренний конец спроектирован.

На Фигуре 21 в качестве примера схематически проиллюстрирован процесс преобразования scFv в IgG. ДНК-вставка (а),  $V_H$ - и  $V_L$ -домены антитела, лигирована с ДНК-фрагментом (b), кодирующим константный участок легкой цепи (CL), последовательность полиаденилированная (Pa), промотор цитомегаловируса (CMV) и сигнальный пептид (SigP). ДНК-фрагмент (b) также может кодировать любой промотор, вместо промотора CMV. Также, кассета pA-CMV может заменяться на участки внутренней посадки рибосомы (IRES) [119] или типа 2A "саморасщепляющиеся" малого пептида [130, 133]. Сшивке молекул ДНК (а) и (b) для получения «миникольцевой ДНК» нерепликативным путем (с) способствует лигированию с использованием «липких концов». На Фигуре 21, сайты NcoI и NotI используются потому, что они используются при создании

библиотеки фагового дисплея McCafferty [7], однако любые подходящие сайты рестрикции могут использоваться для создания нерепликативного "мини-круга" с. После лигирования, "мини-круг" линейризованный с ферментами рестрикции NheI и XhoI, сайты распознавания которых фланкирует линкер между доменами  $V_H$  и  $V_L$ . NheI и XhoI выбраны для иллюстрации данного изобретения, так как они использованы в создании фаговой библиотеки McCafferty [7], однако могут использовать любые подходящие сайты рестрикции.

Линейризованный продукт d затем очищали и лигировали с расщеплением вектора (e). Вектор (e) включает в себя промотор CMV или pEF и сигнальную последовательность выше от сайта NheI, и кодирует антитело с константным участком тяжелой цепи ( $C_H$ ) от 1 до 3 доменов ниже сайта XhoI. Вектор будет также кодировать оригинал бактерии или репликацию и маркер устойчивости к антибиотику (не показан), чтобы иметь возможность выбрать и подвергнуть репликации полученную плазмидную ДНК в бактерии. Продуктом лигирования вставки (d) с вектором (e) будет получена плазида f, которая может использоваться для трансформации бактерии и роста с подходящим селективным маркером, позволяющим производство и очищение плазмидной ДНК стандартными способами. Очищенную плазмиду f можно вводить в клетки млекопитающих [134] для экспрессии гетерогенного антитела Ig. В качестве альтернативы ДНК, кодирующая  $C_{H1-3}$  в векторе (e), можно заменить ДНК, кодирующей один домен  $C_{H1}$  для экспрессии Fab.

В подробном описании ниже способ, используемый для иллюстрации настоящего изобретения, вставка b содержит либо CMV промотор, либо P2A пептид, который позволяет экспрессировать антитело, отделяя легкие и тяжелые цепи из одной информационной РНК (мРНК). Способ не нагляден и уточняли после нескольких экспериментальных попыток. Например, первоначально линейризацию «миникольцевой ДНК» (c) попробовали с помощью ПЦР. Однако это привело к амплификации гомодимерных побочных продуктов, что приводит к низкому выходу целевого продукта (d). В противоположность этому, прямое расщепление «миникольцевой ДНК» (c) при обеспечении достаточного количества материала (d), допускает способ к успешному выполнению. Во-вторых, в попытке предотвратить нежелательный гомодимерный продукт, вставки (a) изначально дефосфорилировали. Однако это требует тщательного контроля, чтобы предотвратить "конец" расщепления, приводящий к продукту не желаемых "липких концов" для лигирования. Оптимальный способ не включает в себя дефосфорилирование, чтобы максимально увеличить долю лигирования компетентного продукта. И, наконец, тщательный контроль соотношений, используемых в лигировании вставок ДНК (a) и (b) требовалось максимально увеличить выход «миникольцевой ДНК» (c).

### 1. Получение вставок scFv методом ПЦР

От глицерина бактериальный штамм, который несет плазмидную ДНК, кодирующая scFv, очищали в 50 мкл воды. Развести это 1 к 10. Использовать 5 мкл этого для реакции ПЦР, содержащей прямой праймер Psang10peIB (CGCTGCCAGCCGGCCATGG SEQ ID NO: 95) (2,5 мкл, 5 мкМ), обратный праймер 2097 (GATGGTGATGATGATGTGCGGATGCG SEQ ID NO: 96) (2,5 мкл, 5 мкМ), 10x KOD буфер (KOD hot start kit from Merck, 71086-4), дНТФ (5 мкл, 2 мкМ), MgSO<sub>4</sub> (2 мкл, 25 мкМ), KOD запуск из горячего старта полимеразы (2,5 единиц) в общем объеме 50 мкл. Циклическую обработку проводят при 94 °С в течение 2 мин, затем 25 циклов при 94 °С в течение 30 секунд, при 54 °С в течение 30 сек, затем 72 °С в течение 1 мин. ПЦР очистку проводили с помощью микроцентрифужной колонки (Qiagen or Fermentas) и реакции ПЦР элюировали в 90 мкл. На Фигуре 22а показано, что 1 мкл реакционной смеси для ПЦР наносили на 1 % гель агарозы TBE. Очищенную scFv ДНК (80 мкл, 8 мкг) подвергали расщеплению, путем добавления буфера 4 (New England Biolabs), БСА (0,1 мг/мл) и 40 единиц NcoI-HF и NotI-HF, в общем объеме получали 100 мкл и инкубировали в течение 2-х часов при температуре 37 °С. Вставки очищали с помощью набора Qiagen для очистки ПЦР и элюировали в 30 мкл и концентрацию ДНК измеряли с помощью измерения оптической плотности при 260 нм с помощью спектрофотометра nanodrop (Thermo).

### 2. Лигирование вставок ДНК (Фигура 21а и b)

Реакцию лигирования проводят для получения «миникольцевой ДНК» (Фигура 21с). Реакция лигирования содержит вставку b (125 нг), scFv вставку (Фигура 21) (125нг), 10x буфера для лигирования (Roche T4 DNA ligase набор, 1,5 мкл), T4 ДНК-лигазу (1 единица) в общем объеме получали 15 мкл. Инкубировали 1-2 часа при 21 °С. Воду (35 мкл) добавляли к смеси для лигирования и очищали с помощью набора Qiagen для очистки ПЦР и элюировали в 30 мкл.

### 3. Расщепление «миникольцевой ДНК» (Фигура 21с) с XhoI / NheI

Очищенную реакцию лигирования (28 мкл) подвергали расщеплению, путем добавления буфера 4 (New England Biolabs, 3,5 мкл), БСА (0,1 мг/мл) и 10 единиц NcoI-HF и NotI-HF, в общем объеме получали 35 мкл и инкубировали в течение 2-х часов при температуре 37 °С. Затем очищали путем разделения на 1 % агарозном геле TBE (Фигура 22b). В качестве альтернативы, на Фигуре 22с показана линейаризованная «миникольцевая ДНК», содержащий последовательность P2A на месте CMV промотора. ДНК полосу при 2,6 т. п.о (Фигура 22b), вырезали и очищали с использованием набора Qiagen для экстракции геля и элюировали в 30 мкл.

4. Лигирование линейаризованной «миникольцевой ДНК» d с pINT3 (XhoI/NheI вырезание) вектора и трансформации *E. coli* DH5 $\alpha$ .

Стандартное лигирование устанавливали с pINT3, вырезанием вектора (50 нг), линейаризованной «миникольцевой» d (20 нг), 10х буфером для лигирования (Roche, 1,5 мкл) и 1 единица T4 ДНК-лигазы (NEB), в конечном объеме получили 15 мкл. Инкубировали при 21°C в течение 2 часов. Трансформацию *E. coli* DH5alpha химически компетентных клеток, эффективно субклонировали (Invitrogen, номер по каталогу 18265017) в соответствии с инструкциями изготовителя. 80 мкл химически компетентных DH5a клеток добавляли в 6 мкл смеси для лигирования, помещали на лед в течение 1 часа и подвергали тепловому шоку при 42 °C в течение 1 мин, со льдом в течение 2 мин и затем переносили 14 мл в полипропиленовую пробирку, содержащую 900 мкл SOC среды и инкубировали при 37 °C в течение 1 часа и высевали на LB amp пластины.

Пример 15 Строительство больших библиотек дисплея на клетках млекопитающих с помощью встраивания, направленного нуклеазой, используя проточную электропорацию

Электропорация является эффективным способом введения ДНК, РНК и белка в клетки и системы проточной электропорацией, предусматривая эффективное введение ДНК в большом количестве клеток млекопитающих. Например, «MaxCyte STX Scalable Transfection System» (Maxcyte) позволяет электропорировать  $10^{10}$  клеток в течение 30 минут, создавая потенциал для трансфекции до  $10^{11}$  клеток в день. Клетки и ДНК смешивали и пропускали из резервуара в электропорационную камеру, электропорировали, откачивали и процесс повторяли со свежей аликвотой клеток и ДНК. Тот же метод может применяться для введения ДНК, РНК, белка или их смесей в культивируемые клетки (например, клетки HEK293 человека или Jurkat клетки) или первичные клетки, например, лимфоциты человека [135]. Проточная электропорации используется для эффективного введения ДНК, РНК и белка в большом количестве первичных и культивированных клеток.

Здесь мы приводим пример использования такой системы для введения донорной ДНК, кодирующая гены антитела, путем котрансфекции с ДНК, кодирующая пару TALE, меченых нуклеаз к локусу AAVS человека, клеток HEK293 человека и клеток Jurkat.

Распределение 2-х различных по специфичности антител определяли с помощью проточной цитометрии с использованием флуоресцентно-меченым антигеном. Создание антител, распознающих FGF-рецепторы FGFR1 или FGFR2, описаны ранее [105]. Клоны  $\alpha$ -FGFR1\_A и  $\alpha$ -FGFR2\_A (описанные здесь) клонированы в pD6, как описано в Примере 6. В дополнение к этому, популяцию scFv антител, выбранных из библиотеки фагового

дисплея «McCafferty» [7], используя один раунд фагового дисплея на  $\beta$ -галактозидазе ( $\beta$ -гал), также клонировали в этот вектор (как описано в Примере 6).

Клетки HEK293 центрифугировали и ресуспендировали в конечном объеме  $10^8$  клеток/мл в буфере от изготовителя для электропорации (Maxcyte Electroporation buffer, Thermo Fisher Scientific номер по каталогу NC0856428). Аликвоту из  $4 \times 10^7$  клеток (0,4 мл) добавляли в кювету для электропорации с 100 мкг ДНК (т.е.  $2,5 \text{ мкг}/10^6$  клеток). Количества, используемых различных компонентов, приведены ниже. Донорная ДНК, кодирующая антитела  $\alpha$ -FGFR1\_A и  $\alpha$ -FGFR2\_A, представлена в виде эквимольной смеси с общей суммой на  $10^6$  клеток показана в Таблице 6 ниже. ДНК, кодирующая AAVS-SBI TALEN (pZT-AAVS1 L1 and pZT-AAVS R1 Systems Bioscience номер по каталогу GE601A-1), используют в виде эквимольной смеси с общей суммой на  $10^6$  клеток показана в Таблице 6 ниже. В образцах без добавления TALEN, ввод ДНК привел к клеткам  $2,5 \text{ мкг}/10^6$  с использованием плазмидного контроля pcDNA3.0.

Эффективность процентного соотношения трансфекции рассчитывали путем подсчета количества колоний из бластицидина, достигнуты учитывая от ввода общего числа клеток. Граница расхождения по сравнению с отрицательным контролем (т.е., ДНК добавляли без TALEN) показана в скобках. В заключение, количество трансформированных колоний, достигаемое путем выполнения полного цикла системы Maxcyte, включая электропорацию из  $10^{10}$  клеток, вычислены в последней колонке. Это означает, что один цикл примерно за 30 минут, дающий возможность работать с многократными циклами в день. Таким образом, суточная производительность может быть в 5-10 раз выше границы. Большой масштаб ферментации и культивирования систем, такие как система Wavebag (GE Healthcare) или система Celltainer (Celltainer Biotech), можно использовать для создания клеток для трансфекции и можно использовать для выращивания библиотек, полученных в результате.

Образец	$\alpha$ -FGFR1_A/ $\alpha$ FGFR2_A нг донорной ДНК/10 <sup>6</sup> клеток	-TALEN нг ДНК/10 <sup>6</sup> клеток	% трансфекции	Количество клонов н 10 <sup>10</sup> клеток
1	580	1920	5,1 (51x)	5,1 x 10 <sup>8</sup>
1b	580	640	3,3 (33x)	3,3 x 10 <sup>8</sup>
2	580	-	0,1	0,1 x 10 <sup>8</sup>
3	194	1920	2,7 (89x)	2,7 x 10 <sup>8</sup>
4	194	-	0,03	0,03 x 10 <sup>8</sup>
5	1185	1315	5,8 (25x)	5,8 x 10 <sup>8</sup>
6	1185	-	0,23	0,23 x 10 <sup>8</sup>
7	1825	675	6,1 (11x)	6,1 x 10 <sup>8</sup>
8	1825	-	0,57	0,57 x 10 <sup>8</sup>
9	580 (FGFR1 один)	1920	5,3	5,3 x 10 <sup>8</sup>
10	580 (FGFR2 один)	1920	5,3	5,3 x 10 <sup>8</sup>

Образец	( $\beta$ -галактозидаза донорной ДНК/10 <sup>6</sup> клеток	TALEN нг ДНК/10 <sup>6</sup> клетки	% трансфекции	Количество клонов на 10 <sup>10</sup> клеток
11	580	1920	4,5	4,5 x 10 <sup>8</sup>
12	580	-	0,21	0,21 x 10 <sup>8</sup>
13	1185	1315	5,5	5,5 x 10 <sup>8</sup>
14	1185	-	0,21	0,21 x 10 <sup>8</sup>

Table 6. Электропорации НЕК293 клеток

Этот пример показывает, что существует возможность сделать очень большие библиотеки клеток со встроенными кассетами антитела. Эффективность трансфекции в диапазоне от 2,7 до 6,1 %. В случае отбранной популяция с  $\beta$ -галактозидазой (образец 13), библиотеки из  $5,5 \times 10^8$  клонов можно создать в одном сеансе потока электропорации. С более чем одного сеанса в день можно создать библиотеку из  $2-5 \times 10^9$  клонов.

После 13-ти дней отбора из бластицидина (10 мкг/мл), клетки метили фикоэритрин-мечеными Fc антителом, как описано ранее (BioLegend, номер по каталогу 409304). В популяции антитела отобранных из  $\beta$ -галактозидазы, 34-36 % клеток были

положительными для Fc экспрессии и 11-13 % были положительными для связывания с Dyelight-633-меченым антигеном при концентрации 10 нМ.

Там, где были использованы FGFR-связывающие клоны, 98-99 % клеток были положительными для Fc экспрессии. Использование смеси из  $\alpha$ -FGFR1\_A и  $\alpha$ -FGFR2\_A антител дает возможность исследовать долю клеток, содержащих множество случаев встраивания. Для отдельной клетки с правильно встроенной кассетой (например,  $\alpha$ -FGFR1\_A), есть приблизительно 50:50 вероятности того, что второе встраивание будет альтернативной специфичности (т.е.  $\alpha$ -FGFR2\_A). Если есть частые многократные встраивания, то будет высока доля двойных положительных клонов, однако, не было обнаружено высокой доли двойных позитивных клонов, иллюстрирующий точность воспроизведения встраивания, направленного нуклеазой, библиотек системы в создании одного гена антитела/на клетку. Способность поверхности, специфически связывать представленные антитела FGFR с их соответствующим антигеном, была подтверждена. Экспрессия антигена была из плазмиды pTT3DestrCD4(d3+4)-His10 [134], кодирующую эктодомен Fgfr1 мыши (ENSMUSP00000063808). Это использовали для трансфекции НЕК293 взвеси клеток и секретируемых Fgfr1-rCd4-His10, очищенными с помощью фиксированием металла аффинной хроматографии, как описано ранее [134]. Эктодомен Fgfr1 мыши амплифицировали методом ПЦР из клона IMAGE 9088089 с использованием праймеров:

2423 (TTTTTTCCATGGGCCGGCCCTCCTTCAGTTTAGTTGAG) (SEQ ID NO: 97) и

2437 (TTTTTTGCGGCCGCGGAAGCCGTGATCTCCTTCTCTCTC) (SEQ ID NO: 98), подвергали расщеплению с NcoI / NotI и клонировали в экспрессионную плазмиду pBIOCAM5 [126]. Fgfr2-Fc, экспрессировали с помощью временной трансфекцией клеток НЕК293, как описано ранее [134], и очищали с помощью аффинной хроматографии.

Трансфецированные популяции испытывали для двойного связывания с использованием обоих меченых антигенов и доля двойных позитива была низкой. В этом эксперименте оптимальный баланс размера библиотеки ( $2,7 \times 10^8$  клонов/на Maxcyte сеанс), с низким процентом двойных позитивов (3,5 %) обнаружили с использованием 197 нг донорной ДНК на  $10^6$  клеток (Фигура 23).

Эта доля двойных позитивов может представлять mis-включение второй кассеты антитела, но принимая во внимание эффективность встраивания, направленного нуклеазой, библиотеки, также возможно, что оба аллеля (далее локус AAVS в данном примере) могут быть направлены с поступающими связующими молекулами в долю клеток. Наличие двух различных генов антител в клетке само по себе не препятствует выделению связующих молекул или их кодирующих генов, но это можно обойти путем первой модификации

клетки-мишени в одном локусе для введения одного встраивания, направленного нуклеазой, например, пре-встраиваемый сайт мегануклеазы SceI, как продемонстрировано в Примере 9.

Пример 16. Выделение генов, кодирующих связывающие молекулы из отсортированной библиотеки популяций.

Отбор методом фагового дисплея проводили на  $\beta$ -галактозидазе (как описано в Примере 6), а также популяции антитела с 1-2 раундов отбора, клонировали в вектор pD6 и вводили в локус AAVS клеток HEK293, как описано в Примере 6. В-галактозидазу метили с использованием Lightning Link Dyelight-633 (Innova Bioscience номер по каталогу No 325-0010) в соответствии с инструкциями производителя. Трансфецированные клеточные популяции отбирали в течение 25 дней на бластицидине (10 мкг/мл) и метили 10 нМ Dyelight-633 меченой  $\beta$ -галактозидазой и фикоэритрин-меченым Fc антителом (Biolegend номер по каталогу 409304). Клетки инкубировали с антителами в течение 30 минут при 4 °C, дважды промывали в PBS/0,1 % БСА, ресуспендировали в PBS/0,1 % БСА и дважды положительные клетки отсортировали с использованием сортирующего потока.

Отсортированные клетки выращивали и второй раунд сортировки проводили с использованием 10 нМ антигена. Клетки выращивали и либо геномную ДНК, либо мРНК выделяли из отсортированных, отобранных популяций.

Там, где связующие молекулы, охватывающие различные цепи (например, антитела, получивший форму IgG), содержатся на той же самой геномной последовательности (например, путем введения на той же плазмиде), но транскрибированы в различные мРНК, могут быть оптимальными для восстановления отдельных генов, кодирующие мультимерную связующую молекулу, путем амплификации из геномной ДНК. В качестве альтернативы, связующие молекулы, включающие множество цепей белка, можно закодировать на той же самой мРНК посредством применения "внутренней посадки последовательности рибосом" (IRES) элементов или последовательностей, таких как вирусные Р2А или Т2А последовательности, которые способствуют поступательному расщеплению/белка [133, 136]. В этом случае и в случае связующих молекул, закодированных на одной цепи белка, также будет возможно выделить кодированные гены из мРНК.

Геномную ДНК получали с использованием «DNeasy blood and tissue kit» (Qiagen номер по каталогу 69504). мРНК получали с использованием «Isolate II RNA mini kit» (Bioline Bio-52072). Для амплификации ПЦР-реакции из геномной ДНК, начинали использование Phusion полимеразы со смесью "2xPhusion GC" в соответствии с инструкциями производителя. Праймеры, которые фланкируют сайты клонирования NcoI



и Not1, например, праймеры 2622 (GAACAGGAACACGGAAGGTC) (SEQ ID NO: 99) и 2623 (TAAAGTAGGCGGTCTTGAGACG) (SEQ ID NO: 82) использовали для амплификации кассеты антитела (при 98°C в течение 10 секунд, при 58°C в течение 20 секунд, при 72°C в течение 90 секунд, при 35 циклов).

Гены, кодирующие отобранные гены для scFv, амплифицировали из мРНК. Общую РНК выделяли из отсортированных клеток с помощью с использованием «Isolate II RNA mini kit» (Bioline Bio-52072). кДНК синтезировали из 2 мкг РНК с использованием набора Superscript II обратной транскриптазы (Life Technologies, номер по каталогу по 180064-022). Отобранные гены для scFv затем амплифицировали из кДНК с помощью ПЦР, с использованием набора KOD Hot Start ДНК-полимеразы (Merck Millipore номер по каталогу 71086-3), с использованием праймеров которые фланкируют сайты клонирования Nco1 и Not1. В этом случае, праймеры: 41679 ATGAGTTGGAGCTGTATCATCC (SEQ ID NO: 100) и 2621 GCATTCCACGGCGGCCGC (SEQ ID NO: 101) использовали для амплификации кассеты антитела (95 °C в течение 20 секунд, при 60 °C в течение 10 секунд, 70 °C в течение 15 секунд, при 25 циклов). Продукты ПЦР подвергали расщеплению с Nco1 и Not 1, до клонирования в сайт Nco/Not1 бактериальной клетки вектора pSANGIO, экспрессирующего антитело. Создание pSANG10, способы экспрессии в клетках бактерий и скрининга с помощью ИФА метода описаны в Martin et al 2006 [127].

ИФА скрининг популяции с 1 раунда отбора с помощью фагового дисплея показал, что 0/90 клонов были положительными. В отличие от этого, когда эта же популяция дополнительно представляла дисплей в клетках млекопитающих и ген scFv популяции извлекали и подвергали скринингу, 27/90 клоны (30 %) были положительными в ИФА. Это свидетельствует о том, что можно осуществлять на библиотеке дисплея на клетках млекопитающих и восстановить обогащенную популяцию связующих молекул.

Пример 15 описывает введение популяции, отобранной с помощью 1 раунда фагового дисплея на β-галактозидазе, в НЕК клетки с использованием потока электропорации. Популяция дисплея клеток млекопитающих отобрана в бластицидин, как и раньше, и подвергали потоку сортировки с использованием 10 нМ меченой β-галактозидазы. Через 9 дней роста с помощью проточной цитометрии 75 % клеток обнаружили положительными, связывающие β-галактозидазу с использованием 10 нм (β-галактозидазы). Их отсортировали и распространили дальше. Чтобы проиллюстрировать управление необходимою возможностью, мечение осуществляли с использованием концентрации антигена либо 1 нМ либо 10 нМ. 20,3 % и 55,9 % клеток соответственно отсортировали из каждой популяции. После сортировки мРНК сразу получают из отсортированной популяции без дополнительной клеточной культуры. После

клонирования, экспрессии в бактериях и ИФА скрининга (как и раньше) установили, что вероятность успеха в ИФА возрастала с увеличением необходимости на потяжении проточной сортировки. Клоны, отображающие самый высокий уровень сигнала, пришли из этой группы, а количество позитивных клонов также повышалось (Фигура 24). Это иллюстрирует способность управлять необходимым отбором в популяциях дисплея, отображая в лучшую производительность полученных антител.

Пример 17. Включение генов рецепторов Т-клеток (TCR) с библиотеки млекопитающих, полученные с помощью встраивания, направляемого нуклеазой

Способы, описанные здесь, имеют применение за пределами дисплея антител. Чтобы продемонстрировать возможность для скрининга библиотек рецепторов Т-клеток с использованием встраивания, направляемого нуклеазой, была построена векторная конструкция (pINT20), которая обеспечивает экспрессию рецепторов Т-клеток.

pINT20 (Фигура 25a) представляет собой вектор с двойным промотором для направленного изменения локуса AAVS человека. В нем содержатся левый и правый гомологичные участки, которые представлены на Фигуре 3. Левый гомологичный участок AAVS фланкируется уникальными сайтами AsiSI и NsiI, а за ними следует акцепторный сайт сплайсинга и гена устойчивости к пурамицину. Последовательность между концом левого гомологичного участка и акцепторного сайта сплайсинга, такого же, как ранее описано, а ген устойчивости к пурамицину начинается с кодона ATG в рамке считывания выше экзона ((Фигура 25b и, как также показано, для гена устойчивости к бластицидину на Фигура 3). Точное встраивание направленного нуклеазой приведет к сплайсингу в рамке гена устойчивости к пиримидину, который эндогенно соединен выше от экзона, который сам приводит в действие эндогенный AAVS промотор, допускающий отбор клонов с точным встраиванием. Ген устойчивости к пурамицину с последующим сайтом полиаденилирования SV40. Правый гомологичный участок AAVS фланкирует уникальные сайты BstZI71 и SBF1.

Конфигурация pINT20 осуществляется с использованием промотора PEF (из pSF-pEF, Oxford Genetics номер по каталогу OG43), что позволяет клонировать представляющие интерес гены в сайты NheI/KpnI. Сайту NheI предшествует лидерная последовательность, обеспечивающая секрецию, а за сайтом KpnI следует сигнал полиаденилирования бычьего гормона роста (полиА BGH), как показано выше на Фигуре 2. Ниже представлен промотор CMV (из pSF-CMV-f1-Pac1, Oxford Genetics, номер по каталогу No OG111), позволяющий клонирование с помощью NcoI и NotI (как показано на Фигуре 2) или HindIII. Сайт NcoI, предшествует лидерной последовательности, обеспечивающей секрецию, а за кассетой следует полиА-сайтом BGH.

В качестве примера дисплей и обогащение Т-клеточных рецепторов (TCR), TCR, распознающий опухолеассоциированный маркер, описан Li *et al.* (2005), а затем Zhao *et al.* (2007) [137, 138]. Этот TCR, называемый c12c2, распознает пептид SLLMWITQV (SE ID NO: 102), презентированный на HLA-A2, с аффинностью 450 мкМ. Этот пептид представляет собой остатки с 157-165 из NY-ESO-1 (NY-ESO-1 157-165). Это вариант, полученный в результате аффинного созревания родительского антитела под названием 1G4 с аффинностью 32 мкМ.

Второй TCR был использован для распознавания другого опухолеассоциированного маркера. Родительский MEL5 TCR распознает пептид MART-1 26-35 («антиген меланомы, распознаваемый Т-клетками 1»), представленный на HLA-A2 с пептидной последовательностью ELAGIGILTV (SEQ ID NO: 103). Этот TCR созрел с помощью фагового дисплея, чтобы получить клон  $\alpha 24/\beta 17$  с 0,5 нМ аффинностью, описанной в Madura *et al.* (2013 г.) [139] Структура комплекса между TCR и МНС: пептидный комплекс должен быть растворимый (pdb code 4JFH), а клон далее называют "4JFH". Тот же родительский TCR также был разработан на основе конструкции Pierce *et al.* (2014) [140], а структура комплекса растворима.

Согласно Debets и коллегами, прикрепление домена CD3  $\zeta$ , как показано, имеют тенденцию вызывать ассоциацию гетерологичного гена, даже в присутствии нативного TCR человека, [141, 142]. Используемый элемент CD3  $\zeta$ , состоит из внеклеточного домена, трансмембранного домена и полного цитоплазматического домена. К тому же, замещения константных доменов человека с помощью константных доменов мыши в гетерологичных генах, также имеет тенденцию вести их объединение по ассоциации с эндогенными константными доменами человека [143]. Наконец, применение константных доменов мыши предоставляет вариант определения гетерологичных цепей TCR на фоне TCR человека. Эти элементы включены в конструкцию экспрессионной кассеты TCR.

Были разработаны и синтезированы два синтетических гена с получением генных конструкций со следующей структурой:

V $\alpha$  TCR человека- константный домен  $\alpha$ -цепи мыши- CD3  $\zeta$  человека

V $\beta$  TCR человека - константный домен  $\beta$ - цепи мыши-CD3  $\zeta$  человека

Последовательность синтетического гена, включающей в себя цепь и  $\beta$  цепь конструкции, включающие переменные домены c12/c2, показан на Фигуре 25c и d. Эти клонировали в сайты Nhe1/Kpn1 и сайты Nco1/Hind3 из pNT20 соответственно. Конструкция, кодирующая этот TCR, называется pNT20-c12/c2. В первом случае синтетический ген разработали для включения V $\alpha$  C $\alpha$  домена (фланкированный Nhe1 и Not 1 сайтами), и домен V $\beta$  (фланкированный сайтами Nco1/Xho1, кодирующие TCR c12/c2).

Эти элементы можно затем заменить альтернативными TCR с использованием этих сайтов рестрикции.

Два дополнительных синтетических гена сделаны для кодирования доменов V $\alpha$  и V $\beta$  из 4JFH [139] (Фигура 25e и f). Конструкция, кодирующая этот TCR, называется pINT20-4JFH.

Клетки НЕК293 в количестве  $10^7$  трансфецировали с использованием 3 $\mu$ g из pINT20-c12/c2 и pINT20-4JFH в качестве донорной ДНК (300 нг донорной ДНК на  $10^6$  клеток), в соотношениях, приведенных в Таблице 7. pINT20-c12/c2 признана как TCR1 и pINT20-4JFH признана как TCR2. 5 мкг каждой из pZT -AAVS1 L1 и pZT- AAVS R1 TALEN добавили к  $10^7$  клеток (500 нг каждого на  $10^6$  клеток), за исключением образца 6, где это было заменено 10 мкг контрольной ДНК (pcDNA3.0). ДНК вводили с помощью полиэтилениминовой трансфекцией, как описано выше.

Образец	Соотношение TCR1/TCR2
1	1:100
2	100:1
3	50:50
4	100 % TCR1
5	100 % TCR2
6	50:50 (без TALE нуклеазы)

Таблица 7

Через 12 дней после отбора (0,25 мкг/мл пурамицина) клетки метили антигеном-мишенью. Комплексы пептид:МНС, распознаваемые ТКР, описанной выше, представлены в виде фикоэритрин-меченого пентамера (ProImmune). c12/c2 распознает пептид SLLMWITQV, представленный на HLA-A2, распознающий NY-ESO-1 157-165 (код продукта ProImmune 390). 4JFH (также известный как  $\alpha$ 24/ $\beta$ 17) распознает пептид ELAGIGILTV, представленный на HLA-A2, распознающий MeIanA/MART 26-35 (код продукта ProImmune 082). В каждом случае МНС:пептидный комплекс метили с фикоэритрином и использовали в соответствии с инструкциями изготовителя. На Фигуре 26 (a-d) показано, что каждый TCR является специфическим, для ожидаемого МНС: пептидный комплекс (a, d), и когда не удается связать не когнантный пептид в комплекс (Фигура 26 b, c). ДНК, кодирующий каждый TCR смешивают с избытком 100-кратным ДНК, кодирующий другие (образцы 1-2, Таблица 7). Клетки НЕК293 трансфецировали и

отобрали в пурамицин. Отбор положительных клеток к антигену проводили через 14 дней после отбора в пурамицин (Фигура 26 g, h).

В целях оценки уровня обогащения популяции клеток, полученных после сортировки на проточном цитофлуориметре, гены TCR были получены с помощью ПЦР амплификации, и относительное количество каждого вида TCR определяли после клонирования. Общую РНК выделяли из отсортированной популяции. Синтез кДНК осуществляли, как описано в Примере 16. Праймерами для амплификации альфа- и бета-цепи TCR были 1999/2782 и 41679/2789, соответственно (Таблица 8). Для ПЦР амплификации применяли KOD hot start polymerase, используя рекомендованный протокол производителя (EMD Millipore, 71086, EMD Millipore). Условиями для реакции ПЦР были 95°C (2 мин) и 25 циклов при 95°C (20 секунд), 60°C (10 секунд), 70°C (15 секунд), а затем 70°C (5 минут). Амплифицированные TCR альфа- и бета-цепи расщепляли с NheI/NotI или NcoI/XhoI и субклонировали в векторы с совместимыми сайтами (в данном случае, pBIOCAM1-Tg-N NheI/NotI или pBIOCAM2-Tg-N (NcoI/XhoI) вырезали векторы соответственно). Отдельные клоны ПЦР амплифицировали с c2c12 (TCR1), специфическим TCR альфа-праймером (2781) или через 4JFH (TCR2), специфическими альфа-праймерами TCR (4JFH-Va-F) и векторных специфических праймеров для анализа идентичности клона TCR. После сортировки образца 1 (Таблица 7), где в соотношении 1:100 TCR1/TCR2 применяли донорную плазмиду (Таблица 7) с обогащением для TCR1 специфических клонов с использованием пептида SLLMWITQV, представленных на HLA-A2 (код товара Proimmune 390), доля TCR1 клонов, как определили с помощью колонии ПЦР, увеличилась до 11/15 (73 %). Обогащения путем сортировки образца 2, где в соотношении 100:1 TCR1/TCR2 применяли донорные плазмиды (Таблица 7), с пептидом ELAGIGILTV, представленные на HLA-A2 (код товара Proimmune 082), привело к увеличению доли TCR2 клонов 4/15 (27 %), определенных с помощью колонии ПЦР.

Для демонстрации отбора библиотеки с использованием встраивания, направленного нуклеазой, мутантная библиотека на основе c12/c2 была создана путем клонирования наборов генов, кодирующих мутантные TCR альфа-цепи вместе с набором генов, кодирующих мутантные TCR бета-цепи. Такая библиотека может быть создана с использованием подходов к мутагенезу, направленного олигонуклеотидом, например, способы, на основе Kunkel мутагенеза [144]. В качестве альтернативы и в качестве примера, подход сборки ПЦР использовали с образованием мутантной TCR альфа-цепи (как фрагмент NHE1/KpnI) и мутантной TCR бета-цепи (как фрагмент KpnI/Hind 3, в том числе CMV промотор), которые клонировали в сайт NHE1/Hind 3 из pINT20. Для этого использовали праймеры, показаны в Таблице 8.

4JFH-Va-F (SEQ ID NO: 104)	ACACACGCTAGCCAGAAAGAGGTGGAACAG
1999 (SEQ ID NO: 87)	AAAAAG CAG GCTACCATGAGG G CCTG GATCTTCI I ICTCC
2770 (SEQ ID NO: 105)	CAAAGAACAGCTCGCCGGTSNNCCCGASSNNGGAGCTG G CGCAAAAGTAC
2771 (SEQ ID NO: 106)	CTCGCCCGAAGGTGGGAATGTANGWTCCSNNSNNAAGT G GGCGCACGGCGCAC
2780 (SEQ ID NO: 107)	CTG G CAG CTAG CAAG CAG GAAG
2781	TACATTCCCACCTTCGG G CGAG

(SEQ ID NO: 108)	
2782 (SEQ ID NO: 109)	TTTTTTGCGGCCGCGGACAGGTTCTG
2783 (SEQ ID NO: 110)	CGTAAGCTGGTACCTTATTATCTAGGG
2785 (SEQ ID NO: 111)	CCCTAGATAATAAGGTACCAGCTTACG
2787 (SEQ ID NO: 112)	ACCG G CGAG CTGTTCI I IG
2788 (SEQ ID NO: 113)	AGTGACAAGCI I I IATTATCTGGGTG
2789 (SEQ ID NO: 114)	CAG GTCCTCGAG CACTGTC
41679 (SEQ ID NO: 100)	ATG AGTTG GAG CTGTATC ATCC

(N= A, C, G, T. S= C OR G, W =A OR T)

Таблица 8. Использование праймеров в создании библиотеки и восстановления клона

Мутантный олигонуклеотид был разработан (праймер 2771), который рандомизированный 2 аминокислотными позициями в CDR3 из c12/c2 альфа-цепи, и обеспечивает вариант либо серина либо, треонина в другом положении (праймер 2771 также представлен в нижнем положении Фигуры 25 г). Праймер 2771 использовали в

сочетании с праймером 2780 с образованием мутантного TCR альфа набора, идущего от сайта клонирования Nhe1, включающего область CDR3 мутагенеза с инвариантной последовательностью в конце. Праймер 2781 является дополнением к инвариантному 5'-концу праймера 2771. ПЦР с праймерами 2781 и 2783 обеспечивают остаток от TCR альфа-CD3 зета-кассеты до сайта Kpn1. Сборка ПЦР из 2-х фрагментов ПЦР используется с образованием TCR альфа-CD3 зета фрагмента, который может быть клонирован в pINT20 после расщепления с Nhe1 и Kpn1.

Второй мутантный олигонуклеотид разработан (праймер 2770), который рандомизированный 2 аминокислотными позициями в CDR3 из c12/c2 бета-цепи, и обеспечивает вариант либо валина, либо лейцина в другом положении (Праймер 2770, представлен в нижнем положении Фигуры 25h). Праймер 2770 использовали в сочетании с праймером 2785 с образованием мутантного TCR бета набора от сайта клонирования KpnI, включающего область CDR3 мутагенеза с инвариантной последовательностью в конце. Праймер 2787 является дополнением к инвариантному 5'-концу праймера 2700. ПЦР с праймерами 2787 и 2788 обеспечивают остаток от TCR бета-CD3 зета-кассеты до сайта Hind 3. Сборка ПЦР из 2-х фрагментов ПЦР используется с образованием TCR бета-CD3 зета фрагмента, который может быть клонирован в pINT20. Полный набор, включающий мутации в обоих CDR 3 альфа- и бета-цепей, создан путем клонирования фрагмента Nhe1/Kpn, фрагмента Kpn1/Hind3 в Nhe1/Hind3, расщепленный pINT20. После лигирования библиотеку клонировали в электрокомпетентных клетках DH10B. Плазмидную ДНК получали и трансфецировали ДНК в клетки HEK293 вместе с векторами, кодирующими TALE нуклеазу, как описано ранее (эквиволярная смесь из pZT-AAVS1 L1 и pZT-AAVS R1 Systems Bioscience, номер по каталогу GE601A-1).

После лигирования мутантных альфа- и бета-цепей с c12/c2 мутантной библиотекой в pINT20, осуществляли электропорацию ДНК в клетки DH10B, получали плазмидную ДНК, и библиотеку котрансфецировали геном и нуклеазы TALE, направленно взаимодействующей с локусом AAVS. Трансфекцию проводили с использованием системы Maxcyte для электропорации. Выращивание и отбор осуществляли, как описано выше. Количественный анализ размера библиотеки, путем титрования трансфецированных клеток и отбора в планшетах с использованием пурамицина, показал, что размер библиотеки  $5 \times 10^5$  был создан. После отбора с использованием пурамицина в течение 11 дней, клетки метили АФЦ-меченым антителом, специфичным к  $\beta$ -цепи TCR мыши (Life Technologies Cat H57-957). На Фигуре 26 i-j показано, что 38 % клонов в популяции экспрессируют T-клеточный рецептор. Из этой TCR-положительной популяции, 13 % также связывают

пептид-1 (5 % от общей численности популяции). При таком подходе могут быть выделены клоны с улучшенной экспрессией или связывающей активностью пептид:МНС.

На Фигуре 26 видно, что каждый Т-клеточный рецептор распознает только его когнантный антиген. Кроме того, когда используется смесь двух разных специфичностей, существует возможность выделить каждый из них посредством меченого антигена. Данный подход также позволяет ТCR клоны с улучшенной аффинностью (или экспрессией), выделять посредством мутантной библиотеки ТCR путем идентификации подмножеств библиотеки, которая была мечена в большей степени, чем родительский клон.

Гены Т-клеточного рецептора также были введены в Jurkat клетки посредством электропорации. Jurkat клетки центрифугировали и ресуспендировали в конечном объеме  $10^8$  клеток/мл в буфере для электропорации изготовителя (Maxcyte Electroporation buffer, Thermo Fisher Scientific. номер по каталогу NC0856428)). Аликвоту из  $4 \times 10^7$  клеток (0,4 мл) добавляли в ОС400 кювету для электропорации с 40 мкг ДНК (т.е.  $1 \text{ мкг}/10^6$  клеток). ДНК состоит из смеси донорной плазмидной ДНК (pINT20-c12/c2 или pINT20-4JHF или pINT20-c12/c2 ТCR библиотеки, 9,2 мкг) и эквимольной смеси ДНК (30,8 мкг общего количества) ДНК, кодирующей AAVS-SBI TALEN (pZT-AAVS1 L1 и pZT-AAVS R1 Systems Bioscience Cat No GE601A-1). В образцах без добавления TALEN, введение ДНК привело к  $1 \text{ мкг}/10^6$  клеток с использованием контроля плазмиды pcDNA3.0. Альтернативный способ введения генов Т-клеточного рецептора в Jurkat клетки использовали 4D-Nucleofector (Lonza). Здесь, протокол трансфекции соблюдали по инструкции изготовителя, в соответствии с набором SE клеточной линии SE (Lonza, Cat. PBC1-02250). В кратком изложении, 2 мкг ДНК, состоящую из смеси донорной плазмидной ДНК (pINT20-c12/c2 или pINT20-4JHF или pINT20-c12/c2 библиотеки ТCR, 0,46 мкг) и эквимольной смеси ДНК (1,54 мкг общего количества) из ДНК, кодирующая AAVS-SBI TALEN (pZT-AAVS1 L1 и pZT-AAVS R1 Systems Bioscience Cat No GE601A-1) трансфецировали на  $10^6$  Jurkat клеток. Настройка кода импульса была CL120 и программирование типа клеток была специфической для Jurkat E6.1 (ATCC) клеток.

На Фигуре 26 показана экспрессия ТCR и распознавание соответствующего пептида:молекул МНС. Это зависит от применения TALE нуклеазы (сравните Фигуру 26 m и n). Передачу сигналов посредством введенной ТCR, также выполнили с использованием соответствующего пептида: молекулы МНС.

pINT20-c12/c2 трансфецированные Jurkat клетки или нетрансфецированные Jurkat клетки высевали в 96-луночные планшеты при плотности  $1 \times 10^6$ /мл, 200 мкл на лунку. Клетки стимулировали либо 2 мкг либо 6 мкг на лунку ПЭ меченым пептидом 1-МНС пентамера (ProImmune) или 2 мкг/мл антитела CD3 человека (BD Pharmingen, номер по



каталогу 555329) в присутствии и в отсутствие антитела CD28 человека (BD Pharmingen, номер по каталогу 555725) в 2 мкг/мл. После 24-часовой инкубации при 37°C и 5 % CO<sub>2</sub>, активацию Jurkat клеток определяли путем исследования экспрессии CD69. Клетки окрашивали 50 мкг PBS + 1 % БСА + 0.5 мкг антитела CD69-APC человека (Invitrogen, номер по каталогу MHCD6905) на лунку в течение 45 минут при температуре 4°C.

На Фигуре 26 (образец о и р) показана повышенная регуляция CD69 при стимуляции с CD3. На Фигуре также показан эффект при добавлении 2 мкг (q) или 6 мкг (r) пептида-1:МНС. Популяция из дважды положительных клеток, которые связывают пептид:МНС и экспрессируют CD69 очевидна. Этот пример показывает клетки, инкубированные при наличии CD28, но тот же эффект наблюдался при отсутствии CD28 (не показан).

Данный пример демонстрирует возможность встраивания, направленного нуклеазой, библиотек альтернативного типа связующей молекулы, т.е. Т-клеточных рецепторов. Мы показали, что можно экспрессировать и определять TCR экспрессию на поверхности клетки с использованием специфических антител. Мы также показали, что эти Т-клеточные рецепторы специфически распознают их соответствующие мишени. Кроме того, мы создали мутантную библиотеку, позволяющую отбор улучшенных связующих молекул. Наконец, мы показали, что возможен скрининг библиотеки на основе активации сигнальных TCR в Т-клетках. Здесь мы использовали культивированные Т-клеточные линии человека. Кроме того, можно ввести ДНК в первичную культуру Т-клеток с помощью Maxcyte электропорации. Способы для выделения и подготовки первичных Т-лимфоцитов известны специалистам в данной области техники (например, *Cribbs et al.*, 2013, *Oelke et al.* 2003 [145, 146]. Воздействие TCR трансфицированных лимфоцитов к мультимерному пептиду: МНС может затем использовать для достижения активации либо посредством воздействия мультимеров пептид:МНС [146] либо антиген- презентующих клеток, нагруженных соответствующим пептидом [146, 147] Активация может быть определена либо посредством экспрессии репортерных генов либо через увеличение регуляции эндогенных генов, такие как CD69 [104, 148].

#### Пример 18. Дисплей библиотек рекомбинантных рецепторов антигенов на клетках млекопитающих

Активация Т-клеток, как правило, происходит посредством взаимодействия Т-клеточного рецептора (TCR) со специфическим пептидом и комплексов МНС. Это в свою очередь, приводит к передаче сигнала, направленного через CD3 и другие Т-клеточные сигнальные молекулы. В качестве альтернативы для распознавания мишеней, направленного посредством TCR, было показано, что альтернативные связывающие молекулы, такие как одноцепочечные Fv, могут быть представлены на Т-клетках в виде

слитых внизу сигнальных молекул таким образом, что перенаправляют активацию Т-клеток в молекулы, распознаваемой scFv (или альтернативной связующей молекулы). Таким образом, активация Т-клеток больше не ограничивается распознаванием молекул, направленных на комплексы пептид:МНС посредством TCR, но может быть направлена на другие молекулы клеточной поверхности. Этот альтернативный формат, в котором молекула, не связывающаяся с TCR, слита с компонентом, участвующим в передаче сигнала, называют «химерным антигенным рецептором» (CAR). В случае Т-клеток, это было показано, что является важным и ценным средством повторного направления для Т-клеточной активации.

Для любой данной мишени до сих пор не ясно, какой оптимальный эпитоп или аффинные признаки должны быть для включения в CAR [103]. Признаки CAR конструкции, такие как длина линкера или выбор трансмембранного домена, могут в свою очередь, влиять на то, что представляет собой оптимальный эпитоп. Комбинация плотности антигена на клетках-мишени и не мишени, вместе с выбором сигнального домена может повлиять на оптимальные требования к аффинности. Способность представлять библиотеки из рецепторов рекомбинантного антигена на Т-клетках дает возможность идентифицировать оптимальную специфичность связывания, формат связующей молекулы, линкер длина/последовательность, варианты слитого сигнального модуля и т.д., либо отдельно, либо в комбинации. В данном документе мы показали полезность встраивания, направляемого нуклеазой, для создания библиотек из рецепторов рекомбинантного антигена в клетках млекопитающих. Вектор pINT21 (Фигура 27а) представляет собой единый CMV промотор вектора для удобной экспрессии и секреции связующих молекул, такие как scFv, фланкированного сайтами рестрикции NcoI/notI, чтобы допустить в рамке экспрессию с вышерасположенной лидерной последовательностью и ниже слияния партнера (как показано ранее на Фигуре 8). Экспрессия CAR кассеты в pINT21 представляет собой фланкированные AAVS гомологичные участки, как было описано ранее на Фигуре 3

Вектор pINT21-CAR1 (Фигура 27 а, с) слит с связующими молекулами, такие как одноцепочечный Fv, в трансмембранный домена и внутриклеточного домена CD3 $\zeta$  (Фигура 27с и, как это описано, для экспрессии TCR на Фигуре 25). Этот формат часто называют «первым покоением» рецептором рекомбинантного антигена. Сигнальные домены от других костимуляторных молекул также использовали для получения дополнительных сигналов и это, как было показано, дает улучшенную передачу сигналов. Они были переданы в качестве второго и третьего поколения рецептора рекомбинантного антигена. Например, pINT21-CAR2 (Фигура 27 b, d) слияние связующего вещества (обычно

клонированный в этом случае в качестве Nco1/Not 1 фрагмент, ранее описанный домен второго поколения (WO 2012/079000 A1), состоящий из:

Петля и трансмембранный домен из CD8

4-1BB сигнальным доменом

CD3  $\zeta$  сигнальным доменом

В качестве примера количество различных групп связывающей молекулы были клонированы в сайты Nco/Not1 из pINT20-CAR1 и pINT20-CAR2. CD19, ранее был использован в ряде различных исследований мишени В-клеток злокачественных образований, литературные источники в Sadelain *et al.* (2013) [103]. Ранее описанное антитело к CD19 (WO 2012/079000 A1) (называемый FMC63) получали в качестве синтетического гена scFv, как в конфигурации VH-линкер-VL или конфигурации VL-линкер-VH (FMC63 H-L или FMC L-H соответственно, на Фигуре 27E показана последовательность FMC63 H-L. FMC63 L-H настроен с переменными доменами в конфигурации VL-линкер-VH фланкированных с помощью Nco1 и Not1 на 5' и 3' концах соответственно.

В качестве контроля, scFv с альтернативным специфичным связыванием, также клонировали в pINT20. Он включает антитела FGFR1\_A [105] и антитела Desmin контрольного антитела [7]. Кроме того, Adhirons [152], распознающий Iox1 (Фигура 29 a, b), введенная в качестве примера альтернативного формата связывающей молекулы сконфигурированной в виде слияния CAR (см. Пример 19 для описания).

Чтобы продемонстрировать создание библиотек связывающих молекул, представленных в формате CAR, клонировали популяции антител, получивших форму scFv, отобранные на мезотелине и CD229. Мезотелин представляет собой клеточный поверхностный гликопротеин, который экспрессируется на высоком уровне в целом ряде злокачественных образований, включая мезотелиому. Целый ряд форматов на основе антител находятся в стадии разработки и в клинических испытаниях, включая CAR, направленный на мезотелин [149]. CD229 представляет собой еще один потенциальный ассоциированный с опухолью антиген, который может быть мишенью иммунной терапии злокачественных образований, таких как хронический лимфоцитарный лейкоз и множественной миеломой [150, 151].

Популяцию антител, распознающих либо мезотелин, либо CD229 создали путем отбора с использованием библиотеки фагового дисплея «McCafferty» (как описано в Примере 6, и ссылка 7). Осуществляли два раунда отбора и гены, кодирующие scFv, выделяли с использованием праймеров M13leadseq и Notmucseq (Пример 6). Продукты вырезали с Nco1/Not 1, очистили в геле и клонировали в pINT21-CAR2. Их направили в

локус AAVS клеток HEK293 с помощью расщепления TALE нуклеазой, для создания библиотеки  $4,8 \times 10^5$  для CD229 и  $6,4 \times 10^5$  из мезотелина (представляющего 30х и 53х увеличение в библиотеке, по сравнению с образцами, трансфицированными при отсутствии TALE нуклеазы).

pINT20-CAR1 и pINT20-CAR2 были введены в клетки HEK293 с помощью PEI трансфекции. В данном документе донорную плазмиду ДНК (pINT20-CAR1 и pINT20-CAR2), 6 мкг смешивали с эквимолярной смесью ДНК (20 мкг общего количества), ДНК, кодирующую AAVS-SBI TALEN (pZT-AAVS1 L1 и pZT-AAVS R1 Systems Bioscience Cat No GE601A-1) в среде Freestyle 293 (Lifetech, Cat. 12338-026), линейный PEI (52 мкл, 1 мг/мл, Polysciences Inc.) добавляли и инкубировали при комнатной температуре в течение 10 минут. Затем смесь добавляли к 20 мл суспензии клеток HEK293 ( $1 \times 10^6$  клеток/мл) в 125 мл оборудованную колбу Эрленмейера. pINT20-CAR1 и pINT20-CAR2 были также введены в клетки Jurkat посредством электропорации. Клетки Jurkat центрифугировали и ресуспендировали в конечном объеме  $10^8$  клеток/мл в буфере для электропорации изготовителя (Maxcyte Electroporation buffer, Thermo Fisher Scientific Cat. no NC0856428). Аликвоту из  $4 \times 10^7$  клеток (0,4 мл) добавляли в кювету OC400 для электропорации с 40 мкг ДНК (т.е. 1 мкг/ $10^6$  клеток). ДНК состоит из смеси донорной плазмидной ДНК (pINT20-CAR1 и pINT20-CAR2, 9,2 мкг) и эквимолярной смеси ДНК (30,8 мкг от общего количества), ДНК, кодирующей AAVS-SBI TALEN (pZT-AAVS1 L1 и pZT-AAVS R1 Systems Bioscience Cat No GE601A-1). В образцах без добавления TALEN, введение ДНК привел к 1 мкг/ $10^6$  клеткам с использованием контроля плазмидной pcDNA3.0.

Флуоресцентное мечение различных антигенов проводили с использованием конъюгационного набора Lightning-Link Rapid Dye- Light 633 (Innova Biosciences, номер по каталогу 325-0010). Получение FGFR1 и FGFR2 описано в Примере 15. Lox1 и CD229 были из R и D Систем (номера по каталогу 1798-LX-050 и 898-CD050, соответственно), CD19-Fc и мезотелин были из (AcroBiosystems, номер по каталогу CD9-H5259 и MSN-H526х соответственно).

На Фигуре 28 b проиллюстрирован дисплей антитела FGFR1, направленного нуклеазой, [105] в пределах конструкции CAR второго поколения (pINT21-CAR2-FGFR1\_A). На Фигуре 28 d также проиллюстрирован дисплей альтернативной каркасной молекулы (ссылка на Adhiron [152]) в виде слияния с рецептором рекомбинантного антигена второго поколения (pINT21-CAR2-lox1). На Фигуре 28 f и g также проиллюстрированы положительные клоны в составе библиотеке scFv, отобранные на мезотелине или CD229.

В этом примере были введены CAR в клетки HEK и клетки Jurkat, но это можно также осуществить путем введения конструкции в клетки первичной культуры, такие как Т-лимфоциты человека (например, как описано Sadelain *et al.* (2013) [103]), например, с помощью электропорации [135]. Экспрессионные конструкции для экспрессии CAR в лимфоциты, могут быть дополнительно оптимизированы, например, путем оптимизации стабильности мРНК и трансляции через вариации в 5' и 3' не транскрибируемые области, поли-А длину и т.д., как это было описано ранее [135]. Передача сигнала из CAR конструкции, введенная в Т-лимфоциты первичной культуры или клеточные линии Т-лимфоцитов, может быть вызвана воздействием клеток, экспрессирующих антигеном-мишенью или с использованием мультимерного антигена, например, антигена, иммобилизованного на поверхности или представленного на гранулах [104, 148].

Пример 19. Дисплей библиотек из альтернативных каркасных белков создан в клетках млекопитающих путём встраивания, направляемого нуклеазой

Способ, описанный для создания библиотек связывающих молекул, может быть использован за пределами дисплея антител и Т-клеточных рецепторов. Целые ряды альтернативных каркасных белков описаны, что позволяет изготавливать варианты библиотек, из которых новые по специфичности связывания были выделены, например, Tiede *et al.* (2014) [152] и литературные источники в данном документе. В примере, описанном согласно Tiede *et al.* (2014), был использован стабильный гибкий каркасный белок на основе консенсусной последовательности из растительного происхождения фитоцистатинов. Это каркасный белок, ссылаясь на Adhiron, и на Фигуре 29а показан синтетический ген, кодирующий Adhiron, который был отобран для связывания с Iох1 (WO 2014125290 A1). На Фигуре 29 В показан альтернативный Iох1 связующей молекулы (Iох1В). Оба были синтезированы и клонированы в сайт NcoI/notI из pINT20\_CAR2 с образованием слияния с партнером ниже.

Библиотеку можно создать путем рандомизации остатков петли (например, путем Kunkel мутагенеза или сборки ПЦР, как описано выше в Примере 17). В качестве примера, на Фигуре 29с показана конструкция мутантных олигонуклеотидов, используемые для создания библиотеки, после такого подхода, как описано в Примере 17. В данном случае рандомизация достигается путем введения переменного количества NNS(HMB?) остатков, хотя могут быть использованы альтернативные стратегии, известны специалистам в данной области техники.

В качестве другого примера, на Фигуре 29 d и e показано средство для создания библиотеки связующих молекул, с помощью встраивания, направляемого нуклеазой на основе каркасного белка ноттина [156]. Ноттины представляют собой пептиды

приблизительно из 30 аминокислот, которые стабилизированы тремя дисульфидными связями, связывающие одну через две других, с образованием «завязанной» структуры. На Фигуре 29 d показан трипсин, связывающий ноттин MCoTI-II с сайтом NcoI на 5'-конце и сайтом NotI на 3'-конце, позволяя экспрессию в рамке считывания с векторами, описанными в данном документе. В качестве примера, для создания библиотеки из 6 аминокислот первой петли (подчеркнуты на Фигуре 29d), может быть мутирована с переменным количеством аминокислот. На Фигуре 29e проиллюстрирована мутагенная стратегия замещения 6 аминокислот 1 петли с 10 рандомизированными аминокислотами, с использованием кодонов VNS (где V = A, C или G и S = C или G).

VNS кодон включает в себя 24-кодона, кодирующих 17 аминокислот, которые исключают цистеины. Эта стратегия для иллюстративных целей и альтернативные мутагенные стратегии известны специалистам в данной области техники.

Пример 20. Встраивание, направляемое нуклеазой, библиотек антител с использованием CRISPR/Cas9

Встраивание, направляемое нуклеазой, с помощью CRISPR/Cas9, продемонстрировано с использованием набора «Geneart CRISPR nuclease vector kit» (Lifetech A21175). В этой системе промотор U6 РНК полимеразы III контролирует экспрессию целевой комплементарной РНК CRISPR (crRNA), которая связана с трансактивирующей crRNA (tracrRNA). CrRNA и tracrRNA вместе составляют направляющую РНК, которая направляет специфичное расщепление белка Cas9, кодируемый на том же наборе «GeneArt CRISPR nuclease vector» (см. инструкции производителя). Вектор предоставлен в виде линейаризованной плазмиды, в которой клонирован короткий двухцепочечный олигонуклеотид с соответствующими 3' концами. Специфичное расщепление, затем определяется последовательность клонированного сегмента. Две различные целевые последовательности предназначены для прямого расщепления локуса AAVS человека, как описано выше.

В качестве последовательностей использовали:

Вставка с помощью CRISPR1 двухцепочечной ДНК:

5' GGGGCCACTAGGGACAGGATGTTTT (SEQ ID NO: 115) 3'  
GTGGCCCCCGGTGATCCCTGTCCTAC (SEQ ID NO: 116)

Вставка с помощью CRISPR2 двухцепочечной ДНК:

5' GTCACCAATCCTGTCCCTAGGTTTT (SEQ ID NO: 117) 3'  
GTGGCCAGTGGTTAGGACAGGGATC (SEQ ID NO: 118)

Полученные в результате направляющие РНК обеспечивали нацеленное расщепление в пределах той же области локуса AAVS, что нуклеазы TALE, описанные

выше (причем CRISPR2 находится в обратной ориентации по отношению к CRISPR1). Таким образом, AAVS гомологичные участки, используемые ранее, чтобы направить встраивание экспрессионной кассеты, могут использоваться для встраивания, направляемого этой CRISPR геновой РНК. Линеаризованный вектор и двухцепочечные олигонуклеотиды лигировали и трансформировали в электрокомпетентные клетки DH10B. Клонирование корректной вставки подтверждали секвенированием и получали плазмидную ДНК. Cas9/CRISPR2 конструкцию (охватывающая олигонуклеотид CRISPR2), трансфецировали вместе с донорной ДНК, кодирующая библиотеку  $\beta$ -галактозидазы, отобранных с помощью 1 раунда отбора фагового дисплея (Пример 15). Они были трансфецированы в клетки HEK293 с использованием системы электропорации Maxcyte с ОС-400 сборки. Клетки в количестве  $4 \times 10^7$  трансфецировали с 23,2 мкг донорной ДНК, представляющая популяцию генов scFv, отобранных с помощью первого раунда отбора фагового дисплея на  $\beta$ -галактозидазе, клонированного в pD6. Клетки котрансфецировали с 77 мкг Cas9-CRISPR2 плазмиды, или 77 мкг TALEN плазмиды (38,5 мкг каждого из pZT-AAVS1 L1 и pZT -AAVS), или 77 мкг контрольной плазмиды

Номер образца	Использование нуклеазы	Эффективность трансфекции %	Количество клонов на $10^{10}$ клеток
1	Cas9/CRISPR 2	10.53 (29x)	$10 \times 10^8$
2	AAVS TALEN	5.1 (14 x)	$5.1 \times 10^8$
3	None (pCDNA3.0)	0.36	$0.36 \times 10^8$

Таблица 9

Титрование количества трансформированных клеток, образованных с помощью cas9/CRISPR2 трансфекции (путем измерения колоний, устойчивых к бластицидину) показали, что 1053 колоний, устойчивых к бластицидину, были созданы из посева 10 000 клеток, приравнивая к эффективности трансфекции 10,5 % (Таблица 9). В случае TALE, направляемого нуклеазой, была достигнута эффективность трансфекции 5,1 %. В противоположность этому, в случае отсутствия конструкции cas9/CRISPR2, эффективность трансфекции была достигнута только 0,36 %

В качестве альтернативы трансфекции с использованием плазмидной ДНК, для введения белка Cas9 и руководящей РНК в клетки, также возможно прямое введение нуклеопротеидного комплекса, состоящего из Cas9 белка (Toolgen, Inc.) и руководящей РНК. Руководящую РНК получали с помощью Toolgen, Inc., используя *in vitro* транскрипцию из промотора T7, в качестве одного транскрипта, который включал последовательность TRACR (подчеркнута) предшествующая последовательности, комплементарной ДНК-мишени (выделено жирным шрифтом), как показано ниже.

CRISPR 1 PНК (SEQ ID NO: 119):

5'

GGGGGGCCACUAGGGACAGGAUGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA  
GGCU AGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU

CRISPR 2 PНК (SEQ ID NO: 120):

5'

GGGUCACCAAUCCUGUCCCUAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAA  
GGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCUUUU

6,6 мкг Cs9 белка, 4.6 мкг PНК и 10 мкг донорной ДНК (кодирующего антитело FGFR1 \_A в PD6) были введены в клетки  $10^7$  HEK293 с помощью электропорации Maxcyte, как описано выше. Эффективность трансфекции 2,2 % и 2,9 % достигнута для CRISPR 1 и CRISPR2 PНК, соответственно, с 0,7 % и 0,8 % в отсутствие добавленного белкового комплекса cas9:PНК.

Эти руководящие PНК нацелены на одни и те же последовательности, кодируемые CRISPR1 и CRISPR2. В качестве альтернативы, crRNA и tracrRNA можно получить путем химического синтеза (например, GE Dharmacon).

Пример 21: Включение гена антитела опосредованного нуклеазой путем лигирования или микрогомологичного опосредованного конечного присоединения (ММЕJ)

Несмотря на то, что гомологичная рекомбинация (ГР) используется для точной вставки больших фрагментов ДНК, это требует создания больших нацеленных векторов, включающих длинные гомологичные участки. Это может сделать создание крупных библиотек более сложным из-за пониженной эффективности трансформации больших конструкций ДНК. В качестве альтернативы, простые реакции лигирования можно осуществить между хромосомной ДНК и нацеленным вектором, если нуклеаза, распознающая последовательность, включена в нацеленный вектор. Реакции лигирования могут быть либо с использованием "липкого конца", например, нуклеаз с доменом "цинковые пальцы" (Orlando *et al.*, 2010) [45] или TALEN (Cristea *et al.*, 2013) [22], который может сделать двунитевые разрывы (DSBs), которые удаляют 5' выступы или "притупленный конец", применяемый CRISPR/Cas9 рибонуклеопротеином. Примером встраивания гена нуклеазой, путем лигирования с использованием I-SceI мегануклеазы, было показано создание вектора pD7-SceI. pD7, полученная из pD6 (Фигура 8), но левые и правые AAVS гомологичные участки заменены короткими двухцепочечными олигонуклеотидами. Левый гомологичный участок AAVS, из серии pD вектора, фланкирует с помощью ферментов рестрикции EcoR1 и NsiI (см. на Фигуре 3). Для преобразования pD6



в pD7-SceI, левый гомологичный участок заменили вставкой двухцепочечного олигонуклеотида, образованной праймерами 2778 и 2779, кодирующие последовательность распознавания мегануклеазы I-SceI с "липкими концами", совместимые с "липкими концами", образованными с помощью EcoRI/NsiI. Правый гомологичный участок AAVS фланкирует с помощью ASCI и Mlu I сайтов рестрикции (Фигура 3). Правый гомологичный участок заменили вставкой двухцепочечного олигонуклеотида с "липкими концами", совместимые с "липкими концами", образованными с помощью расщепления ASCL/MluI и сформировали с помощью праймеров 2723 и 2724.

2723 (SEQ ID NO: 121)	CGCGCCAGAAGTCTCACCAAGCCCA
2724 (SEQ ID NO: 122)	CGCGTGGGCTTGGTGAGACTTCTGG
2768 (SEQ ID NO: 123)	AATTCTCCCCTCCACCCACAGTAGGGACAGTGGGGCCA <b>GGATTGGTGACAGAAAATGCA</b>

2769 (SEQ ID NO: 124)	TTTTCTGTCACCAATCCTGGCCCCACTGTCCCTACTGTGG GGTGGAGGGGAG
2778 (SEQ ID NO: 125)	<b>AATTCTAG G G AT AACAG G GTAAT ATGC A</b>
2779 (SEQ ID NO: 126)	TATTACCCTGTTATCCCTAG
2808 (SEQ ID NO: 127)	<b>AATTCTTTTCTGTCACCAATCCTGGGGCCACTAGGGA CAC TGTGGGGTGGAGGGGATGCA</b>
2809 (SEQ ID NO: 128)	TCCCCTCCACCCACAGTGTCCCTAGTGGCCCCAGGATT G GTGACAGAAAAGAATTG

Таблица 10. Праймеры для конструкции pD7 и pINT19

Антитела, распознающие Fgfr1 и Fgfr2 (Пример 15), клонировали в pD7 с образованием pD7-SceI антитела Fgfr1 и pD7-SceI антитела Fgfr2 соответственно. Они были котрансфицированы с I-SceI, экспрессирующую плазмиду, (Пример 11, Фигура 16) в клон HEK293 клеточной линии 6F (см. Пример 11), которая содержит встроенный I-SceI, распознающий сайт.

Лигирование DSB в хромосоме и нацеливание вектора также можно получить с помощью нуклеаз с доменом цинковые пальцы или TALE нуклеаз. При инвертировании нуклеаз с доменом цинковые пальцы или сайтов распознавания TALEN на нацеленном

векторе, это может обеспечивать, что продукт из вставки больше не является мишенью для расщепления в способе, называемом «Obligate ligation-gated recombination» или ObLiGaRe (Maresca *et al.*, 2013) [153]. pD7-ObLiGaRe векторы можно получить таким же образом, как описано выше, для создания pD7-Sce1. В данном случае, левый гомологичный участок заменяется олигонуклеотидом, состоящим из праймеров 2808 и 2809, кодирующий инвертированный сайт распознавания TALEN (выделен жирным шрифтом) и спейсерную область. Правый гомологичный участок заменяется праймерами 2723 и 2809, как описано выше.

Альтернативой простым реакциям лигирования между DSB в хромосоме и нацеливания вектора, опосредованного негомологичным концом соединения (NHEJ) является опосредованный микрогомологичный конец соединения (MMEJ). MMEJ представляет собой запасной механизм DSB, который использует микрогомологичные последовательности от 5 до 25 п.о. для подтверждения ошибки на конце соединения (McVey and Lee, 2008) [154]. Стратегия для точного встраивания генов была разработана, где геномная последовательность и нацеленный вектор содержат ту же TALE нуклеазную пару, распознающую последовательность, но различный спейсерный вектор последовательности в которой предшествующая и последующая часть переключены. Геномная последовательность и вектор может быть вырезан той же парой TALEN и MMEJ происходит через микрогомологичные концы ДНК. Результат встраивания нацеленного вектора больше не является мишенью для TALE нуклеазы из-за укорочения спейсерной области, которая не является оптимальной для TALE нуклеазного расщепления (Nakade *et al.*, 2014) [155].

MMEJ AAVS направленное встраивание вектора pD7-MMEJ, можно создать таким же образом, как описано выше, для создания pD7-Sce1. В этом случае левая сторона гомологичного участка замещается олигонуклеотидом, состоящим из праймеров 2768 и 2769, кодирующих сайт распознавания TALEN (выделен жирным шрифтом) и переключающий спейсерную область (подчеркнут). Правая сторона гомологичного участка замещается праймерами 2723 и 2809, как описано выше.

Пример 22: Дизайн праймеров для создания единого (CMV) промотора кассеты фланкированный с помощью RQSA26 участков (p1NT19-ROSA)

Этот пример предназначен для того, чтобы показать, что антитело или альтернативные гены связывающей молекулы могут встраиваться в геном клеток млекопитающих, с помощью способов направляемых нуклеазой и полученные в результате скрининга клонов, для требуемой функции, с помощью либо репортерного, либо фенотипического скрининга. Этот пример показывали раньше, где гены антител встраивали

в хромосому эмбриональных стволовых клеток мыши (ES) и проводили отдельный скрининг ES колоний на их способность поддерживать плюрипотентность при воздействии условий дифференцировки [105]. Гены антител, выделенные из ES колоний, которые содержат плюрипотентный фенотип, показаны при блокировании FGFR1/FGF4 сигнального пути. Проблема, связанная с этим, ранее описанным способом, в том, что гомологичная рекомбинация может привести к небольшим размерам библиотеки, тем самым ограничивая ее способность проводить непосредственно скрининг для редких клонов, присутствующих в больших библиотеках связывающей молекулы. Способы встраивания гена, опосредованного нуклеазой, для антитела и встраивания гена связывающей молекулы являются более эффективным результатом, что приводит к большему созданию размера библиотеки и, таким образом, более вероятно, создание библиотек клеток млекопитающих, способных идентифицировать функциональные антитела путем фенотипического или репортерного скрининга клеток.

Донорный нацелеваемый вектор pINT19 предназначен для встраивания генов антител в мышинный локус ROSA26 с помощью способов, направляемых нуклеазой, для прямого функционального скрининга. pINT19 является единственным вектором с одним промотором CMV для экспрессии слитого белка scFv-Fc. Экспрессионную кассету фланкировали с помощью гомологичных участков локуса ROSA26 мыши. Поскольку вышестоящий экзон является неперевернутым, ген устойчивости к пурамицину предшествует акцепторный сайт сплайсинга и, дополнительно низуостоящий, имеет последовательность КОЗАК, ведущую в ген устойчивости к пурамицину.

Левый гомологичный участок AAVS и ген устойчивости к пурамицину из pINT18 заменили на кассету, кодирующую левый гомологичный участок ROSA26, акцепторный сайт сплайсинга, оптимизированную консенсусную последовательность Козак и ген устойчивости к пурамицину. Левый гомологичный участок ROSA26 первоначально амплифицированная в составе pGATOR (Melidoni et al., 2013 (105)) в виде двух фрагментов, что приводило к нокауту внутреннего сайта NotI. Два фрагмента, полученных путем праймеров J60/2716 и 2715/2706, объединили в сборку ПЦР с праймерами J60 и 2706, а расщепляли AsiSI и NsiI. Акцепторный сайт сплайсинга амплифицировали из pGATOR, с использованием праймеров 2709 и 2710, а кассету, устойчивую к пурамицину, амплифицировали с применением праймеров 2745 (которые включают гомологичный участок к акцепторному сайту сплайсинга и оптимизированную консенсусную последовательность Козак) и J59. Участок акцепторного сайта сплайсинга и кассета, устойчивая к пурамицину, объединяли в сборку ПЦР с использованием праймеров 2709 и J59, а расщепляли NsiI и BglII. Левый гомологичный участок ROSA26 и акцепторный сайт

сплайсинга кассеты, устойчивой к пурамицину, лигировали с вектором pINT18 (AsiS1/Bgl2).

Для завершения конструирования нацеливающего вектора ROSA правый гомологичный участок ROSA26, расположенный ниже кассеты CMV-scFv-Fc, был введен для замены правого гомологичного участка pINT18 AAVS. Это осуществляли с помощью ПЦР правого гомологичного участка ROSA, содержащегося в pGATOR (Melidoni et al., 2013), с использованием праймеров J61 и J62 для амплификации фрагмента с BstZI 71 на одном конце и SBF1 на другом. Праймер 61 был установлен, чтобы исключить эндогенный SBF1 сайт 65 п.о., вышестоящий от сайта расщепления ROSA ZFN. На Фигуре 31 показана последовательность ROSA26 левого и правого гомологичных участков.

J60 (SEQ ID NO: 129)	ACACACGGTACCGCGATCGCGCTGATTG G CTTCI I I ICCTC	AsiSI-rosa26-L-F
2706 (SEQ ID NO: 130)	TTTTTT ATG CATTCTAG AAAG ACTG GAGTT GCAGA	Nsil-rosa26-L-R
2715 (SEQ ID NO: 131)	GAGCGTCCGCCACCCTC	ROSA-Left- NotI_knockout_F
2716 (SEQ ID NO: 132)	GAGGGTGGGCGGACGCTC	ROSA-Left- NotI_knockout_R
2709 (SEQ ID NO: 133)	TTTTTT ATG C ATTAAG G G ATCTGTAG GGCG CAG	Splice_acceptor- F-Nsil
2710 (SEQ ID NO: 134)	GTGAATTCCTAGAGCGGCCTC	Splice_acceptor-R
2745	GAGGCCGCTCTAGGAATTCACGCCGCCA C	Overlap-Puro-
(SEQ ID NO: 135)	CATGACCGAGTACAAGCCCAC	F+kozak
J 59 (SEQ ID NO: 136)	AAAAAAAGATCTGTGTGTTTCGAATCAG GC ACCGGGCTTGCGGGTCAT	Bgl2-Puro-R
J61 (SEQ ID NO: 137)	tttttGTATACG G G AATTG AACAG GTGTA AAA TTG	ROSA-Right_F- BstZI 71
J62 (SEQ ID NO: 138)	TTTTTTCCTGCAGGAGGTTGGATTCTCAA T ACATCTATTGTTG	ROSA-Right_R- SbfI

2701 (SEQ ID NO: 139)	GCCGACGTCTCGTCGCTGATGTTTT	
2702 (SEQ ID NO: 140)	ATCAGCGACGAGACGTCCGCCGGTG	
2703 (SEQ ID NO: 141)	CG CCCATCTTCTAG AAAG ACGTTTT	
2704 (SEQ ID NO: 142)	GTCTTTCTAGAAGATGGGCG CGGTG	

Таблица 11. Праймеры для направляемого нуклеазой направленного встраивания в локус ROSA 26 мыши

Встраивание p1NT19, кодирующее антитело или альтернативные связывающие молекулы в мышинном локусе ROSA26, можно достигнуть с помощью встраивания, направляемого нуклеазой, библиотек антитела с использованием CRISPR/Cas9, как описано в Примере 20. Здесь, встраивание, направляемое нуклеазой, с помощью CRISPR/Cas9 можно продемонстрировать, используя набор «Geneart CRISPR nuclease vector kit» (Lifetech A21175). В этой системе промотор U6 РНК-полимераза III приводит в действие экспрессию мишени комплементарной РНК CRISPR (crRNA), которая связана с транс-активирующей crRNA (tracrRNA). CrRNA и tracrRNA вместе составляют гидовую РНК, которая направляет расщепление специфичного белка, кодируемый белком Cas9, на том же наборе «GeneArt CRISPR nuclease vector» (см. инструкции производителя). Вектор, представленный в виде линейизованной плазмиды, в который клонировали короткий двухцепочечный олигонуклеотид с соответствующими 3'-выступами. Специфичность расщепления затем определяет последовательность клонированного сегмента. 2 различные нацеливаемые последовательности предназначены для прямого расщепления ROSA26 мыши, кодируемого праймерами 2701/2702 и 2703/2704 (Таблица 10).

В качестве альтернативы были описаны нуклеазы с доменом «цинковые пальцы», которые расщепляют ДНК в пределах локуса ROSA26, [34]. Они могут быть встроены в соответствующий вектор экспрессии, который описан для мегануклеазы SCE-I (Фигура 16, Пример 11).

Опосредованное нуклеазой встраивание донорной плазмиды p1NT19 обеспечивает получение клонов, экспрессирующих секреторируемое антитело, которое может связываться с эндогенным рецептором или лигандом, что приводит либо к антагонизму [105,107] либо к агонизму [47,106,108] путей передачи сигнала через рецепторы. Чтобы включить связь

между клеточным фенотипом и функциональной активностью секретируемых антител, клетки можно высевать при низкой плотности в полужидкой среде так, чтобы распространить отдельные колонии и экспрессия антитела может быть инициирована с помощью индуцируемого промотора [105]. В качестве альтернативы, конститутивный промотор может быть использован для экспрессии гена антитела. Полужидкая среда будет поддерживать повышенную локальную концентрацию эндогенно экспрессируемого антитела так, что любое фенотипическое изменение специфическое для колонии, возникающей из отдельной клетки, будет вызвано уникальным антителом, экспрессированным из этого определенного клона. При использовании высокочувствительного репортера, как, например, промотора *Rex* или *Nanog*, слитого с репортерным геном, был бы возможен посев клеток при низкой плотности в полужидких средах, сбор и скрининг с помощью проточной цитометрии. В качестве альтернативы, стоп-кодон, расположенный ниже от гена антитела в pINT19, можно заменить трансмембранным доменом, чтобы позволить прикрепление антитела к поверхности клетки. Стоп-кодон, расположенный ниже от гена антитела в pINT19, также можно заменить на сигнальную последовательность для удержания в эндоплазматическом ретикулуме (ЕПР), для обеспечения удержания антител в ЕПР и потенциального снижения экспрессии эндогенно экспрессируемого рецептора-мишени или любого секретируемого белка, или пептида. pINT19 специфически разработан для мишени мышинного локуса ROSA26 и может использоваться для фенотипического скрининга антител или альтернативных связывающих молекул в ЕПР клетках мыши. Однако, направляемое нуклеазой антитело или способы встраивания генов связывающих молекул могут также применяться к другим функциональным скринингам, таких как те, которые описаны с использованием лентивирусного подхода [47,106,107,108].

**Литературные источники**

Все литературные источники, перечисленные ниже, и другие источники, процитированные в других разделах настоящего описания, в полном объеме включены в данный документ посредством ссылки.

- 1 Russell, S. J., Hawkins, R. E., & Winter, G. (1993). Retroviral vectors displaying functional antibody fragments. *Nucleic Acids Res*, 21(5), 1081-1085.
- 2 Boublik, Y., Di Bonito, P., & Jones, I. M. (1995). Eukaryotic virus display: Engineering the major surface glycoprotein of the *Autographa californica* nuclear polyhedrosis virus (AcNPV) for the presentation of foreign proteins on the virus surface. *Nature Biotechnology*, 73(10), 1079-1084.
- 3 Mottershead, D. G., Alftan, K., Ojala, K., Takkinen, K., & Oker-Blom, C. (2000). Baculoviral display of functional scFv and synthetic IgG-binding domains. *Biochemical and Biophysical Research Communications*, 275(1), 84-90.
- 4 Oker-Blom, C., Airene, K. J., & Grabherr, R. (2003). Baculovirus display strategies: Emerging tools for eukaryotic libraries and gene delivery. *Briefings in Functional Genomics and Proteomics*, 2(3), 244-253.
- 5 Edwards, B. M., Barash, S. C, Main, S. H., Choi, G. H., Minter, R., Ullrich, S., et al. (2003). The Remarkable Flexibility of the Human Antibody Repertoire; Isolation of Over One Thousand Different Antibodies to a Single Protein, BLYS. *Journal of Molecular Biology*, 334(1), 103-118. doi:10.1016/j.jmb.2003.09.054
- 6 Pershad, K., Pavlovic, J. D., Graslund, S., Nilsson, P., Colwill, K., Karatt-Vellatt, A., et al. (2010). Generating a panel of highly specific antibodies to 20 human SH2 domains by phage display. *Protein Engineering, Design and Selection*, 23(4), 279-288. doi:10.1093/protein/gzq003
- 7 Schofield, D. J., Pope, A. R., Clementel, V., Buckell, J., Chappie, S., Clarke, K. F., et al. (2007). Application of phage display to high throughput antibody generation and characterization. *Genome Biol*, 8(11), R254.
- 8 Salema, V., Marin, E., Martmez-Arteaga, R., Ruano-Gallego, D., Fraile, S., Margolles, Y., et al. (2013). Selection of Single Domain Antibodies from Immune Libraries Displayed on the Surface of *E. coli* Cells with Two  $\beta$ -Domains of Opposite Topologies. *PLoS ONE*, 8(9), e75126. doi:10.1371/journal.pone.0075126.s007
- 9 Chao, G., Lau, W. L, Hackel, B. J., Sazinsky, S. L, Lippow, S. M., & Wittrup, K. D. (2006). Isolating and engineering human antibodies using yeast surface display. *NatProtoc*, 1(2), 755-768.
- 10 Higuchi, K., Araki, T., Matsuzaki, O., Sato, A., Kanno, K., Kitaguchi, N., & Ito, H. (1997). Cell display library for gene cloning of variable regions of human antibodies to hepatitis B surface antigen. *J Immunol Methods*, 202(2), 193-204.

- 11 Ho, M., Nagata, S., & Pastan, I. (2006). Isolation of anti-CD22 Fv with high affinity by Fv display on human cells. *Proc Natl Acad Sci USA*, 103(25), 9637-9642.  
doi: 10.1073/pnas.0603653103
- 12 Ho, M., & Pastan, I. (2009). Display and selection of scFv antibodies on HEK-293T cells. *Methods Mol Biol*, 562, 99-113. doi: 10.1007/978-1-60327-302-2\_8
- 13 Akamatsu, Y., Pakabunto, K., Xu, Z., Zhang, Y., & Tsurushita, N. (2007). Whole IgG surface display on mammalian cells: Application to isolation of neutralizing chicken monoclonal anti-IL-12 antibodies. *Journal of Immunological Methods*, 327(1-2), 40-52.  
doi:10.1016/j.jim.2007.07.007
- 14 Beerli, R. R., Bauer, M., Buser, R. B., Gwerder, M., Muntwiler, S., Maurer, P., et al. Isolation of human monoclonal antibodies by mammalian cell display. *Proc Natl Acad Sci USA*. 2008 Sep 23; 105(38): 14336-41. doi: 10.1073/pnas.0805942105
- 15 Breous-Nystrom, E., Schultze, K., Meier, M., Flueck, L, Holzer, C, Boll, M., et al. (2013). Retrocyte Display<sup>®</sup> technology: Generation and screening of a high diversity cellular antibody library. *Methods*, 1-11. doi: 10.1016/j.ymeth.2013.09.003
- 16 Grindley, Whiteson & Rice. Mechanisms of Site-Specific Recombination. *Annu Rev Biochem* 2006 75:567-605
- 17 Zhou, C, Jacobsen, F. W., Cai, L, Chen, Q., & Shen, W. D. (2010). Development of a novel mammalian cell surface antibody display platform. *mAbs*, 2(5), 508-518
- 18 Li, C. Z., Liang, Z. K., Chen, Z. R., Lou, H. B., Zhou, Y., Zhang, Z. H., et al. (2012). Identification of HBsAg-specific antibodies from a mammalian cell displayed full-length human antibody library of healthy immunized donor. *Cellular and Molecular Immunology*, 9(2), 184-190.
- 19 Buchholz, F., Ringrose, L, Angrand, P. O., Rossi, F., & Stewart, A. F. (1996). Different thermostabilities of FLP and Cre recombinases: Implications for applied site-specific recombination. *Nucleic Acids Res*, 24(21), 4256-4262.
- 20 Schaft, J., Ashery-Padan, R., Van Hoeven, F. D., Gruss, P., & Francis Stewart, A. (2001). Efficient FLP recombination in mouse ES cells and oocytes. *Genesis*, 37(1), 6-10.
- 21 Moehle, E. A., Moehle, E. A., Rock, J. M., Rock, J. M., Lee, Y.-L, Lee, Y. L, et al. (2007). Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc Natl Acad Sci U S A*, 104(9), 3055-3060. doi:10.1073/pnas.0611478104
- 22 Cristea, S., Freyvert, Y., Santiago, Y., Holmes, M. C, Urnov, F. D., Gregory, P. D., & Cost, G. J. (2013). In vivo cleavage of transgene donors promotes nuclease-mediated targeted integration. *Biotechnology and Bioengineering*, 110(3), 871-880. doi:10.1002/bit.24733
- 23 Letourneur, F., & Malissen, B. (1989). Derivation of a T cell hybridoma variant deprived of functional T cell receptor alpha and beta chain transcripts reveals a nonfunctional alpha-mRNA of



- BW5147 origin. *European Journal of Immunology*, 19(12), 2269-2274. doi:10.1002/eji.1830191214
- 24 Kanayama, N., Todo, K., Reth, M., & Ohmori, H. (2005). Reversible switching of immunoglobulin hypermutation machinery in a chicken B cell line. *Biochem Biophys Res Commun*, 327(1), 70-75. doi:10.1016/j.bbrc.2004.11.143
- 25 Lin, W., Kurosawa, K., Murayama, A., Kagaya, E., & Ohta, K. (2011). B-cell display-based one-step method to generate chimeric human IgG monoclonal antibodies. *Nucleic Acids Research*, 39(3), e14-e14. doi:10.1093/nar/gkq1122
- 26 Adachi, N., So, S., Iizumi, S., Nomura, Y., Murai, K., Yamakawa, C, et al. (2006). The human pre-B cell line Nalm-6 is highly proficient in gene targeting by homologous recombination. *DNA and Cell Biology*, 25(1), 19-24. doi:10.1089/dna.2006.25.19
- 27 Palacios, R., & Steinmetz, M. (1985). IL3-Dependent mouse clones that express B-220 surface antigen, contain Ig genes in germ-line configuration, and generate B lymphocytes in vivo. *Cell*, 41(3), 727-734
- 28 Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A. A., Kim, S., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature*, 483(7391), 603-607. doi:10.1038/nature11003
- 29 Forbes, S. A., Bindal, N., Bamford, S., Cole, C, Kok, C. Y., Beare, D., et al. (2011). COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer. *Nucleic Acids Research*, 39(Database issue), D945-50. doi:10.1093/nar/gkq929
- 30 Silva, G., Poirot, L, Galetto, R., Smith, J., Montoya, G., Duchateau, P., & Paques, F. (2011). Meganucleases and other tools for targeted genome engineering: Perspectives and challenges for gene therapy. *Current Gene Therapy*, 11(1), 11-27
- 31 Epinat, J. C, Silva, G. H., Paques, F., Smith, J., & Duchateau, P. (2013). *Engineered meganucleases for genome engineering purposes. Topics in Current Genetics* (Vol. 23, pp. 147-185).
- 32 Szczepek, M., Brondani, V., Buchel, J., Serrano, L, Segal, D. J., & Cathomen, T. (2007). Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nature Biotechnology*, 25(7), 786-793. doi:10.1038/nbt1317
- 33 Doyon, Y., Vo, T. D., Mendel, M. C, Greenberg, S. G., Wang, J., Xia, D. F., et al. (2011). Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. *Nat Methods*, 8(1), 74-79. doi:10.1038/nmeth.1539
- 34 Perez-Pinera, P., Ousterout, D. G., Brown, M. T., & Gersbach, C. A. (2012). Gene targeting to the ROSA26 locus directed by engineered zinc finger nucleases. *Nucleic Acids Research*, 40(8), 3741-3752. doi:10.1093/nar/gkr1214

- 35 Urnov, F. D., Miller, J. C., Lee, Y.-L., Beausejour, C. M., Rock, J. M., Augustus, S., et al. (2005). Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, 435(7042), 646-651. doi:10.1038/nature03556
- 36 Maresca, M., Lin, V. G., Guo, N., & Yang, Y. (2013). Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining. *Genome Research*, 23(3), 539-546. doi:10.1101/gr.145441.112
- 37 Bogdanove, A. J., & Voytas, D. F. (2011). TAL effectors: customizable proteins for DNA targeting. *Science*, 333(6051), 1843-1846. doi:10.1126/science.1204094
- 38 Reyon, D., Tsai, S. Q., Khgayter, C., Foden, J. A., Sander, J. D., & Joung, J. K. (2012). FLASH assembly of TALENs for high-throughput genome editing. *Nature Biotechnology*, 30(5), 460-465. doi:10.1038/nbt.2170
- 39 Boissel, S., Jarjour, J., Astrakhan, A., Adey, A., Gouble, A., Duchateau, P., et al. (2013). megaTALs: a rare-cleaving nuclease architecture for therapeutic genome engineering. *Nucleic Acids Research*. doi:10.1093/nar/gkt1224
- 40 Beurdeley, M., Bietz, F., Li, J., Thomas, S., Stoddard, T., Juillerat, A., et al. (2013). Compact designer TALENs for efficient genome engineering. *Nature Communications*, 4, 1762. doi: 10.1038/ncomms2782
- 41 Sampson, T. R., & Weiss, D. S. (2014). Exploiting CRISPR/Cas systems for biotechnology. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, 36(1), 34-38. doi:10.1002/bies.201300135
- 42 Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelsen, T. S., et al. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. *Science*, 343(6166), 84-87. doi: 10.1126/science.1247005
- 43 Wang, T., Wei, J. J., Sabatini, D. M., & Lander, E. S. (2014). Genetic screens in human cells using the CRISPR-Cas9 system. *Science*, 343(6166), 80-84. doi: 10.1126/science.1246981
- 44 Beard, C., Hochedlinger, K., Plath, K., Wutz, A., & Jaenisch, R. (2006). Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. *Genesis*, 44(1), 23-28.
- 45 Orlando, S. J., Santiago, Y., DeKever, R. C., Freyvert, Y., Boydston, E. A., Moehle, E. A., et al. (2010). Zinc-finger nuclease-driven targeted integration into mammalian genomes using donors with limited chromosomal homology. *Nucleic Acids Research*, 38(15), e152. doi:10.1093/nar/gkq512
- 46 Cadinanos 2007
- 47 Zhang, H., Yea, K., Xie, J., Ruiz, D., Wilson, I. A., & Lerner, R. A. (2013). Selecting agonists from single cells infected with combinatorial antibody libraries. *Chemistry & Biology*, 20(5), 734-

741.doi:10.1016/j.chembiol.2013.04.012

48 Porteus, M. H., & Baltimore, D. (2003). Chimeric nucleases stimulate gene targeting in human cells. *Science*, 300(5620), 763. doi:10.1126/science.1078395

49 Rouet, P., Smith, F., & Jasin, M. (1994). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Molecular and Cellular Biology*, 14(12), 8096-8106.

50 Jasin, M. (1996). Genetic manipulation of genomes with rare-cutting endonucleases. *Trends in Genetics*, 12(6), 224-228

51 Davis, L., & Maizels, N. (2011). DNA nicks promote efficient and safe targeted gene correction. *PLoS ONE*, 6(9), e23981. doi:10.1371/journal.pone.0023981

52 Fujioka, K., Aratani, Y., Kusano, K., & Koyama, H. (1993). Targeted recombination with single-stranded DNA vectors in mammalian cells. *Nucleic Acids Res*, 27(3), 407-412

53 Khan, I. F., Hirata, R. K., & Russell, D. W. (2011). AAV-mediated gene targeting methods for human cells. *Nat Protoc*, 6(4), 482-501. doi:10.1038/nprot.2011.301

54 Deyle, D. R., & Russell, D. W. (2009). Adeno-associated virus vector integration. *Current Opinion in Molecular Therapeutics*, 11(4), 442-447

55 Benatuil, L., Perez, J. M., Belk, J., & Hsieh, C. M. (2010). An improved yeast transformation method for the generation of very large human antibody libraries. *Protein Engineering, Design and Selection*, 23(4), 155-159.

56 Feldhaus, M. J., Siegel, R. W., Opresko, L. K., Coleman, J. R., Feldhaus, J. M., Yeung, Y. A., et al. (2003). Flow-cytometric isolation of human antibodies from a nonimmune *Saccharomyces cerevisiae* surface display library. *Nat Biotechnol*.

57 Zhao, A., Nunez-Cruz, S., Li, C., Coukos, G., Siegel, D. L., & Scholler, N. (2011). Rapid isolation of high-affinity human antibodies against the tumor vascular marker Endosialin/TEM1, using a paired yeast-display/secretory scFv library platform. *Journal of Immunological Methods*, 363(2), 221-232. doi:10.1016/j.jim.2010.09.001

58 Skerra, A. (2007). Alternative non-antibody scaffolds for molecular recognition. *Curr Opin Biotechnol*, 18(4), 295-304. doi:10.1016/j.copbio.2007.04.010

59 Gebauer, M., & Skerra, A. (2009). Engineered protein scaffolds as next-generation antibody therapeutics. *Current Opinion in Chemical Biology*, 13(3), 245-255. doi:10.1016/j.cbpa.2009.04.627

60 Haan & Maggos (2004) *BioCentury*, 12(5): A1-A6

61 Koide et al. (1998) *Journal of Molecular Biology*, 284: 1141-1151.

62 Nygren et al. (1997) *Current Opinion in Structural Biology*, 7: 463-469

63 Wess, L. In: *BioCentury, The Bernstein Report on BioBusiness*, 12(42), A1-A7, 2004

- 64 Chang, H.-J., Hsu, H.-J., Chang, C.-F., Peng, H.-P., Sun, Y.-K., Yu, H.-M., et al. (2009). Molecular Evolution of Cystine-Stabilized Mini-proteins as Stable Proteinaceous Binders. *Structure*, 17(4), 620-631. doi:10.1016/j.str.2009.01.011
- 65 Ward, E.S. et al., *Nature* 341, 544-546 (1989)
- 66 McCafferty et al *Nature*, 348, 552-554 (1990)
- 67 Holt et al *Trends in Biotechnology* 21, 484-490 (2003)
- 68 Bird et al, *Science*, 242, 423-426, (1988)
- 69 Huston et al, *PNAS USA*, 85, 5879-5883, (1988)
- 70 Holliger, P. et al, *PNAS USA* 90 6444-6448, (1993)
- 71 Reiter, Y. et al, *Nature Biotech*, 14, 1239-1245, (1996)
- 72 Holliger & Hudson, *Nature Biotechnology* 23(9): 1126-1136 (2005)
- 73 Knappik et al. *J. Mol. Biol.* 296, 57-86 (2000)
- 74 Krebs et al. *Journal of Immunological Methods* 254, 67-84 (2001)
- 75 Holliger and Bohlen *Cancer and metastasis rev.* 18: 411-419 (1999)
- 76 Holliger, P. and Winter G. *Current Opinion Biotechnol* 4, 446-449 (1993)
- 77 Glennie M J et al., *J. Immunol.* 139, 2367-2375 (1987)
- 78 Repp R. et al., *J. Hemat.* 377-382 (1995)
- 79 Staerz U. D. and Bevan M. J. *PNAS* 83 (1986)
- 80 Suresh M. R. et al., *MethodEnzymol.* 121: 210-228 (1986)
- 81 Merchand et al., *Nature Biotech.* 16:677-681 (1998)
- 82 Ridgeway, J. B. B. et al, *Protein Eng.*, 9, 616-621, (1996)
- 83 Marks, J. D., Hoogenboom, H. R., Bonnert, T. P., McCafferty, J., Griffiths, A. D., & Winter, G. (1991). By-passing immunization. Human antibodies from V-gene libraries displayed on phage. *J Mol Biol*, 222(3), 581-597.
- 84 Gronwald, R. G. K., Grant, F. J., Haldeman, B. A., Hart, C. E., O'Hara, P. J., Hagen, F. S., et al. (1988). Cloning and expression of a cDNA coding for the human platelet-derived growth factor receptor: Evidence for more than one receptor class (Vol. 85, pp. 3435-3439). Presented at the Proceedings of the National Academy of Sciences of the United States of America.
- 85 Kumar, N., & Borth, N. (2012). Flow-cytometry and cell sorting: an efficient approach to investigate productivity and cell physiology in mammalian cell factories. *Methods*, 56(3), 366-374. doi:10.1016/j.ymeth.2012.03.004
- 86 Brezinsky, S. C. G., Chiang, G. G., Szilvasi, A., Mohan, S., Shapiro, R. I., MacLean, A., et al. (2003). A simple method for enriching populations of transfected CHO cells for cells of higher specific productivity. *J Immunol Methods*, 277(1-2), 141-155.
- 87 Pichler, J., Hesse, F., Wieser, M., Kunert, R., Galosy, S. S., Mott, J. E., & Borth, N. (2009). A

- study on the temperature dependency and time course of the cold capture antibody secretion assay. *Journal of Biotechnology*, 141(1-2), 80-83. doi:10.1016/j.jbiotec.2009.03.001
- 88 Anastassiadis, K., Fu, J., Patsch, C., Hu, S., Weidlich, S., Duerschke, K., et al. (2009). Dre recombinase, like Cre, is a highly efficient site-specific recombinase in *E. coli*, mammalian cells and mice. *Disease Models & Mechanisms*, 2(9-10), 508-515. doi:10.1242/dmm.003087
- 89 Horlick, R. A., Macomber, J. L., Bowers, P. M., Neben, T. Y., Tomlinson, G. L., Krapf, I. P., et al. (2013). Simultaneous surface display and secretion of proteins from mammalian cells facilitate efficient in vitro selection and maturation of antibodies. *J Biol Chem*, 288(27), 19861-19869. doi:10.1074/jbc.M113.452482
- 90 Biffi, G., Tannahill, D., McCafferty, J., & Balasubramanian, S. (2013). Quantitative visualization of DNA G-quadruplex structures in human cells. *Nature Chemistry*, 5(3), 182-186. doi:10.1038/nchem.1548
- 91 Gao, J., Sidhu, S. S., & Wells, J. A. (2009). Two-state selection of conformation-specific antibodies. *Proc Natl Acad Sci U S A*, 106(9), 3071-3076. doi:10.1073/pnas.0812952106
- 92 Gu, G. J., Friedman, M., Jost, C., Johnsson, K., Kamali-Moghaddam, M., Plückthun, A., et al. (2013). Protein tag-mediated conjugation of oligonucleotides to recombinant affinity binders for proximity ligation. *N Biotechnol*, 30(2), 144-152. doi:10.1016/j.nbt.2012.05.005
- 93 Cho, Y. K., & Shusta, E. V. (2010). Antibody library screens using detergent-solubilized mammalian cell lysates as antigen sources. *Protein Eng Des Sel*, 23(7), 567-577. doi:10.1093/protein/gzq029
- 94 Tillotson, B. J., Cho, Y. K., & Shusta, E. V. (2013). Cells and cell lysates: a direct approach for engineering antibodies against membrane proteins using yeast surface display. *Methods*, 60(1), 27-37. doi:10.1016/j.ymeth.2012.03.010
- 95 Kunert, A., Straetemans, T., Govers, C., Lamers, C., Mathijssen, R., Sleijfer, S., & Debets, R. (2013). TCR-Engineered T Cells Meet New Challenges to Treat Solid Tumors: Choice of Antigen, T Cell Fitness, and Sensitization of Tumor Milieu. *Frontiers in Immunology*, 4, 363. doi: 10.3389/fimmu.2013.00363
- 96 Liddy, N., Bossi, G., Adams, K. J., Lissina, A., Mahon, T. M., Hassan, N. J., et al. (2012). Monoclonal TCR-redirectioned tumor cell killing. *Nature Medicine*, 18(6), 980-987. doi:10.1038/nm.2764
- 97 Holler, P. D., Holman, P. O., Shusta, E. V., O'Herrin, S., Wittrup, K. D., & Kranz, D. M. (2000). In vitro evolution of a T cell receptor with high affinity for peptide/MHC. *Proc Natl Acad Sci USA*, 97(10), 5387-5392. doi:10.1073/pnas.080078297
- 98 Weber, K. S., Donermeyer, D. L., Allen, P. M., & Kranz, D. M. (2005). Class II-restricted T cell receptor engineered in vitro for higher affinity retains peptide specificity and function. *Proc*

*Natl Acad Sci U S A*, 102(52), 19033-19038. doi:10.1073/pnas.0507554102

99 Kessels, H. W., van Den Boom, M. D., Spits, H., Hooijberg, E., & Schumacher, T. N. (2000). Changing T cell specificity by retroviral T cell receptor display. *Proc Natl Acad Sci USA*, 97(26), 14578-14583. doi: 10.1073/pnas.97.26.14578

100 Chervin, A. S., Aggen, D. H., Raseman, J. M., & Kranz, D. M. (2008). Engineering higher affinity T cell receptors using a T cell display system. *Journal of Immunological Methods*, 339(2), 175-184.

101 Crawford, F., Jordan, K. R., Stadinski, B., Wang, Y., Huseby, E., Marrack, P., et al. (2006). Use of baculovirus MHC/peptide display libraries to characterize T-cell receptor ligands. *Immunological Reviews*, 210, 156-170. doi:10.1111/j.0105-2896.2006.00365.x

102 Hinrichs, C. S., & Restifo, N. P. (2013). Reassessing target antigens for adoptive T-cell therapy. *Nat Biotechnol*, 37(11), 999-1008. doi:10.1038/nbt.2725

103 Sadelain, M., Brentjens, R., & Rivière, I. (2013). The basic principles of chimeric antigen receptor design. *Cancer Discovery*, 3(4), 388-398. doi: 10.1158/2159-8290.CD-12-0548

104 Alonso-Camino, V., Sánchez-Martín, D., Compte, M., Sanz, L., & Alvarez-Vallina, L. (2009). Lymphocyte display: a novel antibody selection platform based on T cell activation. *PLoS ONE*, 4(9), e7174. doi:10.1371/journal.pone.0007174

105 Melidoni, A. N., Dyson, M. R., Wormald, S., & McCafferty, J. (2013). Selecting antagonistic antibodies that control differentiation through inducible expression in embryonic stem cells. *Proceedings of the National Academy of Sciences*, 110(44), 17802-17807. doi: 10.1073/pnas.1312062110

106 Zhang, H., Wilson, I. A., & Lerner, R. A. (2012). Selection of antibodies that regulate phenotype from intracellular combinatorial antibody libraries. *Proceedings of the National Academy of Sciences*, 709(39), 15728-15733. doi:10.1073/pnas.1214275109

107 Xie, J., Yea, K., Zhang, H., Moldt, B., He, L., Zhu, J., & Lerner, R. A. (2014). Prevention of cell death by antibodies selected from intracellular combinatorial libraries. *Chemistry & Biology*, 21(2), 274-283.

108 Yea, K., Zhang, H., Xie, J., Jones, T. M., Yang, G., Song, B. D., & Lerner, R. A. (2013). Converting stem cells to dendritic cells by agonist antibodies from unbiased morphogenic selections (Vol. 110, pp. 14966-14971). Presented at the Proceedings of the National Academy of Sciences of the United States of America. doi:10.1073/pnas.1313671110

109 Kawahara, M., Kimura, H., Ueda, H., & Nagamune, T. (2004). Selection of genetically modified cell population using hapten-specific antibody/receptor chimera. *Biochem Biophys Res Commun*, 375(1), 132-138. doi:10.1016/j.bbrc.2004.01.030

110 Kawahara, M., Shimo, Y., Sogo, T., Hitomi, A., Ueda, H., & Nagamune, T. (2008). Antigen-

- mediated migration of murine pro-B Ba/F3 cells via an antibody/receptor chimera. *Journal of Biotechnology*, 133(1), 154-161. doi:10.1016/j.jbiotec.2007.09.009
- 111 Sogo, T., Kawahara, M., Ueda, H., Otsu, M., Onodera, M., Nakauchi, H., & Nagamune, T. (2009). T cell growth control using hapten-specific antibody/interleukin-2 receptor chimera. *Cytokine*, 46(1), 127-136. doi: 10.1016/j.cyto.2008.12.020
- 112 Kawahara, M., Chen, J., Sogo, T., Teng, J., Otsu, M., Onodera, M., et al. (2011). Growth promotion of genetically modified hematopoietic progenitors using an antibody/c-MpI chimera. *Cytokine*, 55(3), 402-408. doi:10.1016/j.cyto.2011.05.024
- 113 Ueda, H., Kawahara, M., Aburatani, T., Tsumoto, K., Todokoro, K., Suzuki, E., et al. (2000). Cell-growth control by monomeric antigen: the cell surface expression of lysozyme-specific Ig V-domains fused to truncated Epo receptor. *J Immunol Methods*, 241(1-2), 159-170.
- 114 Kerppola, T. K. (2009). Visualization of molecular interactions using bimolecular fluorescence complementation analysis: Characteristics of protein fragment complementation. *Chemical Society Reviews*, 38(10), 2876-2886.
- 115 Michnick, S. W., Ear, P. H., Manderson, E. N., Remy, I., & Stefan, E. (2007). Universal strategies in research and drug discovery based on protein-fragment complementation assays. *Nature Reviews Drug Discovery*, 6(7), 569-582. doi:10.1038/nrd231
- 116 Petschnigg, J., Groisman, B., Kotlyar, M., Taipale, M., Zheng, Y., Kurat, C. F., et al. (2014). The mammalian-membrane two-hybrid assay (MaMTH) for probing membrane-protein interactions in human cells. *Nat Methods*. doi:10.1038/nmeth.2895
- 117 Renaut, L, Monnet, C, Dubreuil, O., Zaki, O., Crozet, F., Bouayadi, K., et al. (2012). *Affinity maturation of antibodies: Optimized methods to generate high-quality scfv libraries and isolate igg candidates by high-throughput screening. Methods in Molecular Biology* (Vol. 907, pp. 451-461)
- 118 Dyson, M. R., Zheng, Y., Zhang, C, Colwill, K., Pershad, K., Kay, B. K., et al. (2011). Mapping protein interactions by combining antibody affinity maturation and mass spectrometry. *AnalBiochem*, 417(1), 25-35. doi:10.1016/j.ab.2011.05.005
- 119 de Felipe P (2002) Polycistronic viral vectors. *Curr Gene Ther* 2: 355-378. doi: 10.2174/1566523023347742.
- 120 Foote, J., & Winter, G. (1992). Antibody framework residues affecting the conformation of the hypervariable loops. *J Mol Biol*, 224(2), 487-499.
- 121 Massie, B., Dionne, J., Lamarche, N., Fleurent, J., & Langelier, Y. (1995). Improved adenovirus vector provides herpes simplex virus ribonucleotide reductase R1 and R2 subunits very efficiently. *Nature Biotechnology*, 13(6), 602-608.
- 122 Kim, D. W., Uetsuki, T., Kaziro, Y., Yamaguchi, N., & Sugano, S. (1990). Use of the human

elongation factor 1 alpha promoter as a versatile and efficient expression system. *Gene*, 91(2), 217-223.

123 Holden, P., Keene, D. R., Lunstrum, G. P., Bachinger, H. P., & Horton, W. A. (2005). Secretion of cartilage oligomeric matrix protein is affected by the signal peptide. *J Biol Chem*, 280(17), 17172-17179.

124 Sadelain, M., Papapetrou, E. P., & Bushman, F. D. (2012). Safe harbours for the integration of new DNA in the human genome. *Nature Reviews Cancer*, 12(1), 51-58.

125 Sanjana, N. E., Cong, L., Zhou, Y. (Cunniff, M. M., Feng, G., & Zhang, F. (2012). A transcription activator-like effector toolbox for genome engineering. *NatProtoc*, 7(1), 171-192. doi: 10.1038/nprot.2011.431

126 Falk, R., Falk, A., Dyson, M. R., Melidoni, A. N., Parthiban, K., Young, J. L., et al. (2012). Generation of anti-Notch antibodies and their application in blocking Notch signalling in neural stem cells. *Methods*, 58(1), 69-78. doi: 10.1016/j.ymeth.2012.07.008

127 Martin, C. D., Rojas, G., Mitchell, J. N., Vincent, K. J., Wu, J., McCafferty, J., & Schofield, D. J. (2006). A simple vector system to improve performance and utilisation of recombinant antibodies. *BMC Biotechnology*, 6, 46.

128 Reyon, D., Tsai, S. Q., Khgayter, C, Foden, J. A., Sander, J. D., & Joung, J. K. (2012). FLASH assembly of TALENs for high-throughput genome editing. *Nature Biotechnology*, 30(5), 460-465. doi:10.1038/nbt.2170

129 Van Der Weyden, L, Adams, D. J., Harris, L. W., Tannahill, D., Arends, M. J., & Bradley, A. (2005). Null and conditional Semaphorin 3B alleles using a flexible puroAtk LoxP/FRT vector. *Genesis*, 41(4), 171-178

130 de Felipe, P., & Ryan, M. D. (2004). Targeting of proteins derived from self-processing polyproteins containing multiple signal sequences. *Traffic*, 5(8), 616-626.

131 Raymond, C. S., & Soriano, P. (2007). High-efficiency FLP and PhiC31 site-specific recombination in mammalian cells. *PLoS ONE*, 2(1), e162. doi:10.1371/journal.pone.0000162

132 Kranz, A., Fu, J., Duerschke, K., Weidlich, S., Naumann, R., Stewart, A. F., & Anastassiadis, K. (2010). An improved Flp deleter mouse in C57BI/6 based on Flpo recombinase. *Genesis*, 48(8), 512-520. doi:10.1002/dvg.20641

133 Szymczak AL, Vignali DA (2005) Development of 2A peptide-based strategies in the design of multicistronic vectors. *Expert Opin Biol Ther* 5: 627-638. doi: 10.1517/14712598.5.5.627

134 Chappie, S.D., Crofts, A.M., Shadbolt, S.P., McCafferty, J., and Dyson, M.R. (2006). Multiplexed expression and screening for recombinant protein production in mammalian cells. *BMC biotechnology* 6, 49.

135 Zhao, Y. *et al.* Multiple injections of electroporated autologous T cells expressing a chimeric



- antigen receptor mediate regression of human disseminated tumor. *Cancer Research* 70, 9053-9061 (2010).
136. Szymczak, A. L. *et al.* Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. *Nature biotechnology* **22**, 589-594 (2004).
137. Li, Y. *et al.* Directed evolution of human T-cell receptors with picomolar affinities by phage display. *Nat. Biotechnol.* **23**, 349-354 (2005).
138. Zhao, Y. *et al.* High-affinity TCRs generated by phage display provide CD4<sup>+</sup> T cells with the ability to recognize and kill tumor cell lines. *J. Immunol.* **179**, 5845-5854 (2007).
139. Madura, F. *et al.* T-cell receptor specificity maintained by altered thermodynamics. *The Journal of biological chemistry* **288**, 18766-18775 (2013).
140. Pierce, B. G. *et al.* Computational design of the affinity and specificity of a therapeutic T cell receptor. *PLoS Comput Biol* **10**, e1003478 (2014).
141. Sebestyen, Z. *et al.* Human TCR that incorporate CD3zeta induce highly preferred pairing between TCRalpha and beta chains following gene transfer. *J. Immunol.* **180**, 7736-7746 (2008).
142. Roszik, J. *et al.* T-cell synapse formation depends on antigen recognition but not CD3 interaction: studies with TCR:£, a candidate transgene for TCR gene therapy. *Eur. J. Immunol.* **41**, 1288-1297 (2011).
143. Cohen, C. J., Zhao, Y., Zheng, Z., Rosenberg, S. A. & Morgan, R. A. Enhanced antitumor activity of murine-human hybrid T-cell receptor (TCR) in human lymphocytes is associated with improved pairing and TCR/CD3 stability. *Cancer Research* **66**, 8878-8886 (2006).
144. Huovinen, T. *et al.* Primer extension mutagenesis powered by selective rolling circle amplification. *PLoS ONE* **7**, e31817 (2012).
145. Cribbs, A. P., Kennedy, A., Gregory, B. & Brennan, F. M. Simplified production and concentration of lentiviral vectors to achieve high transduction in primary human T cells. *BMC biotechnology* **13**, 98 (2013).
146. Oelke, M. *et al.* Ex vivo induction and expansion of antigen-specific cytotoxic T cells by HLA-Ig-coated artificial antigen-presenting cells. *Nat Med* **9**, 619-624 (2003).
147. Wölf I, M. & Greenberg, P. D. Antigen-specific activation and cytokine-facilitated expansion of naive, human CD8<sup>+</sup> T cells. *Nature Protocols* **9**, 950-966 (2014).
148. Lipowska-Bhalla, G., Gilham, D. E., Hawkins, R. E. & Rothwell, D. G. Isolation of tumor antigen-specific single-chain variable fragments using a chimeric antigen receptor bicistronic retroviral vector in a Mammalian screening protocol. *Hum Gene Ther Methods* **24**, 381-391 (2013).
149. Kelly, R. J., Sharon, E., Pastan, I. & Hassan, R. Mesothelin-targeted agents in clinical trials

- and in preclinical development. *Mol. Cancer Ther.* 11, 517-525 (2012).
150. Atanackovic, D. *et al.* Surface molecule CD229 as a novel target for the diagnosis and treatment of multiple myeloma. *Haematologica* 96,1512-1520 (2011).
151. Bund, D., Mayr, C, Kofler, D. M., Haliak, M. & Wendtner, C.-M. Human Ly9 (CD229) as novel tumor-associated antigen (TAA) in chronic lymphocytic leukemia (B-CLL) recognized by autologous CD8+ T cells. *Exp. Hematol.* 34, 860-869 (2006).
152. Tiede, C. *et al.* Adhiron: a stable and versatile peptide display scaffold for molecular recognition applications. *Protein Eng. Des. Sel.* 27,145-155 (2014).
153. Maresca, M., Lin, V. G., Guo, N. & Yang, Y. Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining. *Genome Res.* 23, 539-546 (2013).
154. McVey, M. & Lee, S. E. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet.* 24, 529-538 (2008).
155. Nakade, S. *et al.* Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. *Nat Commun* 5, 5560 (2014).
156. Chiche, L. *et al.* Squash inhibitors: From structural motifs to macrocyclic knottins. *Current Protein & Peptide Science* 5, 341-349.

ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

<110> Iontas Limited

<120> ПОЛУЧЕНИЕ БИБЛИОТЕК ВАРИАНТОВ БЕЛКОВ, ЭКСПРЕССИРОВАННЫХ В ЭУКАРИОТИЧЕСКИХ КЛЕТКАХ, И ПРИМЕНЕНИЕ ИХ ДЛЯ ОТБОРА СВЯЗЫВАЮЩИХ МОЛЕКУЛ

<130> SMW/CP7119530

<150> GB1407852.1

<151> 2014-05-02

<160> 142

<170> PatentIn version 3.5

<210> 1

<211> 5396

<212> DNA

<213> Artificial Sequence

<220>

<223> pD1 dual promoter antibody expression cassette for surface expression

<400> 1

ggtaccgaat tccgtgaggc tccggtgcc gtcagtgggc agagcgcaca tcgcccacag	60
tccccgagaa gttgggggga ggggtcggca attgaaccgg tgccatagaga aggtggcgcg	120
gggtaaactg ggaaagtgat gtcgtgtact ggctccgcct ttttcccgag ggtgggggag	180
aaccgtatat aagtgcagta gtcgccgtga acgttctttt tcgcaacggg tttgccgcca	240
gaacacaggt aagtgccgtg tgtggttccc gcgggcctgg cctctttacg ggttatggcc	300
cttgcgtgcc ttgaattact tccacctggc tccagtacgt gattcttgat cccgagctgg	360
agccaggggc gggccttgcg ctttaggagc cccttcgcct cgtgcttgag ttgaggcctg	420
gcctgggagc tggggccgcc gcgtgcgaat ctgggtggcac cttcgcgcct gtctcgctgc	480
tttcgataag tctctagcca tttaaaattt ttgatgacct gctgcgacgc ttttttctg	540
gcaagatagt cttgtaaagc cgggccagga tctgcacact ggtatctcgg tttttgggccc	600
cgcgccgggc gacggggccc gtgctgccca gcgcacatgt tcggcgaggc ggggcctgag	660
agcgcgcca ccgagaatcg gacgggggta gtctcaagct ggccggcctg ctctgggtgc	720
tggcctcgcg ccgccgtgta tcgccccgcc ctggggcgga aggctggccc ggtcggcacc	780
agttgctgta gcggaagat ggccgcttcc cggccctgct ccagggggct caaaatggag	840
gacgcggcgc tcgggagagc gggcggtgta gtcaccaca caaaggaaaa gggcctttcc	900
gtcctcagcc gtcgcttcat gtgactccac ggagtaccgg gcgcccgtcca ggcacctcga	960
ttagttctgg agcttttgga gtacgtcgtc tttaggttgg ggggaggggt tttatgcat	1020
ggagtttccc cacactgagt ggggtggagac tgaagttagg ccagcttggc acttgatgta	1080
attctcgttg gaatttgccc tttttgagtt tggatcttgg ttcattctca agcctcagac	1140

agtggttcaa	agtttttttc	ttccatttca	ggtgtcgtga	gacgtggcca	ccatgagggc	1200
ctggatcttc	tttctccttt	gcctggcccg	gagggctctg	gcagctagcg	acatccagat	1260
gaccagagc	ccaagcagcc	tgagcgccag	cgtgggtgac	agagtgacca	tcacctgtag	1320
agccagcgg	aacatccaca	actacctggc	ttggtaccag	cagaagccag	gtaaggctcc	1380
aaagctgctg	atctactaca	ccaccaccct	ggctgacgg	gtgccaaagca	gattcagcgg	1440
tagcggtagc	ggtaccgact	acaccttcac	catcagcagc	ctccagccag	aggacatcgc	1500
cacctactac	tgccagcact	tctggagcac	cccaaggacg	ttcggccaag	ggaccaaggt	1560
ggaaatcaaa	cgtaccgcg	ccgccccttc	cgtgttcac	ttccctccct	ccgacgagca	1620
gctgaagtcc	ggcaccgcct	ctgtggtgtg	cctgctgaac	aacttctacc	ctcgggaggc	1680
caaggtgcag	tggaaggtgg	acaacgccct	gcagtcggc	aactcccagg	aatccgtcac	1740
cgagcaggac	tccaaggact	ctacctactc	cctgtcctcc	accctgacct	tgtccaaggc	1800
cgactacgag	aagcacaagc	tgtacgcctg	cgaagtgacc	caccagggcc	tgtcctctcc	1860
cgtgaccaag	tccttcaacc	ggggcgagtg	ctaataaaag	cttacgacgt	gatcagcctc	1920
gactgtgcct	tctagttgcc	agccatctgt	tgtttgcccc	tcccccgctg	cttccttgac	1980
cctggaaggt	gccactccca	ctgtcctttc	ctaataaaat	gaggaaattg	catcgcattg	2040
tctgagtagg	tgtcattcta	ttctgggggg	tggggtgggg	caggacagca	agggggagga	2100
ttgggaagac	aatagcaggc	atgctgggga	acattgatta	ttgactagtt	attaatagta	2160
atcaattacg	gggtcattag	ttcatagccc	atatatggag	ttccgcgtta	cataacttac	2220
ggtaaattgg	ccgcctggct	gaccgccc	cgacccccgc	ccattgacgt	caataatgac	2280
gtatgttccc	atagtaacgc	caatagggac	tttccattga	cgtcaatggg	tggagtat	2340
acggtaaact	gcccacttgg	cagtacatca	agtgtatcat	atgccaagtc	cgccccctat	2400
tgacgtcaat	gacggtaaat	ggcccgcctg	gcattatgcc	cagtacatga	ccttacggga	2460
ctttcctact	tggcagtaca	tctacgtatt	agtcatecgt	attaccatag	tgatgcgggt	2520
ttggcagtac	accaatgggc	gtggatagcg	gtttgactca	cggggatttc	caagtctcca	2580
cccattgac	gtcaatggga	gtttgttttg	gcacccaaat	caacgggact	ttccaaaatg	2640
tcgtaataac	cccgccccgt	tgacgcaaat	ggcggttagg	cgtgtacgg	gggaggtcta	2700
tataagcaga	gctcgtttag	tgaaccgtca	gatcctcact	ctcttccgca	tcgctgtctg	2760
cgagggccag	ctgttgggct	cgcggttgag	gacaaactct	tcgcggtctt	tccagtactc	2820
ttggatcgg	aaccgcctcg	cctccgaacg	gtactccgcc	accgagggac	ctgagcgagt	2880
ccgcatcgac	cggatcggaa	aacctctcgt	gaaaggcgtc	taaccagtca	cagtcgcaag	2940
gtaggctgag	caccgtggcg	ggcggcagcg	ggtggcggtc	ggggttggtt	ctggcggagg	3000
tgctgctgat	gatgtaatta	aagtaggcgg	tcttgagacg	gcggtgggtc	gaggtgaggt	3060

gtggcaggct	tgagatccag	ctgttgggggt	gagtactccc	tctcaaaagc	gggcattact	3120
tctgcgctaa	gattgtcagt	ttccaaaaac	gaggaggatt	tgatattcac	ctggcccgat	3180
ctggccatac	acttgagtga	caatgacatc	cactttgcct	ttctctccac	aggtgtccac	3240
tcccaggtcc	aagtttgtgg	aaattaatac	gacgtggcca	ccatgagttg	gagctgtatc	3300
atcctcttct	tggtagcaac	agctacaggt	aaggggttaa	cagtagcagg	cttgaggctt	3360
ggacataat	atgggtgaca	atgacatcca	ctttgccttt	ctctccacag	gcgccatggc	3420
ccaggtccaa	ctgcaggaga	gcggtccagg	tcttgtgaga	cctagccaga	ccctgagcct	3480
gacctgcacc	gtgtctggca	gcaccttcag	cggtatgggt	gtaaactggg	tgagacagcc	3540
acctggacga	ggtcttgagt	ggattggaat	gatttgggggt	gatggaaaca	cagactataa	3600
ttcagctctc	aatccagag	tgacaatgct	ggtagacacc	agcaagaacc	agttcagcct	3660
gagactcagc	agcgtgacag	ccgccgacac	cgcggtctat	tattgtgcaa	gagagagaga	3720
ttataggctt	gactactggg	gtcaaggcag	cctcgtcaca	gtctcgagtg	ccagcaccaa	3780
gggccccagc	gtgttcctc	tggccccctg	tagcagaagc	accagcgaga	gcacagccgc	3840
cctgggctgc	ctgggtcaagg	actacttccc	cgagcccgtg	accgtgtcct	ggaactctgg	3900
cgctctgacc	agcggcgtgc	acacctttcc	agccgtgctg	cagagcagcg	gcctgtacag	3960
cctgagcagc	gtggtcaccg	tgcccagcag	caacttcggc	accagacct	acacctgtaa	4020
cgtggaccac	aagcccagca	acaccaaggt	ggacaagacc	gtggaacgga	agtgctgcgt	4080
ggaatgcccc	ccctgtcccg	ctcctccagt	ggctggacct	tccgtgttcc	tgttcccccc	4140
aaagcccaag	gacaccctga	tgatcagccg	gacccccgaa	gtgacctgcg	tggtgggtgga	4200
cgtgtcccac	gaggaccccg	aggtgcagtt	caattggtac	gtggacggcg	tggaagtgca	4260
caacgccaag	accaagccca	gagaggaaca	gttcaacagc	accttccggg	tggtgtccgt	4320
gctgaccgtg	gtgcaccagg	actggctgaa	cggcaaagag	tacaagtgcg	ccgtctccaa	4380
caagggcctg	cctgccccca	tcgagaaaac	catcagcaag	accaagggcc	agcctcgoga	4440
gcctcaggtg	tacacactgc	ccccagccg	ggaagagatg	accaagaacc	aggtgtccct	4500
gacctgcctc	gtgaagggtt	tctaccccag	cgatatcgcc	gtggaatggg	agagcaacgg	4560
ccagcccag	aacaactaca	agaccacccc	ccccatgctg	gacagcgacg	gctcattctt	4620
cctgtacagc	aagctgacag	tggacaagag	ccggtggcag	cagggcaacg	tgttcagctg	4680
cagcgtgatg	cacgaggccc	tgcacaacca	ctacaccag	aagtccttga	gcctgagccc	4740
cggcaaggg	tccaaggtaa	gtttaaacat	atatataact	ttaaataatt	ggcattattt	4800
aaagttacta	ctaactaacc	ctgattattt	aaatthtcag	gaacaaaaac	tcattctcaga	4860
agaggatctg	aatgctgtgg	gccaggacac	gcaggaggtc	atcgtgggtg	cacactcctt	4920
gccctttaag	gtgggtggtg	tctcagccat	cctggccctg	gtgggtgctc	ccatcatctc	4980

ccttatcatc ctcatcatgc tttggcagaa gaagccacgt tagtaacagg taagagtgta 5040  
 actttaaata atgccaatta tttaaagtta ctgactctct ctgcttacga cgctttctct 5100  
 tttttttttc ctgcaggggt agtaatcagc ctcgactgtg ccttctagtt gccagccatc 5160  
 tgttgtttgc ccctcccccg tgccttcctt gaccctggaa ggtgccactc ccaactgtcct 5220  
 ttccctaataa aatgaggaaa ttgcatcgca ttgtctgagt aggtgtcatt ctattctggg 5280  
 ggggtgggggtg gggcaggaca gcaaggggga ggattgggaa gacaatagca ggcattgctgg 5340  
 ggatggcccg ggcattgataa cttcgtataa tgtatgctat acgaagttat gtatac 5396

<210> 2

<211> 233

<212> PRT

<213> Artificial sequence

<220>

<223> BM40 Leader, Humanised D1.3 VL and Human C Kappa sequence

<400> 2

Met Arg Ala Trp Ile Phe Phe Leu Leu Cys Leu Ala Gly Arg Ala Leu  
 1 5 10 15

Ala Ala Ser Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala  
 20 25 30

Ser Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gly Asn Ile  
 35 40 45

His Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys  
 50 55 60

Leu Leu Ile Tyr Tyr Thr Thr Thr Leu Ala Asp Gly Val Pro Ser Arg  
 65 70 75 80

Phe Ser Gly Ser Gly Ser Gly Thr Asp Tyr Thr Phe Thr Ile Ser Ser  
 85 90 95

Leu Gln Pro Glu Asp Ile Ala Thr Tyr Tyr Cys Gln His Phe Trp Ser  
 100 105 110

Thr Pro Arg Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr  
 115 120 125

Ala Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu  
 130 135 140

Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr Pro  
 145 150 155 160

Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser Gly  
165 170 175

Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr Tyr  
180 185 190

Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys His  
195 200 205

Lys Leu Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro Val  
210 215 220

Thr Lys Ser Phe Asn Arg Gly Glu Cys  
225 230

<210> 3  
<211> 15  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Leader 1 sequence

<400> 3

Met Ser Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr  
1 5 10 15

<210> 4  
<211> 449  
<212> PRT  
<213> Artificial sequence

<220>  
<223> Leader 2 and D1.3 VH sequence, CH1 domain, Hinge, CH2 and CH3 domains

<400> 4

Gly Ala Met Ala Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val  
1 5 10 15

Arg Pro Ser Gln Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Ser Thr  
20 25 30

Phe Ser Gly Tyr Gly Val Asn Trp Val Arg Gln Pro Pro Gly Arg Gly  
35 40 45

Leu Glu Trp Ile Gly Met Ile Trp Gly Asp Gly Asn Thr Asp Tyr Asn  
50 55 60

Ser Ala Leu Lys Ser Arg Val Thr Met Leu Val Asp Thr Ser Lys Asn  
65 70 75 80

Gln Phe Ser Leu Arg Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val  
85 90 95

Tyr Tyr Cys Ala Arg Glu Arg Asp Tyr Arg Leu Asp Tyr Trp Gly Gln  
100 105 110

Gly Ser Leu Val Thr Val Ser Ser Ala Ser Thr Lys Gly Pro Ser Val  
115 120 125

Phe Pro Leu Ala Pro Cys Ser Arg Ser Thr Ser Glu Ser Thr Ala Ala  
130 135 140

Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val Thr Val Ser  
145 150 155 160

Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe Pro Ala Val  
165 170 175

Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val Thr Val Pro  
180 185 190

Ser Ser Asn Phe Gly Thr Gln Thr Tyr Thr Cys Asn Val Asp His Lys  
195 200 205

Pro Ser Asn Thr Lys Val Asp Lys Thr Val Glu Arg Lys Cys Cys Val  
210 215 220

Glu Cys Pro Pro Cys Pro Ala Pro Pro Val Ala Gly Pro Ser Val Phe  
225 230 235 240

Leu Phe Pro Pro Lys Pro Lys Asp Thr Leu Met Ile Ser Arg Thr Pro  
245 250 255

Glu Val Thr Cys Val Val Val Asp Val Ser His Glu Asp Pro Glu Val  
260 265 270

Gln Phe Asn Trp Tyr Val Asp Gly Val Glu Val His Asn Ala Lys Thr  
275 280 285

Lys Pro Arg Glu Glu Gln Phe Asn Ser Thr Phe Arg Val Val Ser Val  
290 295 300

Leu Thr Val Val His Gln Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys  
305 310 315 320

Ala Val Ser Asn Lys Gly Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser  
325 330 335



Lys Thr Lys Gly Gln Pro Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro  
340 345 350

Ser Arg Glu Glu Met Thr Lys Asn Gln Val Ser Leu Thr Cys Leu Val  
355 360 365

Lys Gly Phe Tyr Pro Ser Asp Ile Ala Val Glu Trp Glu Ser Asn Gly  
370 375 380

Gln Pro Glu Asn Asn Tyr Lys Thr Thr Pro Pro Met Leu Asp Ser Asp  
385 390 395 400

Gly Ser Phe Phe Leu Tyr Ser Lys Leu Thr Val Asp Lys Ser Arg Trp  
405 410 415

Gln Gln Gly Asn Val Phe Ser Cys Ser Val Met His Glu Ala Leu His  
420 425 430

Asn His Tyr Thr Gln Lys Ser Leu Ser Leu Ser Pro Gly Lys Gly Ser  
435 440 445

Lys

<210> 5  
<211> 60  
<212> PRT  
<213> Artificial

<220>  
<223> Myc epitope, PDGFR spacer and PDGFR Transmembrane region (TM)  
<400> 5

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Ala Val Gly Gln Asp  
1 5 10 15

Thr Gln Glu Val Ile Val Val Pro His Ser Leu Pro Phe Lys Val Val  
20 25 30

Val Ile Ser Ala Ile Leu Ala Leu Val Val Leu Thr Ile Ile Ser Leu  
35 40 45

Ile Ile Leu Ile Met Leu Trp Gln Lys Lys Pro Arg  
50 55 60

<210> 6  
<211> 1542  
<212> DNA  
<213> Artificial Sequence

<220>

<223> pD1-huD1.3 donor construct fragment containing AAVS1 left  
homology arm, splice acceptor site, Blastocidin gene, SV40 polyA

<400> 6

```
ggtaccgaat tcgccctttg ctttctctga cctgcattct ctcccctggg cctgtgcccgc      60
tttctgtctg cagcttgtgg cctgggtcac ctctacggct ggcccagatc cttccctgcc      120
gcctccttca ggttccgtct tcctccactc cctcttcccc ttgctctctg ctgtgttctg      180
gccaaggat gctctttccg gagcacttcc ttctcggcgc tgcaccacgt gatgtcctct      240
gagcggatcc tcccctgtgc tgggtcctct cggggcatct ctctccctc acccaacccc      300
atgccgtctt cactcgtctg gttccctttt ctttctcctt ctggggcctg tgccatctct      360
cgtttcttag gatggccttc tccgacggat gtctcccttg cgtcccgcct ccccttcttg      420
taggcctgca tcatcacctg ttttctggac aaccccaaag taccctctct cctggccttt      480
agccacctct ccctctctt gctttctttg cctggacacc cgtttctctt gtggattcgg      540
gtcacctctc actcctttca tttgggcagc tcccctaccc cccttacctc tctagtctgt      600
gcaagctctt ccagccccct gtcatggcat cttccagggg tccgagagct cagctagtct      660
tcttctcca acccggggcc ctatgtccac ttcaggacag catgtttgct gcctccaggg      720
atcctgtgtc cccgagctgg gaccacctta tattcccagg gccggttaat gtggctctgg      780
ttctgggtac ttttatctgt cccctccacc ccacagtggg gcaagatgca tcttctgacc      840
tcttctcttc ctcccacagg gcatggcaaa acctctgagc caggaagaaa gcacactgat      900
tgaaagagca accgctacta tcaacagcat ccccatctcc gaagactatt ctgtggctag      960
tgccgctctg tccagcgacg ggagaatctt caccgggtgtg aacgtctacc actttacagg     1020
cggaccatgc gcagagctgg tggctctggg gactgcagcc gctgcagccg ctggtaatct     1080
gacctgtatc gtggccattg gcaacgaaaa taggggcatc ctgtcccat gcggcaggtg     1140
tcggcaggtg ctgctggatc tgcactctgg catcaaggca attgtcaaag actctgatgg     1200
acagcctacc gccgtcggta tccgtgaact gctgcctagc ggctatgtct gggagggata     1260
atgagcttgg cttcgaaatg accgaccaag cgacgcccaa cctgccatca cgagatttctg     1320
attccaccgc cgccttctat gaaagggttg gcttcggaat cgttttccgg gacgcccggct     1380
ggatgatcct ccagcgcggg gatctcatgc tggagttctt cgcccacccc aacttgttta     1440
ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca aataaagcat     1500
ttttttcact gcattctagt tgtggtttaa ttaagtcaat tc                               1542
```

<210> 7

<211> 924

<212> DNA

<213> Artificial Sequence

<220>

<223> pD1-huD1.3 fragment containing AAVS1 right homology arm

<400> 7

tggcccgggc atgataactt cgtataatgt atgctatacg aagttatgta tacggcgcgc 60  
ccactagggg caggattggg gacagaaaag ccccatcctt aggcctcctc cttcctagtc 120  
tcctgatatt gggctcaacc cccacctcct gttaggcaga ttccttatct ggtgacacac 180  
ccccatttcc tggagccatc tctctccttg ccagaacctc taaggtttgc ttacgatgga 240  
gccagagagg atcctgggag ggagagcttg gcaggggggtg ggaggggaagg gggggatgcg 300  
tgacctgccg ggttctcagt ggccaccctg cgctaccctc tcccagaacc tgagctgctc 360  
tgacgcggtt gtctgggtgcg tttcactgat cctgggtgctg cagcttcctt acacttccca 420  
agaggagaag cagtttggaa aaacaaaatc agaataagtt ggtcctgagt tctaactttg 480  
gctcttcacc tttctagtcc ccaatttata ttgttctctc gtgcgtcagt tttacctgtg 540  
agataaggcc agtagccagc cccgtcctgg cagggctgtg gtgaggaggg ggggtgtccgt 600  
gtggaaaact ccctttgtga gaatggtgcg tcctagggtg tcaccaggtc gtggccgcct 660  
ctactccctt tctctttctc catccttctt tccttaaaga gtccccagtg ctatctggga 720  
catattctc cgcccagagc agggctccgc ttccctaagg cctgctctg ggcttctggg 780  
tttgagtctt tggcaagccc aggagagggc ctcaggcttc cctgtcccc ttctctgtcc 840  
accatctcat gccctgggt ctctgcccc ttccctacag gggttcctgg ctctgctctg 900  
acgcgtgat actcgatctt tccg 924

<210> 8

<211> 132

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic Blastocidin resistance gene

<400> 8

Met Ala Lys Pro Leu Ser Gln Glu Glu Ser Thr Leu Ile Glu Arg Ala  
1 5 10 15

Thr Ala Thr Ile Asn Ser Ile Pro Ile Ser Glu Asp Tyr Ser Val Ala  
20 25 30

Ser Ala Ala Leu Ser Ser Asp Gly Arg Ile Phe Thr Gly Val Asn Val  
35 40 45

Tyr His Phe Thr Gly Gly Pro Cys Ala Glu Leu Val Val Leu Gly Thr  
50 55 60

Ala Ala Ala Ala Ala Ala Gly Asn Leu Thr Cys Ile Val Ala Ile Gly

65

70

75

80

Asn Glu Asn Arg Gly Ile Leu Ser Pro Cys Gly Arg Cys Arg Gln Val  
85 90 95

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

Gly Gln Pro Thr Ala Val Gly Ile Arg Glu Leu Leu Pro Ser Gly Tyr  
115 120 125

Val Trp Glu Gly  
130

<210> 9  
<211> 734  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> pD6 vector for scFv-Fc expression, fragment from NcoI site to PmeI site containing a fragment (last 4 amino acids) of the leader sequence, the complete CH2 domain and CH3 domain

<400> 9  
ggcgccatgg cccaggtcgc ggccgccgtg gaatgcccc cctgtcccgc tcctccagtg 60  
gctggacctt ccgtgttctt gttcccccca aagcccaagg acaccctgat gatcagccgg 120  
acccccgaag tgacctgcgt ggtggtggac gtgtcccacg aggaccccga ggtgcagttc 180  
aattggtacg tggacggcgt ggaagtgcac aacgccaaga ccaagcccag agaggaacag 240  
ttcaacagca ctttccgggt ggtgtccgtg ctgaccgtgg tgcaccagga ctggctgaac 300  
ggcaaagagt acaagtgcgc cgtctccaac aagggcctgc ctgcccccat cgagaaaacc 360  
atcagcaaga ccaagggcca gcctcgcgag cctcaggtgt acacactgcc ccccagccgg 420  
gaagagatga ccaagaacca ggtgtccctg acctgcctcg tgaagggcct ctaccccage 480  
gatatcgccg tggaatggga gagcaacggc cagcccgaga acaactacaa gaccaccccc 540  
cccatgctgg acagcgacgg ctattcttc ctgtacagca agctgacagt ggacaagagc 600  
cgggtggcagc agggcaacgt gttcagctgc agcgtgatgc acgaggcct gcacaaccac 660  
tacaccaga agtccttgag cctgagcccc ggcaagggat ccaaggtaag tttaaacata 720  
tatataactt taaa 734

<210> 10  
<211> 4  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Vector pD6, fragment at the end of leader sequence

<400> 10

Gly Ala Met Ala  
1

<210> 11

<211> 219

<212> PRT

<213> Artificial Sequence

<220>

<223> Vector pD6 CH2 and CH3 domains

<400> 11

Ala Pro Pro Val Ala Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro  
1 5 10 15

Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val  
20 25 30

Val Asp Val Ser His Glu Asp Pro Glu Val Gln Phe Asn Trp Tyr Val  
35 40 45

Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln  
50 55 60

Phe Asn Ser Thr Phe Arg Val Val Ser Val Leu Thr Val Val His Gln  
65 70 75 80

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Ala Val Ser Asn Lys Gly  
85 90 95

Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Thr Lys Gly Gln Pro  
100 105 110

Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr  
115 120 125

Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser  
130 135 140

Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr  
145 150 155 160

Lys Thr Thr Pro Pro Met Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr  
165 170 175

Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe  
180 185 190

Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys  
195 200 205

Ser Leu Ser Leu Ser Pro Gly Lys Gly Ser Lys  
210 215

<210> 12  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Fragment of sequence of pD4 vector preceding and including the start of the AAVS1 left homology arm

<400> 12  
ggtaccgaat tcgccctttg ctttctctga cctgcattct ctcccctggg cctgtgcccg 60

<210> 13  
<211> 3933  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Fragment of sequence of vector pD4 including the end of the AAVS1 LHA, FRT site, Lox2272 site, 1-SceI meganuclease site, GFP left TALEN binding site, GFP right TALEN binding site, T2A, GFP, PGK promoter, puromycin delta TK, loxP site and start of AAVS1 RHA

<400> 13  
ttctgggtac ttttatctgt cccctccacc ccacagtggg gcaagatgca tgaagttcct 60  
attccgaagt tcctattctc tagaaagtat aggaacttcg accataactt cgtataaagt 120  
atcctatacg aagttatgcg atcgctcgcg cgtagggata acagggtaat aagtccaccg 180  
gtcgccacca tggtgagcaa gggcgaggag ctgttcactt ctgacctctt ctcttctctc 240  
cacagggcct agagagatct ggcagcggag agggcagagg aagtcttcta acatgcgggtg 300  
acgtggagga gaatcccga cccgatggtga gcaagggaga agaactcttc accgggggtg 360  
tgcccatcct ggtcgagctg gacggcgacg tgaacggcca caagttcagc gtgtccggcg 420  
agggcgaggg cgatgccacc tacggcaagc tgaccctgaa gttcatctgc accaccggca 480  
agctgcccgt gccctggccc accctcgtga ccaccctgac ctacggcgtg cagtgttca 540  
gccgtacc caccacatg aagcagcagc acttcttcaa gtccgcatg cccgaaggct 600  
acgtccagga gcgcaccatc ttcttcaagg acgacggcaa ctacaagacc cgcgcccagg 660  
tgaagttcga gggcgacacc ctggtgaacc gcatcgagct gaagggcatc gacttcaagg 720  
aggacggcaa catcctgggg cacaagctgg agtacaacta caacagccac aacgtctata 780  
tcatggccga caagcagaag aacggcatca aggtgaactt caagatccgc cacaacatcg 840  
aggacggcag cgtgcagctc gccgaccact accagcagaa ccccccatc ggcgacggcc 900

ccgtgctgct	gcccgacaac	cactacctga	gcacccagtc	cgccctgagc	aaagacocca	960
acgagaagcg	cgatcacatg	gtcctgctgg	agttcgtgac	cgccgcccgg	atcaactcacg	1020
gcatggacga	gctgtacaag	aagctgagcc	acggcttccc	gccggcgggtg	gcggcgcagg	1080
atgatggcac	gctgcccacg	tcttgctgcc	aggagagcgg	gatggaccgt	caccctgcag	1140
cctgtgcttc	tgctaggatc	aatgtgtagg	tgatcagcct	cgactgtgcc	ttctagttgc	1200
cagccatctg	ttgtttgcc	ctccccctg	ccttccttga	ccctggaagg	tgccaactccc	1260
actgtccttt	cctaataaaa	tgaggaaatt	gcatcgcatt	gtctgagtag	gtgtcattct	1320
attctggggg	gtgggggtgg	gcaggacagc	aagggggagg	attgggaaga	caatagcagg	1380
catgctgggg	atggcccacg	tctaccgggt	aggggagggc	cttttcccaa	ggcagctctgg	1440
agcatgcgct	ttagcagccc	cgctgggcac	ttggcgctac	acaagtggcc	tctggcctcg	1500
cacacattcc	acatccaccg	gtagggccca	accggctccg	ttctttgggtg	gccccttcgc	1560
gccaccttct	actcctcccc	tagtcaggaa	gttccccccc	gccccgcagc	tcgcgtcgtg	1620
caggacgtga	caaatggaag	tagcacgtct	cactagtctc	gtgcagatgg	acagcaccgc	1680
tgagcaatgg	aagcgggtag	gcctttgggg	cagcggccaa	tagcagcttt	gctccttcgc	1740
tttctggggc	cagaggctgg	gaaggggtgg	gtccgggggc	gggctcaggg	gcgggctcag	1800
gggccccggc	ggcggccgaa	ggtcctccgg	aggccccgca	ttctgcacgc	ttcaaaagcg	1860
cacgtctgcc	gcgctgttct	cctcttctct	atctccgggc	ctttcgacct	gcagccaacg	1920
ccacatggg	gaccgagtac	aagcccacgg	tgcgccctgc	caccgcgcac	gacgtcccc	1980
gggcccgtac	caccctcgcc	gccgcgttcg	ccgactacce	cgccacgcgc	cacaccgtcg	2040
acccggaccg	ccacatcgag	cgggtcaccg	agctgcaaga	actcttctct	acgcgcgtcg	2100
ggctcgacat	cggcaaggtg	tgggtcgcgg	acgacggcgc	cgcggtggcg	gtctggacca	2160
cgccggagag	cgtcgaagcg	ggggcgggtg	tcgccgagat	cggcccgcgc	atggccgagt	2220
tgagcggttc	ccggctggcc	gcgcagcaac	agatggaagg	cctcctggcg	ccgcaccggc	2280
ccaaggagcc	cgctgtggtc	ctggccaccg	tcggcgtctc	gcccgaccac	cagggcaagg	2340
gtctgggcag	cgccgtcgtg	ctccccggag	tggaggcggc	cgagcgcgcc	ggggtgcccg	2400
ccttctgga	gacctccgcg	ccccgcaacc	tccccttcta	cgagcggctc	ggcttcaccg	2460
tcaccgccga	cgctgaggtg	cccgaaggac	cgcgcacctg	gtgcatgacc	cgcaagcccc	2520
gtgccggatc	catgcccacg	ctactgcggg	tttatataga	cggctctcac	gggatgggga	2580
aaaccaccac	cacgcaactg	ctggtggccc	tgggttcgcg	cgacgatatc	gtctacgtac	2640
ccgagccgat	gacttactgg	caggtgctgg	gggcttccga	gacaatcgcg	aacatctaca	2700
ccacacaaca	ccgcctcgac	cagggtgaga	tatcggccgg	ggacgcggcg	gtggtaatga	2760
caagcgcca	gataacaatg	ggcatgcctt	atgccgtgac	cgacgcggtt	ctggctcctc	2820

atatcggggg ggaggctggg agctcacatg ccccgccccc ggccctcacc ctcatcttcg 2880  
 accgccatcc catcgccgcc ctctgtgct acccgccgc gcgatacctt atgggcagca 2940  
 tgacccccca ggccgtgctg gcgttcgtgg ccctcatccc gccgacctg cccggcacia 3000  
 acatcgtgtt gggggccctt ccggaggaca gacacatcga ccgcctggcc aaacgccagc 3060  
 gccccggcga gcggcttgac ctggctatgc tggccgcgat tcgccgcggt tacgggctgc 3120  
 ttgccaatac ggtgcgggat ctgcagggcg gcgggtcgtg gcgggaggat tggggacagc 3180  
 tttcggggac ggccgtgccg ccccagggcg ccgagcccca gagcaacgcg ggccccagac 3240  
 cccatatcgg ggacacgtta tttaccctgt ttcgggcccc cgagttgctg gcccccaacg 3300  
 gcgacctgta caacgtgttt gcctgggcct tggacgtctt ggccaaacgc ctccgtccca 3360  
 tgcacgtctt tctctggat tacgaccaat cggccgcgg ctgccgggac gccctgctgc 3420  
 aacttacctc cgggatggtc cagaccacg tcaccacccc cggctccata ccgacgatct 3480  
 gcgacctggc gcgcacgttt gcccgggaga tgggggaggc taactgagct ctagagctcg 3540  
 ctgatcagcc tcgactgtgc cttctagttg ccagccatct gttgtttgcc cctccccctg 3600  
 gccttccttg accctggaag gtgccactcc cactgtcctt tcctaataaa atgaggaaat 3660  
 tgcacgcat tgtctgagta ggtgtcattc tattctgggg ggtgggggtg ggccaggacag 3720  
 caagggggag gattgggaag acaatagcag gcatgctggg gatgcgggtg gctctatggc 3780  
 ttctgaggcg gaaagaacca gctggggctc gagatccact agttctagcc tcgaggctag 3840  
 agcggccggc cctataactt cgtataatgt atgctatacg aagttatcag gtaagttaac 3900  
 agggcgcgcc cactagggac aggattggtg aca 3933

<210> 14  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Fragment of pD4 including the end of the AAVS1 right homology arm (RHA)

<400> 14  
 tctgctctga cgcgtgtata 20

<210> 15  
 <211> 53  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Fragment of pD4 comprising T2A and start of EGFP sequence

<400> 15

Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu  
 1 5 10 15



Glu Asn Pro Gly Pro Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly  
20 25 30

Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys  
35 40 45

Phe Ser Val Ser Gly  
50

<210> 16

<211> 69

<212> PRT

<213> Artificial Sequence

<220>

<223> Fragment of pD4 comprising terminal part of EGFP sequence and Mu  
ornithine decarboxylase PEST sequence

<400> 16

Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val Thr Ala Ala  
1 5 10 15

Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys Lys Leu Ser His Gly  
20 25 30

Phe Pro Pro Ala Val Ala Ala Gln Asp Asp Gly Thr Leu Pro Met Ser  
35 40 45

Cys Ala Gln Glu Ser Gly Met Asp Arg His Pro Ala Ala Cys Ala Ser  
50 55 60

Ala Arg Ile Asn Val  
65

<210> 17

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> Initial fragment of of puroldelta TK sequence encoded by pD4

<400> 17

Met Gly Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp  
1 5 10 15

Val Pro

<210> 18

<211> 15  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Terminal fragment of puromycin delta TK sequence encoded by pD4

<400> 18

Cys Asp Leu Ala Arg Thr Phe Ala Arg Glu Met Gly Glu Ala Asn  
1                  5                  10                  15

<210> 19  
<211> 1098  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Sequence of donor plasmid for integration into Flp/GFP TALEN sites (pD5), fragment preceding D1.3 antibody expression cassette

<400> 19  
ggtaccgaat tcctagctct tccagcccc tgatcatggca tcttccaggg gtccgagagc 60  
tcagctagtc ttcttctctc aaccgaggcc cctatgtcca cttcaggaca gcatgttttc 120  
tgcctccagg gatcctgtgt ccccgagctg ggaccacctt atattcccag ggccgggtaa 180  
tgtggctctg gttctgggta cttttatctg tcccctccac cccacagtgg ggcaagatgc 240  
atgaagttcc tattccgaag ttcttattct ctagaaagta taggaacttc gaccataact 300  
tcgtataaag tatkctatac gaagtattgc gatcgctcgc gcgtagggat aacagggtaa 360  
taagtccacc ggtcgccacc atggctttct gacctttct cttcctcca cagggcatgg 420  
caaacctct gagccaggaa gaaagcacac tgattgaaag agcaaccgct actatcaaca 480  
gcatcccat ctccgaagac tattctgtgg ctagtgccgc tctgtccagc gacgggagaa 540  
tcttcaccgg tgtgaacgtc taccacttta caggcggacc atgacagag ctggtggtcc 600  
tggggactgc agccgctgca gccgctggta atctgacctg tatcgtggcc attggcaacg 660  
aaaatagggg catcctgtcc ccatgaggca ggtgtcggca ggtgctgctg gatctgcatc 720  
ctggcatcaa ggcaattgtc aaagactctg atggacagcc taccgcccgc ggtatccgtg 780  
aactgctgcc tagcggctat gtctgggagg gataatgaat gagcttggct tcgaaatgac 840  
cgaccaagcg acgccaacc tgccatcacg agatttcgat tccaccgccg cttctatga 900  
aaggttggc ttcggaatcg ttttccggga cggcgctgg atgacccccc agcgggggga 960  
tctcatgctg gagttcttcg cccaccccaa cttgtttatt gcagcttata atggttacia 1020  
ataaagcaat agcatcacia atttcaciaa taaagcattt ttttactgc attctagttg 1080  
tggtttaatt aacaattc 1098

<210> 20  
<211> 931

<212> DNA  
<213> Artificial Sequence

<220>  
<223> Sequence of donor plasmid for integration into Flp/GFP TALEN sites (pD5), fragment following D1.3 antibody expression cassette

<400> 20  
caggcatgct ggggatggcc cgggcatgat aacttcgtat aatgtatgct atacgaagtt 60  
atgtatacgg cgcgcccagag caagggcgag gagctgttca cttctgacct cttctcttcc 120  
tcccacctga gcctagagag atctggcagc ggagagggca gaggaagtct tctaacaatgc 180  
ggtgacgtgg aggagaatcc cggaccgtga gtgagcaagg gagaagaact cttcaccggg 240  
gtggtgcca tcctggtcga gctggacggc gacgtgaacg gccacaagtt cagcgtgtcc 300  
ggcgagggcg agggcgatgc cacctacggc aagctgacct tgaagttcat ctgcaccacc 360  
ggcaagctgc ccgtgccctg gccaccctc gtgaccacct tgacctacgg cgtgcagtgc 420  
ttcagccgct accccgacca catgaagcag cagcattct tcaagtccgc catgcccgaa 480  
ggctacgtcc aggagcgcac catcttcttc aaggacgacg gcaactaaa gaccgcgccc 540  
gaggtgaagt tcgagggcga caccctggtg aaccgcatcg agctgaaggg catcgacttc 600  
aaggaggacg gcaacatcct ggggcacaag ctggagtaca actacaacag ccacaacgtc 660  
tatatcatgg ccgacaagca gaagaacggc atcaaggtga acttcaagat ccgccacaac 720  
atcgaggacg gcagcgtgca gctcgccgac cactaccagc agaacacccc catcgggcagc 780  
ggccccgtgc tgctgcccga caaccactac ctgagcacct agtccgccct gagcaaagac 840  
cccaacgaga agcgcgatca catggtcctg ctggagttcg tgaccgcccgc cgggatcact 900  
cacggcatgg acgagcctga cgcgtgtata c 931

<210> 21  
<211> 131  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Sequence of Blasticidin resistance protein encoded by donor plasmid for integration into FLP/GFP TALEN sites (pD5)

<400> 21  
Ala Lys Pro Leu Ser Gln Glu Glu Ser Thr Leu Ile Glu Arg Ala Thr  
1 5 10 15  
Ala Thr Ile Asn Ser Ile Pro Ile Ser Glu Asp Tyr Ser Val Ala Ser  
20 25 30  
Ala Ala Leu Ser Ser Asp Gly Arg Ile Phe Thr Gly Val Asn Val Tyr  
35 40 45

His Phe Thr Gly Gly Pro Cys Ala Glu Leu Val Val Leu Gly Thr Ala  
50 55 60

Ala Ala Ala Ala Ala Gly Asn Leu Thr Cys Ile Val Ala Ile Gly Asn  
65 70 75 80

Glu Asn Arg Gly Ile Leu Ser Pro Cys Gly Arg Cys Arg Gln Val Leu  
85 90 95

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

Gln Pro Thr Ala Val Gly Ile Arg Glu Leu Leu Pro Ser Gly Tyr Val  
115 120 125

Trp Glu Gly  
130

<210> 22  
<211> 813  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Sequence of I-Sce-1 meganuclease construct

<400> 22  
ccatgggcta tccttacgat gtcacctgatt acgccaacag tcctgggtatc cctgggtatgg 60  
gtcctaaaaa gaagcgaaaa gtgggtagac tggaaccgag catgaagaac attaagaaaa 120  
atcaggtgat gaacctggga cctaattcca agctgctgaa agagtacaag tctcagctga 180  
tcgaactgaa cattgagcag tttgaagcag ggatcgggtct gattctgggg gacgcctaca 240  
tccggagcag ggatgagggc aagacttatt gcatgcagtt cgaatggaag aataaggcct 300  
acatggacca cgtgtgtctg ctgtatgatc agtgggtcct gtctccccct cacaagaaa 360  
agagagtgaa ccatctgggc aatctgggtca ttacttgggg agcacagacc ttcaagcatc 420  
aggcctttaa caaactggct aacctgttca tcgtgaacaa caagaaaacc atccctaaca 480  
atctggtcga aaactacctg acaccaatga gtctggccta ttggttcatg gacgatggcg 540  
gaaaatggga ctacaacaag aacagcacia acaaaagcat cgtgctgaat acccagtcct 600  
tcacatttga ggaagtggag tatctgggtca agggcctgag gaacaaattc cagctgaact 660  
gctacgtgaa gatcaacaag aacaagccaa tcatctacat cgattctatg agttacctga 720  
tcttttataa cctgattaag ccataacctga tccccagat gatgtataaa ctgcccata 780  
caatcagctc cgagactttc ctgaaggtct aga 813

<210> 23

<211> 269  
<212> PRT  
<213> Artificial Sequence

<220>

<223> Sequence encoded by I-SceI meganuclease construct: HA tag, NLS,  
and I-SceI

<400> 23

Met Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Asn Ser Pro Gly Ile  
1 5 10 15

Pro Gly Met Gly Pro Lys Lys Lys Arg Lys Val Gly Arg Leu Glu Pro  
20 25 30

Gly Met Lys Asn Ile Lys Lys Asn Gln Val Met Asn Leu Gly Pro Asn  
35 40 45

Ser Lys Leu Leu Lys Glu Tyr Lys Ser Gln Leu Ile Glu Leu Asn Ile  
50 55 60

Glu Gln Phe Glu Ala Gly Ile Gly Leu Ile Leu Gly Asp Ala Tyr Ile  
65 70 75 80

Arg Ser Arg Asp Glu Gly Lys Thr Tyr Cys Met Gln Phe Glu Trp Lys  
85 90 95

Asn Lys Ala Tyr Met Asp His Val Cys Leu Leu Tyr Asp Gln Trp Val  
100 105 110

Leu Ser Pro Pro His Lys Lys Glu Arg Val Asn His Leu Gly Asn Leu  
115 120 125

Val Ile Thr Trp Gly Ala Gln Thr Phe Lys His Gln Ala Phe Asn Lys  
130 135 140

Leu Ala Asn Leu Phe Ile Val Asn Asn Lys Lys Thr Ile Pro Asn Asn  
145 150 155 160

Leu Val Glu Asn Tyr Leu Thr Pro Met Ser Leu Ala Tyr Trp Phe Met  
165 170 175

Asp Asp Gly Gly Lys Trp Asp Tyr Asn Lys Asn Ser Thr Asn Lys Ser  
180 185 190

Ile Val Leu Asn Thr Gln Ser Phe Thr Phe Glu Glu Val Glu Tyr Leu  
195 200 205

Val Lys Gly Leu Arg Asn Lys Phe Gln Leu Asn Cys Tyr Val Lys Ile  
210 215 220

Asn Lys Asn Lys Pro Ile Ile Tyr Ile Asp Ser Met Ser Tyr Leu Ile  
225 230 235 240

Phe Tyr Asn Leu Ile Lys Pro Tyr Leu Ile Pro Gln Met Met Tyr Lys  
245 250 255

Leu Pro Asn Thr Ile Ser Ser Glu Thr Phe Leu Lys Val  
260 265

<210> 24  
<211> 67  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> pINT20 Vector fragment showing NsiI site, splice acceptor site  
and start of Puromycin resistance gene sequence

<400> 24  
atgcatcttc tgacctcttc tcttcctccc acagggcatg accgagtaca agcccacggt 60  
gcgccctc 67

<210> 25  
<211> 10  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Initial fragment of puromycin resistance protein encoded by  
pINT20 vector

<400> 25

Met Thr Glu Tyr Lys Pro Thr Val Arg Leu  
1 5 10

<210> 26  
<211> 1125  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Sequence of T cell receptor clone c12/c2 alpha chain construct  
containing NheI, NotI and Acc651 restriction sites

<400> 26  
gctagcaagc aggaagtgc tcagatccca gccgctctga gcgctgctga gggagaaaac 60  
ctggctctga attgcagttt caccgactca gccatctata acctgcagtg gtttcgccag 120  
gatccaggca agggactgac ctccctgctg ctgattcaga gctcccagag ggaacagaca 180  
tctggcagac tgaatgctag tctggacaaa tctagtggac ggtctaccct gtacatogca 240  
gccagccagc ctggagattc cgcaacatat ctgtgcgccc tgcgcccact tacaggcgga 300

agctacattc ccaccttcgg gcgaggtaca agcctgatcg tgcacccaga catccagaat 360  
 ccggagcccc ccgtatacca gctgaaggac cccagaagcc aggacagcac cctgtgacctg 420  
 ttcaccgact tcgacagcca gatcaacgtg cccaagacaa tggaaagcgg caccttcac 480  
 accgacaaga ccgtgctgga catgaaggct atggacagca agagcaacgg cgccattgcc 540  
 tggccaacc agaccagctt cacatgccag gacatcttca aagagacaaa cgccacctac 600  
 cccagcagcg acgtgccctg tgatgccacc ctgaccgaga agtccttoga gacagacatg 660  
 aacctgaact tccagaacct gtccgcggcc gcaggcctgc tggatcccaa gctgtgctac 720  
 ctgctggagc ggatcctggt catctacggg gtgatcctga ctgccctggt cctgagcagtc 780  
 aaattttctc ggagtgccga cgctcctgca taccagcagg ggcagaacca gctgtataac 840  
 gagctgaatc tgggtcggag ggaggaatat gacgtgctgg ataagagacg cggcagggat 900  
 ccagaaatgg ggggcaagcc ccagcgacgg aaaaaccctc aggagggact gtataatgaa 960  
 ctgcagaagg acaaaatggc cgaggcttac tctgaaattg ggatgaaggg cgagaggaga 1020  
 cgcgcaaaag gacacgatgg cctgtaccag ggactgagca ctgctaccaa ggacacatat 1080  
 gatgctctgc atatgcaggc actgccccct agataataag gtacc 1125

<210> 27  
 <211> 369  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Sequence of T cell receptor clone c12/c2 alpha chain

<400> 27

Lys Gln Glu Val Thr Gln Ile Pro Ala Ala Leu Ser Val Pro Glu Gly  
 1 5 10 15

Glu Asn Leu Val Leu Asn Cys Ser Phe Thr Asp Ser Ala Ile Tyr Asn  
 20 25 30

Leu Gln Trp Phe Arg Gln Asp Pro Gly Lys Gly Leu Thr Ser Leu Leu  
 35 40 45

Leu Ile Gln Ser Ser Gln Arg Glu Gln Thr Ser Gly Arg Leu Asn Ala  
 50 55 60

Ser Leu Asp Lys Ser Ser Gly Arg Ser Thr Leu Tyr Ile Ala Ala Ser  
 65 70 75 80

Gln Pro Gly Asp Ser Ala Thr Tyr Leu Cys Ala Val Arg Pro Leu Thr  
 85 90 95

Gly Gly Ser Tyr Ile Pro Thr Phe Gly Arg Gly Thr Ser Leu Ile Val

100

105

110

His Pro Asp Ile Gln Asn Pro Glu Pro Ala Val Tyr Gln Leu Lys Asp  
 115 120 125

Pro Arg Ser Gln Asp Ser Thr Leu Cys Leu Phe Thr Asp Phe Asp Ser  
 130 135 140

Gln Ile Asn Val Pro Lys Thr Met Glu Ser Gly Thr Phe Ile Thr Asp  
 145 150 155 160

Lys Thr Val Leu Asp Met Lys Ala Met Asp Ser Lys Ser Asn Gly Ala  
 165 170 175

Ile Ala Trp Ser Asn Gln Thr Ser Phe Thr Cys Gln Asp Ile Phe Lys  
 180 185 190

Glu Thr Asn Ala Thr Tyr Pro Ser Ser Asp Val Pro Cys Asp Ala Thr  
 195 200 205

Leu Thr Glu Lys Ser Phe Glu Thr Asp Met Asn Leu Asn Phe Gln Asn  
 210 215 220

Leu Ser Ala Ala Ala Gly Leu Leu Asp Pro Lys Leu Cys Tyr Leu Leu  
 225 230 235 240

Asp Gly Ile Leu Phe Ile Tyr Gly Val Ile Leu Thr Ala Leu Phe Leu  
 245 250 255

Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Gln Gln Gly  
 260 265 270

Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr  
 275 280 285

Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys  
 290 295 300

Pro Gln Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln  
 305 310 315 320

Lys Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu  
 325 330 335

Arg Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr  
 340 345 350

Ala Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro



Arg

<210> 28  
 <211> 1226  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Sequence of T cell receptor clone c12/c2 beta chain construct  
 containing Nco1, Xho1 and Hind3 restriction sites

<400> 28  
 ccatggccaa cgctggagtg actcagaccc ctaagttcca ggtcctgaaa actgggcaga 60  
 gtatgaccct gcagtgcgca caggacatga atcacgagta catgtcatgg tatcggcagg 120  
 atccagggat gggctctgagg ctgatccatt acagcgtggg cgctggaact accgaccagg 180  
 gcgaggtgcc caacggatat aatgtctcaa gaagcaccac agaagatttc ccaactgcgac 240  
 tgctgagcgc cgctcctagc cagacatccg tgtacttttg cgccagctcc aatgtcggga 300  
 acaccggcga gctgtttcttt ggggaagggt cccgcctgac agtgctcgag gacctgagaa 360  
 acgtgacccc cccaagggtg tccctgttcg agcctagcaa ggccgagatc gccaacaagc 420  
 agaaagccac cctcgtgtgc ctggccagag gcttcttccc cgaccacgtg gaactgtctt 480  
 ggtgggtcaa cggcaaagag gtgcacagcg gcgtgtccac cgatccccag gcctacaaag 540  
 agagcaacta cagctactgc ctgagcagca gactgcgggt gtccgccacc ttctggcaca 600  
 acccccggaa ccacttcaga tgccagggtgc agtttcacgg cctgagcgaa gaggacaagt 660  
 ggcccgaggg cagccctaag cccgtgacct agaatatctc tgccgaagcc tggggcagag 720  
 ccgactgtgg cattaccagc gccagctacc agcagggcgt gctgtctgcc accatcctgt 780  
 acgaggtcgc gagcggactg ctggacccaa agctgtgcta cctgctggat gggatcctgt 840  
 tcatctacgg tgtgattctg acagccctgt tctgctgagt caagttcagc cggagcgccg 900  
 acgcaccagc ataccagcag gggcagaatc agctgtataa cgagctgaat ctgggtcgga 960  
 gggaggaata cgacgtgctg gataagagac gcggcagggg tcccgaaatg ggcggaaagc 1020  
 ctcagcgacg gaaaaaccca caggagggac tgtacaatga actgcagaag gacaaaatgg 1080  
 ctgaggcata ttctgaaatc ggcatgaagg gagagaggag acgcggcaaa ggacacgatg 1140  
 ggctgtacca gggctctgagt acagccacta aggacaccta tgatgccttg catatgcagg 1200  
 ctctgccacc cagataataa aagctt 1226

<210> 29  
 <211> 404  
 <212> PRT  
 <213> Artificial Sequence

<220>

<223> Sequence of T cell receptor clone c12/c2 beta chain construct

<400> 29

Met Ala Asn Ala Gly Val Thr Gln Thr Pro Lys Phe Gln Val Leu Lys  
1 5 10 15

Thr Gly Gln Ser Met Thr Leu Gln Cys Ala Gln Asp Met Asn His Glu  
20 25 30

Tyr Met Ser Trp Tyr Arg Gln Asp Pro Gly Met Gly Leu Arg Leu Ile  
35 40 45

His Tyr Ser Val Gly Ala Gly Thr Thr Asp Gln Gly Glu Val Pro Asn  
50 55 60

Gly Tyr Asn Val Ser Arg Ser Thr Thr Glu Asp Phe Pro Leu Arg Leu  
65 70 75 80

Leu Ser Ala Ala Pro Ser Gln Thr Ser Val Tyr Phe Cys Ala Ser Ser  
85 90 95

Asn Val Gly Asn Thr Gly Glu Leu Phe Phe Gly Glu Gly Ser Arg Leu  
100 105 110

Thr Val Leu Glu Asp Leu Arg Asn Val Thr Pro Pro Lys Val Ser Leu  
115 120 125

Phe Glu Pro Ser Lys Ala Glu Ile Ala Asn Lys Gln Lys Ala Thr Leu  
130 135 140

Val Cys Leu Ala Arg Gly Phe Phe Pro Asp His Val Glu Leu Ser Trp  
145 150 155 160

Trp Val Asn Gly Lys Glu Val His Ser Gly Val Ser Thr Asp Pro Gln  
165 170 175

Ala Tyr Lys Glu Ser Asn Tyr Ser Tyr Cys Leu Ser Ser Arg Leu Arg  
180 185 190

Val Ser Ala Thr Phe Trp His Asn Pro Arg Asn His Phe Arg Cys Gln  
195 200 205

Val Gln Phe His Gly Leu Ser Glu Glu Asp Lys Trp Pro Glu Gly Ser  
210 215 220

Pro Lys Pro Val Thr Gln Asn Ile Ser Ala Glu Ala Trp Gly Arg Ala  
225 230 235 240

Asp Cys Gly Ile Thr Ser Ala Ser Tyr Gln Gln Gly Val Leu Ser Ala  
245 250 255

Thr Ile Leu Tyr Glu Val Ala Ser Gly Leu Leu Asp Pro Lys Leu Cys  
260 265 270

Tyr Leu Leu Asp Gly Ile Leu Phe Ile Tyr Gly Val Ile Leu Thr Ala  
275 280 285

Leu Phe Leu Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr  
290 295 300

Gln Gln Gly Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg  
305 310 315 320

Glu Glu Tyr Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met  
325 330 335

Gly Gly Lys Pro Gln Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn  
340 345 350

Glu Leu Gln Lys Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met  
355 360 365

Lys Gly Glu Arg Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly  
370 375 380

Leu Ser Thr Ala Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala  
385 390 395 400

Leu Pro Pro Arg

<210> 30

<211> 677

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of T cell receptor clone 4JFH alpha chain construct with  
initial NheI and terminal NotI restriction sites

<400> 30

gctagccaga aagaggtgga acagaacagc ggcctctga gcgtgccaga aggcgctatc 60

gccagcctga actgcaccta cagctttctg ggcagccaga gcttcttctg gtacagacag 120

tacagcggca agagccccga gctgatcatg ttcacctaca gagagggcga caaagaggac 180

ggcagattca ccgccagct gaacaaggcc agccagcacg tgtccctgct gatcagagac 240

```

agccagccta gcgacagcgc cacctacctg tgcgccgtga atgatggcgg cagactgacc      300
tttggcgacg gcaccacctt gaccgtgaag cctgacatcc agaatccgga gcccgccgta      360
taccagctga aggaccccag aagccaggac agcaccctgt gcctgttcac cgacttgcac      420
agccagatca acgtgcccga gacaatggaa agcggcacct tcatcaccga caagaccgtg      480
ctggacatga aggctatgga cagcaagagc aacggcgcca ttgcctggtc caaccagacc      540
agcttcacat gccaggacat cttcaaagag acaaacgcca cctaccccag cagcgacgtg      600
ccctgtgatg ccaccctgac cgagaagtcc ttcgagacag acatgaacct gaacttcag      660
aacctgtccg cggccgc                                                    677

```

```

<210> 31
<211> 221
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Sequence of T cell receptor clone 4JFH alpha chain construct

```

```

<400> 31

```

```

Gln Lys Glu Val Glu Gln Asn Ser Gly Pro Leu Ser Val Pro Glu Gly
1           5           10           15

```

```

Ala Ile Ala Ser Leu Asn Cys Thr Tyr Ser Phe Leu Gly Ser Gln Ser
          20           25           30

```

```

Phe Phe Trp Tyr Arg Gln Tyr Ser Gly Lys Ser Pro Glu Leu Ile Met
          35           40           45

```

```

Phe Thr Tyr Arg Glu Gly Asp Lys Glu Asp Gly Arg Phe Thr Ala Gln
          50           55           60

```

```

Leu Asn Lys Ala Ser Gln His Val Ser Leu Leu Ile Arg Asp Ser Gln
65           70           75           80

```

```

Pro Ser Asp Ser Ala Thr Tyr Leu Cys Ala Val Asn Asp Gly Gly Arg
          85           90           95

```

```

Leu Thr Phe Gly Asp Gly Thr Thr Leu Thr Val Lys Pro Asp Ile Gln
          100          105          110

```

```

Asn Pro Glu Pro Ala Val Tyr Gln Leu Lys Asp Pro Arg Ser Gln Asp
          115          120          125

```

```

Ser Thr Leu Cys Leu Phe Thr Asp Phe Asp Ser Gln Ile Asn Val Pro
          130          135          140

```

```

Lys Thr Met Glu Ser Gly Thr Phe Ile Thr Asp Lys Thr Val Leu Asp

```

145 150 155 160

Met Lys Ala Met Asp Ser Lys Ser Asn Gly Ala Ile Ala Trp Ser Asn  
165 170 175

Gln Thr Ser Phe Thr Cys Gln Asp Ile Phe Lys Glu Thr Asn Ala Thr  
180 185 190

Tyr Pro Ser Ser Asp Val Pro Cys Asp Ala Thr Leu Thr Glu Lys Ser  
195 200 205

Phe Glu Thr Asp Met Asn Leu Asn Phe Gln Asn Leu Ser  
210 215 220

<210> 32  
<211> 353  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Sequence of T cell receptor clone4JFH beta chain construct with  
initial Nco1 restriction site and terminal Xho1 restriction site

<400> 32  
ccatggccag ccagaccatc catcagtggc ctgccaccct ggtgcagcct gtgggatctc 60  
ctctgagcct ggaatgcacc gtggaaggca ccagcaaccc caacctgtac tggtacagac 120  
aggccgctgg cagaggcccc cagctgctgt ttactgggg cccctttggc cagatcagca 180  
gcgaggtgcc ccagaacctg agcgccagca gaccccagga ccggcagttt atcctgagca 240  
gcaagaagct gctgctgagc gacagcggct tctacctgtg cgcttggagc gagacaggcc 300  
tgggcatggg cggatggcag tttggcgagg gcagcagact gacagtgctc gag 353

<210> 33  
<211> 117  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Sequence of T cell receptor clone 4JFH beta chain construct

<400> 33

Met Ala Ser Gln Thr Ile His Gln Trp Pro Ala Thr Leu Val Gln Pro  
1 5 10 15

Val Gly Ser Pro Leu Ser Leu Glu Cys Thr Val Glu Gly Thr Ser Asn  
20 25 30

Pro Asn Leu Tyr Trp Tyr Arg Gln Ala Ala Gly Arg Gly Pro Gln Leu  
35 40 45

Leu Phe Tyr Trp Gly Pro Phe Gly Gln Ile Ser Ser Glu Val Pro Gln  
50 55 60

Asn Leu Ser Ala Ser Arg Pro Gln Asp Arg Gln Phe Ile Leu Ser Ser  
65 70 75 80

Lys Lys Leu Leu Leu Ser Asp Ser Gly Phe Tyr Leu Cys Ala Trp Ser  
85 90 95

Glu Thr Gly Leu Gly Met Gly Gly Trp Gln Phe Gly Glu Gly Ser Arg  
100 105 110

Leu Thr Val Leu Glu  
115

<210> 34  
<211> 53  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer for mutation of CDR3 of c12/c2 TCR alpha chain

<220>  
<221> misc\_feature  
<222> (20)..(21)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (23)..(24)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (31)..(31)  
<223> n is a, c, g, or t

<400> 34  
cacgcggcac gcgggtgaan nsnscctwc natgtaaggg tggaagcccg etc

53

<210> 35  
<211> 52  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer used to mutate CDR3 of c12/c2 TCR alpha chain

<220>  
<221> misc\_feature  
<222> (23)..(23)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature

<222> (30)..(31)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (33)..(34)  
<223> n is a, c, g, or t

<400> 35  
ctcgcccgaag ggtgggaatg tangwtccsn nsnnaagtgg gcgcacggcg ca

52

<210> 36  
<211> 17  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Strategy used to mutate CDR3 of c12/c2 TCR alpha chain

<220>  
<221> misc\_feature  
<222> (7)..(8)  
<223> Xaa can be any naturally occurring amino acid

<400> 36

Cys Ala Val Arg Pro Leu Xaa Xaa Gly Ser Tyr Ile Pro Thr Phe Gly  
1 5 10 15

Arg

<210> 37  
<211> 17  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Strategy used to mutate CDR3 of c12/c2 TCR alpha chain

<220>  
<221> misc\_feature  
<222> (7)..(8)  
<223> Xaa can be any naturally occurring amino acid

<400> 37

Cys Ala Val Arg Pro Leu Xaa Xaa Gly Thr Tyr Ile Pro Thr Phe Gly  
1 5 10 15

Arg

<210> 38  
<211> 50  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer used to mutate CDR3 of c12/c2 TCR beta chain

<220>  
<221> misc\_feature  
<222> (20)..(21)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (29)..(30)  
<223> n is a, c, g, or t

<400> 38  
gtacttttgc gccagctccn nsstcggggn saccggcgag ctggttctttg

50

<210> 39  
<211> 49  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer used to mutate CDR3 of c12/c2 TCR beta chain

<220>  
<221> misc\_feature  
<222> (21)..(22)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (30)..(31)  
<223> n is a, c, g, or t

<400> 39  
caaagaacag ctcgccgts nccccgassn nggagctggc gcaaaagta

49

<210> 40  
<211> 16  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Strategy used to mutate CDR3 of c12/c2 TCR beta chain

<400> 40

Tyr Phe Cys Ala Ser Ser Asn Val Gly Asn Thr Gly Glu Leu Phe Phe  
1                    5                                    10                                    15

<210> 41  
<211> 16  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Strategy used to mutate CDR3 of c12/c2 TCR beta chain



<400> 41

Tyr Phe Cys Ala Ser Ser Asn Leu Gly Asn Thr Gly Glu Leu Phe Phe  
1 5 10 15

<210> 42

<211> 462

<212> DNA

<213> Artificial Sequence

<220>

<223> Sequence of CD3 zeta in pINT21 CAR1

<400> 42

ggcgccatgg cccaggtcgc ggccgcaagc ggactgctgg acccaaagct gtgctacctg 60  
ctggatggga tcctgttcat ctacggtgtg attctgacag ccctgttcct gcgagtcaag 120  
ttcagccgga gcgccgacgc accagcatac cagcaggggc agaatcagct gtataacgag 180  
ctgaatctgg gtcggagggga ggaatacgcac gtgctggata agagacgcgg cagggatccc 240  
gaaatgggcg gaaagcctca gcgacggaaa aaccacagc agggactgta caatgaactg 300  
cagaaggaca aaatggctga ggcatattct gaaatcggca tgaagggaga gaggagacgc 360  
ggcaaaggac acgatgggct gtaccagggt ctgagtacag ccactaagga cacctatgat 420  
gcctgcata tgcaggctct gccaccaga taataaaagc tt 462

<210> 43

<211> 4

<212> PRT

<213> Artificial Sequence

<220>

<223> Fragment encoded by pINT21 CAR1 vector

<400> 43

Gly Ala Met Ala  
1

<210> 44

<211> 141

<212> PRT

<213> Artificial Sequence

<220>

<223> CDR3 zeta sequence encoded by pINT21 CAR1

<400> 44

Ser Gly Leu Leu Asp Pro Lys Leu Cys Tyr Leu Leu Asp Gly Ile Leu  
1 5 10 15

Phe Ile Tyr Gly Val Ile Leu Thr Ala Leu Phe Leu Arg Val Lys Phe  
20 25 30

Ser Arg Ser Ala Asp Ala Pro Ala Tyr Gln Gln Gly Gln Asn Gln Leu  
35 40 45

Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp Val Leu Asp  
50 55 60

Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro Gln Arg Arg  
65 70 75 80

Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp Lys Met  
85 90 95

Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg Arg Gly  
100 105 110

Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys Asp  
115 120 125

Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg  
130 135 140

<210> 45  
<211> 705  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> pINT21 CAR2 Sequence encoding CD3 zeta

<400> 45  
gccatggccc aggtcgcggc cgcaacaaca accccagccc ccagacctcc taccctgccc 60  
cctacaattg ccagccagcc tctgagcctg aggcccgagg cttgtagacc agctgctggc 120  
ggagccgtgc acaccagagg actggatttc gcctgcgaca tctacatctg ggcccctctg 180  
gccggcacat gtggcgtgct gctgctgagc ctcgatgatca ccctgtactg caagcggggc 240  
agaaagaaac tgctgtacat ctttaagcag cccttcatgc ggcccgtgca gaccaccagc 300  
gaagaggagc gctgctcctg cagattcccc gaggaagaag aaggcggctg cgagctgaga 360  
gtgaagttca gcagatccgc cgacgccctt gcctacaagc agggccagaa ccagctgtac 420  
aacgagctga acctgggcag acgggaagag tacgacgtgc tggacaagcg gagaggccgg 480  
gaccagaga tgggcggaaa gccagaaga aagaaccccc aggaaggcct gtataacgaa 540  
ctgcagaaaag acaaaatggc cgaggcctac agcgagatcg gaatgaaggc cgagcggaga 600  
agaggcaagg ggcacgatgg cctgtaccag ggctgagca ccgccaccaa ggacacctat 660  
gacgccctgc acatgcaggc cctgccccct agataataaa agctt 705

<210> 46  
<211> 223

<212> PRT  
<213> Artificial Sequence

<220>

<223> Sequence encoded by pINT21 CAR2: CD8 hinge / transmembrane region, 4-1BB and CD3 zeta

<400> 46

Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala  
1 5 10 15

Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly  
20 25 30

Gly Ala Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr Ile  
35 40 45

Trp Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu Ser Leu Val  
50 55 60

Ile Thr Leu Tyr Cys Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe  
65 70 75 80

Lys Gln Pro Phe Met Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly  
85 90 95

Cys Ser Cys Arg Phe Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu Arg  
100 105 110

Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly Gln  
115 120 125

Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp  
130 135 140

Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro  
145 150 155 160

Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp  
165 170 175

Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg  
180 185 190

Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr  
195 200 205

Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg  
210 215 220

<210> 47  
<211> 750  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Sequence of FMC63 H-L (anti-CD19 antibody)

<400> 47  
gccatggccg aagtgaaact gcaggagtct ggaccggcc tgggtggcccc atctcagtct 60  
ctgagcgtga cctgtaccgt gtccggcgtg tccctgcctg actatggcgt gtctctggatc 120  
agacagcccc ccagaaaggg cctggaatgg ctgggagtga tctggggcag cgaaaccacc 180  
tactacaaca gcgccctgaa gtcccggctg accatcatca aggacaactc caagagccag 240  
gtgttctga agatgaacag cctgcagacc gacgacaccg ccatctacta ctgcgccaag 300  
cactactact acggcggcag ctacgctatg gactactggg gccagggcac ctcggtcacc 360  
gtctcgagtg gtggaggcgg ttcaggcggg ggtggctctg gcggtggcgc tagcgacatc 420  
cagatgacc agaccaccag cagcctgagc gccagcctgg gcgatagagt gaccatcagc 480  
tgcaagacca gccaggacat cagcaagtac ctgaactggt atcagcagaa acccgacggc 540  
accgtgaagc tgctgatcta ccacaccagc agactgcaca gggcgctgcc cagcagattt 600  
tccggctctg gcagcggcac cgactacagc ctgaccatct ccaacctgga acaggaagat 660  
atcgctacct acttctgtca gcaaggcaac accctgcctt acaccttcgg cggagggacc 720  
aagctggaga tcaaacgtac cgcggccgca 750

<210> 48  
<211> 250  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Sequence of FMC63 H-L (anti-CD19 antibody)

<400> 48

Ala Met Ala Glu Val Lys Leu Gln Glu Ser Gly Pro Gly Leu Val Ala  
1 5 10 15  
Pro Ser Gln Ser Leu Ser Val Thr Cys Thr Val Ser Gly Val Ser Leu  
20 25 30  
Pro Asp Tyr Gly Val Ser Trp Ile Arg Gln Pro Pro Arg Lys Gly Leu  
35 40 45  
Glu Trp Leu Gly Val Ile Trp Gly Ser Glu Thr Thr Tyr Tyr Asn Ser  
50 55 60  
Ala Leu Lys Ser Arg Leu Thr Ile Ile Lys Asp Asn Ser Lys Ser Gln

65					70														75				80
Val	Phe	Leu	Lys	Met	Asn	Ser	Leu	Gln	Thr	Asp	Asp	Thr	Ala	Ile	Tyr								
				85						90				95									
Tyr	Cys	Ala	Lys	His	Tyr	Tyr	Tyr	Gly	Gly	Ser	Tyr	Ala	Met	Asp	Tyr								
			100					105					110										
Trp	Gly	Gln	Gly	Thr	Ser	Val	Thr	Val	Ser	Ser	Gly	Gly	Gly	Gly	Ser								
		115					120						125										
Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Ala	Ser	Asp	Ile	Gln	Met	Thr	Gln								
		130				135					140												
Thr	Thr	Ser	Ser	Leu	Ser	Ala	Ser	Leu	Gly	Asp	Arg	Val	Thr	Ile	Ser								
145					150					155					160								
Cys	Arg	Ala	Ser	Gln	Asp	Ile	Ser	Lys	Tyr	Leu	Asn	Trp	Tyr	Gln	Gln								
				165					170					175									
Lys	Pro	Asp	Gly	Thr	Val	Lys	Leu	Leu	Ile	Tyr	His	Thr	Ser	Arg	Leu								
			180				185						190										
His	Ser	Gly	Val	Pro	Ser	Arg	Phe	Ser	Gly	Ser	Gly	Ser	Gly	Thr	Asp								
		195					200						205										
Tyr	Ser	Leu	Thr	Ile	Ser	Asn	Leu	Glu	Gln	Glu	Asp	Ile	Ala	Thr	Tyr								
	210					215					220												
Phe	Cys	Gln	Gln	Gly	Asn	Thr	Leu	Pro	Tyr	Thr	Phe	Gly	Gly	Gly	Thr								
225					230					235					240								
Lys	Leu	Glu	Ile	Lys	Arg	Thr	Ala	Ala	Ala														
			245						250														

```

<210> 49
<211> 327
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence of lox1 binding Adhiron lox1A

<400> 49
gccatggctg ctacaggcgt gcgggctgtg cccggcaatg agaacagcct ggaaatcgag          60
gaactggcca gattcgccgt ggacgagcac aacaagaaag agaacgcctt gctggaattc          120
gtgCGGGtCG tgaaggccaa agagcagtgg agcgaggccg acaacgactg gcacaccatg          180
tactacctga ccctggaagc caaggacggc ggcaagaaga agctgtacga ggccaaagtg          240

```

tgggtcaagc tggacctgga aacctggcag cacttcaact tcaaagagct ccaggaattc 300  
aagcccgtgg gcgacgctgc ggccgcg 327

<210> 50  
<211> 109  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Sequence of Lox1 binding Adhiron lox1A

<400> 50

Ala Met Ala Ala Thr Gly Val Arg Ala Val Pro Gly Asn Glu Asn Ser  
1 5 10 15

Leu Glu Ile Glu Glu Leu Ala Arg Phe Ala Val Asp Glu His Asn Lys  
20 25 30

Lys Glu Asn Ala Leu Leu Glu Phe Val Arg Val Val Lys Ala Lys Glu  
35 40 45

Gln Trp Ser Glu Ala Asp Asn Asp Trp His Thr Met Tyr Tyr Leu Thr  
50 55 60

Leu Glu Ala Lys Asp Gly Gly Lys Lys Lys Leu Tyr Glu Ala Lys Val  
65 70 75 80

Trp Val Lys Leu Asp Leu Glu Thr Trp Gln His Phe Asn Phe Lys Glu  
85 90 95

Leu Gln Glu Phe Lys Pro Val Gly Asp Ala Ala Ala Ala  
100 105

<210> 51  
<211> 333  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Sequence of lox1 binding Adhiron lox1B

<400> 51

gccatggctg ctacaggcgt gcgggctgtg cccggcaatg agaacagcct ggaaatcgag 60  
gaactggcca gattcgccgt ggacgagcac aacaagaaag agaacgcct gctggaattc 120  
gtgcggtcg tgaaggcaa agacgaggaa cagcccatcg gcgagcacc cgtgaacgac 180  
accatgtact acctgaccct ggaagccaag gacggcggca agaagaagct gtacgaggcc 240  
aaagtgtggg tcaagcgggtg gctgcggttc accgagatct acaacttcaa agagctocag 300  
gaattcaagc ccgtgggcca cgctgcggcc gcg 333

<210> 52  
<211> 111  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Sequence of lox1 binding Adhiron lox1B

<400> 52

Ala Met Ala Ala Thr Gly Val Arg Ala Val Pro Gly Asn Glu Asn Ser  
1 5 10 15

Leu Glu Ile Glu Glu Leu Ala Arg Phe Ala Val Asp Glu His Asn Lys  
20 25 30

Lys Glu Asn Ala Leu Leu Glu Phe Val Arg Val Val Lys Ala Lys Glu  
35 40 45

Gln Glu Gln Pro Ile Gly Glu His Pro Val Asn Asp Thr Met Tyr Tyr  
50 55 60

Leu Thr Leu Glu Ala Lys Asp Gly Gly Lys Lys Lys Leu Tyr Glu Ala  
65 70 75 80

Lys Val Trp Val Lys Arg Trp Leu Arg Phe Thr Glu Ile Tyr Asn Phe  
85 90 95

Lys Glu Leu Gln Glu Phe Lys Pro Val Gly Asp Ala Ala Ala Ala  
100 105 110

<210> 53  
<211> 54  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Adhiron\_mut1 primer

<220>  
<221> misc\_feature  
<222> (23)..(24)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (26)..(27)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (29)..(30)  
<223> n is a, c, g, or t

<220>

<221> misc\_feature  
<222> (32)..(33)  
<223> n is a, c, g, or t

<400> 53  
cagggtcagg tagtacatgg tsnnsnnsnn snnctgctct ttggccttca cgac

54

<210> 54  
<211> 48  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Adhiron\_mut2 primer

<220>  
<221> misc\_feature  
<222> (20)..(21)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (23)..(24)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (26)..(27)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (29)..(30)  
<223> n is a, c, g, or t

<400> 54  
ctggagctct ttgaagttsn nsnnsnnsnn cttgaccac actttggc

48

<210> 55  
<211> 54  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Mutagenic primer for construction of library of binders within  
loop 1 (adhiron mut 1)

<220>  
<221> misc\_feature  
<222> (22)..(23)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (25)..(26)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (28)..(29)



<223> n is a, c, g, or t

<220>

<221> misc\_feature

<222> (31)..(32)

<223> n is a, c, g, or t

<400> 55

gtcgtgaagg ccaaagagca gnnsnnsnns nnsaccatgt actacctgac cctg

54

<210> 56

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> Adhiron mut1 loop library variant sequence

<220>

<221> misc\_feature

<222> (8)..(11)

<223> Xaa can be any naturally occurring amino acid

<400> 56

Val Val Lys Ala Lys Glu Gln Xaa Xaa Xaa Xaa Thr Met Tyr Tyr Leu  
1 5 10 15

Thr Leu

<210> 57

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Mutagenic primer for construction of library of binders within  
Adhiron mut2 loop

<220>

<221> misc\_feature

<222> (19)..(20)

<223> n is a, c, g, or t

<220>

<221> misc\_feature

<222> (22)..(23)

<223> n is a, c, g, or t

<220>

<221> misc\_feature

<222> (25)..(26)

<223> n is a, c, g, or t

<220>

<221> misc\_feature

<222> (28)..(29)

<223> n is a, c, g, or t

<400> 57  
gccaaagtgt ggggtcaagnn snnsnnsnns aacttcaaag agctccag 48

<210> 58  
<211> 16  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Adhiron mut 2 loop library variant sequence

<220>  
<221> misc\_feature  
<222> (7)..(10)  
<223> Xaa can be any naturally occurring amino acid

<400> 58

Ala Lys Val Trp Val Lys Xaa Xaa Xaa Xaa Asn Phe Lys Glu Leu Gln  
1 5 10 15

<210> 59  
<211> 105  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Trypsin binding Knottin MCoTI-IIsequence

<400> 59  
gccatggccg gtgtgtgccc caagatcttg aaaaagtgcc gccgtgacag cgattgtccc 60

ggcgctgca tctgccgcg caatggctat tgcggagcgg ccgca 105

<210> 60  
<211> 32  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Trypsin binding Knottin MCo-TI-II sequence

<400> 60

Ala Met Ala Gly Val Cys Pro Lys Ile Leu Lys Lys Cys Arg Arg Asp  
1 5 10 15

Ser Asp Cys Pro Gly Ala Cys Ile Cys Arg Gly Asn Gly Tyr Cys Gly  
20 25 30

<210> 61  
<211> 99  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer for creation of library of Knottin mutants

<220>  
<221> misc\_feature  
<222> (11)..(11)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (14)..(14)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (17)..(17)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (20)..(20)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (23)..(23)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (26)..(26)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (29)..(29)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (32)..(32)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (35)..(35)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (38)..(38)  
<223> n is a, c, g, or t

<400> 61  
ggtgtgtgcv nsvnsvnsvn svnsvnsvns vnsvnsvns ggcgcctga cagcgattgt 60

cccgcgctt gcattgccg cggcaatggc tattgcgga 99

<210> 62  
<211> 99  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Sequence varied in library of knottin mutants

<220>

<221> misc\_feature

<222> (62)..(62)

<223> n is a, c, g, or t

<220>

<221> misc\_feature

<222> (65)..(65)

<223> n is a, c, g, or t

<220>

<221> misc\_feature

<222> (68)..(68)

<223> n is a, c, g, or t

<220>

<221> misc\_feature

<222> (71)..(71)

<223> n is a, c, g, or t

<220>

<221> misc\_feature

<222> (74)..(74)

<223> n is a, c, g, or t

<220>

<221> misc\_feature

<222> (77)..(77)

<223> n is a, c, g, or t

<220>

<221> misc\_feature

<222> (80)..(80)

<223> n is a, c, g, or t

<220>

<221> misc\_feature

<222> (83)..(83)

<223> n is a, c, g, or t

<220>

<221> misc\_feature

<222> (86)..(86)

<223> n is a, c, g, or t

<220>

<221> misc\_feature

<222> (89)..(89)

<223> n is a, c, g, or t

<400> 62

tccgcaatag ccattgccgc ggcagatgca ggcgccggga caatcgctgt cacggcggca 60

snbsnbsnbs nbsnbsnbsn bsnbsnbsnb gcacacacc 99

<210> 63

<211> 33

<212> PRT

<213> Artificial Sequence

<220>  
<223> Sequence of Knottin varied in library

<220>  
<221> misc\_feature  
<222> (4)..(13)  
<223> Xaa can be any naturally occurring amino acid

<400> 63

Gly Val Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Cys Arg Arg  
1 5 10 15

Asp Ser Asp Cys Pro Gly Ala Cys Ile Cys Arg Gly Asn Gly Tyr Cys  
20 25 30

Gly

<210> 64  
<211> 121  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> pD7-Sce1 sequence nucleotides 1 - 120

<400> 64  
ggtaccgaat tctaggata acagggtaat atgcatcttc tgacctcttc tcttctctcc 60  
acagggcatg gcaaacctc tgagccagga agaaagcaca ctgattgaaa gagcaaccgc 120  
t 121

<210> 65  
<211> 18  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Fragment of blasticidin resistance gene encoded by pD7-Sce1

<400> 65

Met Ala Lys Pro Leu Ser Gln Glu Glu Ser Thr Leu Ile Glu Arg Ala  
1 5 10 15

Thr Ala

<210> 66  
<211> 120  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> pD7-ObLiGaRe sequence fragment

<400> 66  
ggtaccgaat tcttttctgt caccaatcct ggggccacta gggacactgt ggggtggagg 60  
aaatgcatct tctgacctct tctcttcctc ccacagggca tggcaaaacc tctgagccag 120

<210> 67  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Fragment of Blasticidin resistance gene encoded by pD7-ObLiGaRe

<400> 67

Met Ala Lys Pro Leu Ser Gln  
1 5

<210> 68  
<211> 1061  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> ROSA 26 left homology arm sequence

<400> 68  
gcgatcgcgc tgattggctt cttttcctcc cgccgtgtgt gaaaacacaa atggcgtggt 60  
ttggttggcg taaggcgcct gtcagttaac ggcagccgga gtgctgcagcc gccggcagcc 120  
tcgctctgcc cactgggtgg ggcgggaggt aggtgggggtg aggcgagctg gacgtgcggg 180  
cgcggtcggc ctctggcggg gcgggggagg ggagggaggg tcagcgaaag tagctcgcgc 240  
gcgagcggcc gccaccctc cccttcctct gggggagtcg ttttaccgc cgccggccgg 300  
gcctcgtcgt ctgattggct ctcggggccc agaaaactgg cccttgccat tggctcgtgt 360  
tcgtgcaagt tgagtccatc cgccggccag cgggggcggc gaggaggcgc tcccaggttc 420  
cgccctccc ctcgccccg cgccgcagag tctggcgcg cgccctgcg caacgtggca 480  
ggaagcgcgc gctggggggcg gggacgggca gtagggctga ggggctgcgg ggggggtgca 540  
agcacgtttc cgacttgagt tgctcaaga gggcgtgct gagccagacc tccatcgcgc 600  
actccgggga gtggaggaa ggagcgagg ctcagttgg ctgttttgg ggcaggaagc 660  
acttgcctc ccaaagtcgc tctgagttgt tatcagtaag ggagctgcag tggagtaggc 720  
ggggagaagg ccgcaccctt ctccggagg gggaggggag tgttgcaata cctttctggg 780  
agttctctgc tgctcctgg cttctgagga ccgccctgg cctgggagaa tcccttcccc 840  
ctcttcctc gtgatctgca actccagtct ttctagaatg cattaagga tctgtagggc 900  
gcagtagtcc agggtttctc tgatgatgtc atacttatcc tgtccctttt ttttccacag 960  
ctcgcggttg aggacaaact cttcgcggtc tttccagtgg ggatcgacgg tatcgtagag 1020

tcgaggccgc tctaggaatt cacgccgcca ccatgaccga g

1061

<210> 69  
<211> 703  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> ROSA 26 right homology arm sequence

<400> 69  
gtatacggga attgaacagg tgtaaaattg gagggacaag acttcccaca gattttcggg 60  
tttgtcggga agttttttaa taggggcaaa taaggaaaat gggaggatag gtagtcatct 120  
ggggttttat gcagcaaac tacaggttat tattgcttgt gatccgcctc ggagtatfff 180  
ccatcgaggg agattaaaga catgctcacc cgagttttat actctcctgc ttgagatcct 240  
tactacagta tgaaattaca gtgtcgcgag ttagactatg taagcagaat tttaatcatt 300  
tttaaagagc ccagtacttc atatccattt ctcccgtctc ttctgcagcc ttatcaaaag 360  
gtattttaga aactcattt tagccccatt ttcatttatt atactggctt atccaacccc 420  
tagacagagc attggcattt tccctttcct gatcttagaa gtctgatgac tcatgaaacc 480  
agacagatta gttacataca ccacaaatcg aggctgtagc tggggcctca aactgcagt 540  
tcttttataa ctcttagta cactttttgt tgatcctttg ccttgatcct taatfttcag 600  
tgtctatcac ctctcccgtc aggtggtggt ccacatttgg gcctattctc agtccaggga 660  
gttttacaac aatagatgta ttgagaatcc aacctcctgc agg 703

<210> 70  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Left TALEN recognition sequence

<400> 70  
tcccctccac cccacagt 18

<210> 71  
<211> 15  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> TALEN spacer

<400> 71  
ggggcacta gggac 15

<210> 72  
<211> 18  
<212> DNA

<213> Artificial Sequence  
 <220>  
 <223> Right TALEN target sequence  
 <400> 72  
 aggattggtg acagaaaa 18

<210> 73  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Right TALEN sequence  
 <400> 73  
 ttttctgtca ccaatcct 18

<210> 74  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> AAVS-Left-arm-junction-PCR-Forw (9625) primer  
 <400> 74  
 ccggaactct gccctctaac 20

<210> 75  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> BSD\_Junction PCR-rev (9626) primer  
 <400> 75  
 tagccacaga atagtcttcg gag 23

<210> 76  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Donor\_plasmid\_seq\_PDGFRTM-2 Forw primer  
 <400> 76  
 acacgcagga ggccatcgtg g 21

<210> 77  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> AAVS1\_right arm\_junction\_PCR\_rev primer



<400> 77  
 tcctgggata ccccgaagag 20

<210> 78  
 <211> 54  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2598

<400> 78  
 ttttttttaa ttaagattat tgactagtta ttaatagtaa tcaattacgg ggtc 54

<210> 79  
 <211> 52  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2619

<400> 79  
 ttttttgttt aaacttacct tggatccctt gccggggctc aggctcaggg ac 52

<210> 80  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> M13Leadseq primer

<400> 80  
 aaattattat tcgcaattcc tttggttggt cct 33

<210> 81  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Notmycseq primer

<400> 81  
 ggccccattc agatcctctt ctgagatgag 30

<210> 82  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2623

<400> 82  
 taaagtaggc ggtcttgaga cg 22

<210> 83  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 2624  
  
 <400> 83  
 gaaggtgctg ttgaactggt cc 22

<210> 84  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 2595  
  
 <400> 84  
 gagggctctg gcagctagc 19

<210> 85  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 2597  
  
 <400> 85  
 tcgagactgt gacgaggctg 20

<210> 86  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 2625  
  
 <400> 86  
 ccttggtgct ggcactcga 19

<210> 87  
 <211> 40  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 1999  
  
 <400> 87  
 aaaaagcagg ctaccatgag ggcttgatc ttctttctcc 40

<210> 88  
 <211> 44  
 <212> DNA

<213> Artificial Sequence  
 <220>  
 <223> TALEN recognition sequence  
 <400> 88  
 tccaccggtc gccaccatgg tgagcaaggg cgaggagctg ttca 44

<210> 89  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> AAVS1\_HA-L\_Nested\_Forw1 primer  
 <400> 89  
 gtgcccttgc tgtgccgccg gaactctgcc ctc 33

<210> 90  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> EGFP\_Synthetic\_gene\_Rev\_Assembly primer  
 <400> 90  
 ttcacgtcgc cgtccagctc gac 23

<210> 91  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Purotk\_seq\_fow2 primer  
 <400> 91  
 tccataaccga cgatctgcga c 21

<210> 92  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> J44 primer  
 <400> 92  
 aaaagcgcct cccctaccgg gtagaat 27

<210> 93  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> J46 primer

<400> 93  
 ggcgacacgg aatggtgaa tactcat 27

<210> 94  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> J48 primer

<400> 94  
 cactacacc agaagtcct ggcctg 27

<210> 95  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> pSANG10pelB primer

<400> 95  
 cgctgccag ccggccatgg 20

<210> 96  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2097

<400> 96  
 gatggtgatg atgatgtgcg gatgcg 26

<210> 97  
 <211> 38  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2423

<400> 97  
 tttttccat gggccggccc tccttcagtt tagttgag 38

<210> 98  
 <211> 39  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2437

<400> 98  
 ttttttgagg ccgcggaagc cgtgatctcc ttctctctc 39

<210> 99  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer 2622

<400> 99  
gaacaggaac acggaaggtc 20

<210> 100  
<211> 22  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer 41679

<400> 100  
atgagttgga gctgtatcat cc 22

<210> 101  
<211> 18  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer 2621

<400> 101  
gcattccacg gcggccgc 18

<210> 102  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Peptide SLLMWITQV (NY-ESO-1 157-165) recognised by c12c2

<400> 102

Ser Leu Leu Met Trp Ile Thr Gln Val  
1 5

<210> 103  
<211> 10  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Peptide MART-1 26-35 recognised by MEL5 TCR

<400> 103

Glu Leu Ala Gly Ile Gly Ile Leu Thr Val  
1 5 10

<210> 104  
<211> 30  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Primer 4JFH-Valpha-F  
  
<400> 104  
acacacgcta gccagaaaga ggtggaacag 30

<210> 105  
<211> 50  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Primer 2770  
  
<220>  
<221> misc\_feature  
<222> (21)..(22)  
<223> n is a, c, g, or t  
  
<220>  
<221> misc\_feature  
<222> (30)..(31)  
<223> n is a, c, g, or t  
  
<400> 105  
caaagaacag ctcgccgts nccccgassn nggagctggc gcaaaagtac 50

<210> 106  
<211> 53  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Primer 2771  
  
<220>  
<221> misc\_feature  
<222> (23)..(23)  
<223> n is a, c, g, or t  
  
<220>  
<221> misc\_feature  
<222> (30)..(31)  
<223> n is a, c, g, or t  
  
<220>  
<221> misc\_feature  
<222> (33)..(34)  
<223> n is a, c, g, or t  
  
<400> 106  
ctcgcccga ggtgggaatg tangwtccsn nsnaagtgg gcgcacggcg cac 53

<210> 107  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 2780  
  
 <400> 107  
 ctggcagcta gcaagcagga ag 22

<210> 108  
 <211> 22  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 2781  
  
 <400> 108  
 tacattccca ccttcgggcg ag 22

<210> 109  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 2782  
  
 <400> 109  
 ttttttgcgg ccgcgacag gttctg 26

<210> 110  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 2783  
  
 <400> 110  
 cgtaagctgg taccttatta tctaggg 27

<210> 111  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 2785  
  
 <400> 111  
 ccctagataa taaggtacca gcttacg 27

<210> 112  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2787

<400> 112  
 accggcgagc tgttctttg 19

<210> 113  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2788

<400> 113  
 agtgacaagc ttttattatc tgggtg 26

<210> 114  
 <211> 19  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2789

<400> 114  
 caggtcctcg agcactgtc 19

<210> 115  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> CRISPR1 double-stranded DNA insert strand

<400> 115  
 ggggccacta gggacaggat gtttt 25

<210> 116  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> CRISPR1 double-stranded DNA insert strand

<400> 116  
 catcctgtcc ctagtggccc ccggtg 26

<210> 117  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> CRISPR2 double-stranded DNA insert strand



<400> 117  
 gtcaccaatc ctgtccctag gtttt 25

<210> 118  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> CRISPR2 double-stranded DNA insert strand

<400> 118  
 ctagggacag gattggtgac cggtg 25

<210> 119  
 <211> 102  
 <212> RNA  
 <213> Artificial Sequence

<220>  
 <223> CRISPR 1 RNA

<400> 119  
 gggggggccac uagggacagg auguuuuaga gcuagaaaua gcaaguuaaa auaaggcuag 60  
 uccguuauca acuugaaaaa guggcaccga gucggugcuu uu 102

<210> 120  
 <211> 102  
 <212> RNA  
 <213> Artificial Sequence

<220>  
 <223> CRISPR 2 RNA

<400> 120  
 gggucaccaa uccuguccu agguuuuaga gcuagaaaua gcaaguuaaa auaaggcuag 60  
 uccguuauca acuugaaaaa guggcaccga gucggugcuu uu 102

<210> 121  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2723

<400> 121  
 cgcgccagaa gtctcaccaa gccca 25

<210> 122  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2724

<400> 122  
cgcgtgggct tggtagact tctgg 25

<210> 123  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer 2768

<400> 123  
aattctccc tccacccac agtagggaca gtggggccag gattggtgac agaaaatgca 60

<210> 124  
<211> 52  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer 2769

<400> 124  
ttttctgtca ccaatcctgg cccactgtc cctactgtgg ggtggagggg ag 52

<210> 125  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer 2778

<400> 125  
aattctaggg ataacagggt aatatgca 28

<210> 126  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer 2779

<400> 126  
tattaccctg ttatccctag 20

<210> 127  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer 2808

<400> 127  
aattcttttc tgtcaccaat cctggggcca ctagggacac tgtggggtgg aggggatgca 60

<210> 128  
 <211> 57  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 2809  
  
 <400> 128  
 tcccctccac cccacagtgt ccctagtggc cccaggattg gtgacagaaa agaattg 57

<210> 129  
 <211> 41  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer J60  
  
 <400> 129  
 acacacggta ccgcatcgc gctgattggc ttcttttct c 41

<210> 130  
 <211> 35  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 2706  
  
 <400> 130  
 tttttatgc attctagaaa gactggagtt gcaga 35

<210> 131  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 2715  
  
 <400> 131  
 gagcgtccgc ccaccctc 18

<210> 132  
 <211> 18  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Primer 2716  
  
 <400> 132  
 gaggtgggc ggacgctc 18

<210> 133  
 <211> 33  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2709  
  
 <400> 133  
 ttttttatgc attaagggat ctgtagggcg cag 33

<210> 134  
 <211> 21  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2710  
  
 <400> 134  
 gtgaattcct agagcggcct c 21

<210> 135  
 <211> 50  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer 2745  
  
 <400> 135  
 gaggccgctc taggaattca cgccgccacc atgaccgagt acaagcccac 50

<210> 136  
 <211> 48  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer J59  
  
 <400> 136  
 aaaaaaagat ctgtgtgttt cgaatcaggc accgggcttg cgggtcat 48

<210> 137  
 <211> 36  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer J61  
  
 <400> 137  
 tttttgtat acgggaattg aacaggtgta aaattg 36

<210> 138  
 <211> 43  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Primer J62

<400> 138  
ttttttcctg caggaggttg gattctcaat acatctattg ttg 43

<210> 139  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer 2701

<400> 139  
gccgacgtct cgtcgctgat gtttt 25

<210> 140  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer 2702

<400> 140  
atcagcgacg agacgtcggc cggtg 25

<210> 141  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer 2703

<400> 141  
cgcccatctt ctagaaagac gtttt 25

<210> 142  
<211> 25  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Primer 2704

<400> 142  
gtctttctag aagatgggcg cggtg 25

**ФОРМУЛА ИЗОБРЕТЕНИЯ**

1. Способ получения библиотеки клонов эукариотических клеток, содержащих ДНК, кодирующую широкий набор связывающих молекул, включающий этапы в которых обеспечивают молекулы донорной ДНК, кодирующие связывающие молекулы, и эукариотические клетки,

вводят донорную ДНК в клетки и обеспечивают присутствие в клетке сайт-специфической нуклеазы, причем указанная нуклеаза расщепляет распознаваемую последовательность в клеточной ДНК с образованием сайта встраивания, в котором донорная ДНК встраивается в клеточную ДНК, при этом встраивание осуществляется посредством эндогенных клеточных механизмов репарации ДНК, в результате чего создаются рекомбинантные клетки, содержащие донорную ДНК, встроенную в клеточную ДНК, и

культивируют рекомбинантные клетки для получения клонов, таким образом обеспечивают библиотеку клонов эукариотических клеток, содержащих донорную ДНК, кодирующую набор связывающих молекул.

2. Способ по п. 1, отличающийся тем, что связывающие молекулы представляют собой молекулы антител.

3. Способ по п. 2, отличающийся тем, что молекулы антител представляют собой полноразмерные иммуноглобулины, IgG, Fab, scFv-Fc или scFv.

4. Способ по п. 1 или п. 2, отличающийся тем, что связывающие молекулы представляют собой мультимерные молекулы, содержащие по меньшей мере первую и вторую субъединицу.

5. Способ получения библиотеки клонов эукариотических клеток, содержащих ДНК, кодирующую широкий набор мультимерных связывающих молекул, в котором каждая связывающая молекула содержит первую и вторую субъединицу, причем указанный способ включает этапы в которых

обеспечивают эукариотические клетки, содержащие ДНК, кодирующую первую субъединицу, и обеспечивают молекулы донорной ДНК, кодирующие вторую субъединицу связывающей молекулы,

вводят донорную ДНК в клетки и обеспечивают присутствие в клетке сайт-специфической нуклеазы, причем указанная нуклеаза расщепляет распознаваемую последовательность в клеточной ДНК с образованием сайта встраивания, в котором донорная ДНК встраивается в клеточную ДНК, при этом встраивание осуществляется посредством эндогенных клеточных механизмов репарации ДНК, в результате чего

создаются рекомбинантные клетки, которые содержат донорную ДНК, встроенную в клеточную ДНК, и

культивируют рекомбинантные клетки для получения клонов, содержащих ДНК, кодирующую первую и вторую субъединицы мультимерной связывающей молекулы,

таким образом обеспечивают библиотеку клонов эукариотических клеток, содержащих донорную ДНК, кодирующую набор мультимерных связывающих молекул.

6. Способ получения библиотеки клонов эукариотических клеток, содержащих ДНК, кодирующую широкий набор мультимерных связывающих молекул, в котором каждая связывающая молекула содержит по меньшей мере первую и вторую субъединицу, причем указанный способ включает этапы в которых

обеспечивают молекулы первой донорной ДНК, кодирующие первую субъединицу, и обеспечивают эукариотические клетки,

вводят первую донорную ДНК в клетки и обеспечивают присутствие в клетке сайт-специфической нуклеазы, причем указанная нуклеаза расщепляет распознаваемую последовательность в клеточной ДНК с образованием сайта встраивания, в котором донорная ДНК встраивается в клеточную ДНК, при этом встраивание осуществляется посредством эндогенных клеточных механизмов репарации ДНК, в результате чего создается первый набор рекомбинантных клеток, содержащих первую донорную ДНК, встроенную в клеточную ДНК,

культивируют первый набор рекомбинантных клеток для получения первого набора клонов, содержащих ДНК, кодирующую первую субъединицу,

вводят молекулы второй донорной ДНК, кодирующие вторую субъединицу, в клетки первого набора клонов, причем вторая донорная ДНК встраивается в клеточную ДНК первого набора клонов, в результате чего создают второй набор рекомбинантных клеток, содержащих первую и вторую донорную ДНК, встроенную в клеточную ДНК, и

культивируют второй набор рекомбинантных клеток для получения второго набора клонов, причем эти клоны содержат ДНК, кодирующую первую и вторую субъединицы мультимерной связывающей молекулы,

таким образом обеспечивают библиотеку клонов эукариотических клеток, содержащих донорную ДНК, кодирующую набор мультимерных связывающих молекул.

7. Способ по п. 6, отличающийся тем, что встраивание молекул второй донорной ДНК осуществляется по способу, включающему обеспечение присутствия в клетке сайт-специфической нуклеазы, причем указанная нуклеаза расщепляет распознаваемую последовательность в клеточной ДНК с образованием сайта встраивания,

в котором донорная ДНК встраивается в клеточную ДНК, при этом встраивание осуществляется посредством эндогенных клеточных механизмов репарации ДНК.

8. Способ по любому из предшествующих пунктов, отличающийся тем, что набор связывающих молекул представляет собой ряд полипептидов, которые имеют общую структуру и одну или более областей аминокислотной последовательности, в которых имеются различия.

9. Способ по любому из предшествующих пунктов, отличающийся тем, что набор связывающих молекул представляет собой ряд молекул антител, различающихся по по одной или более областей, определяющих комплементарность.

10. Способ по любому из пп. 3-9, отличающийся тем, что мультимерные связывающие молекулы представляют собой молекулы антител, содержащие вариабельный домен тяжелой цепи (VH) и вариабельный домен легкой цепи (VL) в виде отдельных субъединиц.

11. Способ по п. 10, отличающийся тем, что мультимерные связывающие молекулы представляют собой цельные молекулы иммуноглобулинов.

12. Способ по п. 10, отличающийся тем, что мультимерные связывающие молекулы представляют собой IgG.

13. Способ по п. 10, отличающийся тем, что мультимерные связывающие молекулы представляют собой Fab.

14. Способ по любому из пп. 10-13, отличающийся тем, что первая субъединица содержит VH-домен, и при этом вторая субъединица содержит VL-домен.

15. Способ по любому из пп. 10-13, отличающийся тем, что первая субъединица содержит VL-домен, и при этом вторая субъединица содержит VH-домен.

16. Способ по любому из пп. 10-15, отличающийся тем, что молекула антитела дополнительно содержит одну или более дополнительных субъединиц, которые могут быть введены в ту же самую донорную ДНК, что и первая или вторая субъединица, или которые могут быть встроены в независимые сайты клеточной ДНК.

17. Способ по любому из предшествующих пунктов, отличающийся тем, что клетки представляют собой клетки высших эукариот с размером генома более чем  $2 \times 10^7$  пар оснований.

18. Способ по любому из предшествующих пунктов, отличающийся тем, что клетки представляют собой клетки млекопитающих, птиц, насекомых или растений.

19. Способ по п. 18, отличающийся тем, что клетки представляют собой клетки млекопитающих.



20. Способ по п. 19, отличающийся тем, что клетки представляют собой клетки НЕК293, клетки яичника китайского хомячка (CHO), клетки линии Т-лимфоцитов или клетки линии В-лимфоцитов, или любые клеточные линии, приведенные в источниках «Cancer Cell Line Encyclopedia» или «COSMIC catalogue of somatic mutations in cancer».

21. Способ по п. 20, отличающийся тем, что клетки представляют собой первичную культуру Т-клеток или линию Т-клеток.

22. Способ по п. 20, отличающийся тем, что клетки представляют собой первичную культуру В-клеток, линию В-клеток, линию пре-В-клеток или про-В-клеток.

23. Способ по п. 22, отличающийся тем, что клетки представляют собой линию мышинных пре-В-клеток 1624-5, линию ИЛ-3-зависимых про-В-клеток Ва/Г3 или В-клеток DT40 курицы.

24. Способ по любому из предшествующих пунктов, отличающийся тем, что распознаваемая последовательность находится в геномной ДНК клеток.

25. Способ по любому из пп. 1-23, отличающийся тем, что распознаваемая последовательность находится в эписомальной ДНК внутри клеток.

26. Способ по любому из предшествующих пунктов, отличающийся тем, что последовательность, распознаваемая сайт-специфической нуклеазой, встречается в клеточной ДНК только один или два раза.

27. Способ по любому из предшествующих пунктов, отличающийся тем, что сайт-специфическая нуклеаза расщепляет клеточную ДНК с образованием двухцепочечного разрыва, служащего в качестве сайта встраивания.

28. Способ по любому из предшествующих пунктов, отличающийся тем, что нуклеаза представляет собой мегануклеазу.

29. Способ по любому из пп. 1-27, отличающийся тем, что нуклеаза представляет собой нуклеазу «цинковые пальцы» (ZFN).

30. Способ по любому из пп. 1-27, отличающийся тем, что нуклеаза представляет собой нуклеазу TALE.

31. Способ по любому из пп. 1-27, отличающийся тем, что нуклеаза представляет собой нуклеазу, направляемую нуклеиновой кислотой.

32. Способ по п. 31, отличающийся тем, что расщепление ДНК осуществляется системой CRISPR/Cas.

33. Способ по любому из предшествующих пунктов, отличающийся тем, что донорная ДНК встраивается в клеточную ДНК путем гомологичной рекомбинации.

34. Способ по любому из пп. 1-32, отличающийся тем, что донорная ДНК встраивается в геномную ДНК путем негомологичного соединения концов или микрогомологично-направленного соединения концов.

35. Способ по любому из предшествующих пунктов, отличающийся тем, что донорная ДНК содержит генетический элемент для селекции клеток, в которые встраивается донорная ДНК.

36. Способ по любому из предшествующих пунктов, отличающийся тем, что встраивание донорной ДНК в клеточную ДНК приводит к экспрессии связывающей молекулы и/или экспрессии генетического селективного элемента под контролем промотора, содержащегося в клеточной ДНК.

37. Способ по любому из пп. 1-34, отличающийся тем, что донорная ДНК содержит последовательность, кодирующую связывающую молекулу, функционально связанную с промотором.

38. Способ по любому из предшествующих пунктов, отличающийся тем, что библиотека содержит по меньшей мере  $100$ ,  $10^3$ ,  $10^4$ ,  $10^5$  или  $10^6$  клонов, причем каждый клон образуется из отдельной рекомбинантной клетки, полученной в результате встраивания донорной ДНК.

39. Способ по любому из предшествующих пунктов, отличающийся тем, что библиотека содержит клоны, кодирующие по меньшей мере  $100$ ,  $10^3$ ,  $10^4$ ,  $10^5$  или  $10^6$  различных связывающих молекул.

40. Способ по любому из предшествующих пунктов, отличающийся тем, что каждый клон содержит встроенную донорную ДНК, кодирующую только одну или две молекулы из набора связывающих молекул.

41. Способ по любому из предшествующих пунктов, отличающийся тем, что эукариотические клетки являются диплоидными и содержат распознаваемую сайт-специфической нуклеазой последовательность, содержащуюся в дубликатных фиксированных локусах клеточной ДНК.

42. Способ по любому из пп. 1-39, отличающийся тем, что каждый клон содержит встроенную донорную ДНК, кодирующую одну молекулу из набора связывающих молекул.

43. Способ по любому из предшествующих пунктов, отличающийся тем, что каждая молекула донорной ДНК кодирует одну связывающую молекулу или субъединицу связывающей молекулы.

44. Способ по любому из предшествующих пунктов, отличающийся тем, что связывающие молекулы представлены на клеточной поверхности.

45. Способ по любому из предшествующих пунктов, отличающийся тем, что связывающие молекулы секретируются из клеток.

46. Способ по любому из предшествующих пунктов, дополнительно включающий:

культивирование библиотеки клонов для экспрессии связывающих молекул, выделение одного или более клонов, экспрессирующих представляющую интерес связывающую молекулу, и

создание вторичной библиотеки из одного или более выделенных клонов, причем вторичная библиотека содержит ДНК, кодирующую второй набор связывающих молекул.

47. Способ по п. 46, отличающийся тем, что создание вторичной библиотеки включает выделение донорной ДНК из одного или более выделенных клонов, введение мутации в ДНК для получения вторичной популяции молекул донорной ДНК, кодирующих второй набор связывающих молекул, и введение вторичной популяции молекул донорной ДНК в клетки с образованием вторичной библиотеки клеток, содержащих ДНК, кодирующую второй набор связывающих молекул.

48. Способ по п. 46, отличающийся тем, что создание вторичной библиотеки включает введение мутации в донорную ДНК одного или более выделенных клонов путем индукции мутации ДНК в этих клонах.

49. Способ получения широкого набора связывающих молекул, включающий получение библиотеки с помощью способа по любому из пп. 1-48 и культивирование библиотеки клонов для экспрессии связывающих молекул.

50. Способ скрининга клеток требуемого фенотипа, в котором фенотип обусловлен экспрессией клеткой связывающей молекулы, при этом способ включает получение библиотеки с помощью способа по любому из пп. 1-48, культивирование библиотеки клонов для экспрессии связывающих молекул и определение того, проявляется ли требуемый фенотип.

51. Способ по п. 50, отличающийся тем, что фенотип характеризуется экспрессией репортерного гена в клетке, которая экспрессирует связывающую молекулу.

52. Способ по п. 50 или п. 51, дополнительно включающий выделение клеток клона, который экспрессирует связывающую молекулу, которая обеспечивает получение требуемого фенотипа.

53. Способ по п. 52, дополнительно включающий выделение ДНК, кодирующей связывающую молекулу, из отобранного клона, в результате чего получая ДНК, кодирующую связывающую молекулу, которая обеспечивает получение требуемого фенотипа.

54. Способ скрининга связывающих молекул, которые распознают мишень, включающий: получение библиотеки с помощью способа по любому из пп. 1-48, культивирование клеток библиотеки для экспрессии связывающих молекул, экспонирование мишеней для воздействия связывающих молекул, обеспечивающего распознавание мишени одной или более когнатными связывающими молекулами, в случае наличия, и

обнаружение факта распознавания мишени когнатной связывающей молекулой.

55. Способ по п. 50 или п. 54, отличающийся тем, что мишень представлена в растворимой форме.

56. Способ по п. 50 или п. 54, отличающийся тем, что мишень представлена на поверхности популяции клеток-мишеней, а связывающие молекулы представлены на поверхности клеток библиотеки, причем способ включает экспонирование мишеней для воздействия связывающих молекул путем приведения клеток библиотеки в контакт с клетками-мишенями.

57. Способ по любому из пп. 50-56, отличающийся тем, что связывающие молекулы представляют собой молекулы антител, а мишень представляет собой антиген.

58. Способ по любому из пп. 50- 56, отличающийся тем, что связывающие молекулы представляют собой TCR (рецепторы Т-клеток), а мишень представляет собой комплекс ГКГС (главный комплекс гистосовместимости): пептид.

59. Способ по любому из пп. 54-58, дополнительно включающий обнаружение распознавания мишени когнатной связывающей молекулой и выделение клеток клона, содержащего ДНК, кодирующую когнатную связывающую молекулу.

60. Способ по п. 59, дополнительно включающий выделение ДНК, кодирующей связывающую молекулу, из отобранного клона, таким образом обеспечивая ДНК, кодирующую связывающую молекулу, которая распознает мишень.

61. Способ по п. 53 или п. 60, включающий введение мутации или превращение ДНК в модифицированную ДНК, кодирующую реструктурированную связывающую молекулу.

62. Способ по п. 61, отличающийся тем, что связывающая молекула представляет собой scFv, а указанный способ включает превращение ДНК, кодирующей scFv, в ДНК, кодирующую Ig или его фрагмент, с сохранением попарного объединения оригинальных переменных VH и VL доменов цепей.

63. Способ по п. 53, 60, 61 или 62, дополнительно включающий введение ДНК в клетку-хозяина.

64. Способ по п. 52, 59 или 63, дополнительно включающий культивирование клеток и концентрирование клеток для получения клеточного осадка или концентрированной клеточной суспензии.

65. Способ по п. 52, 59 или 63, дополнительно включающий культивирование клеток для экспрессии связывающей молекулы, и очистку связывающей молекулы.

66. Библиотека, полученная с помощью способа по любому из пп. 1-48.

67. Полученная *in vitro* библиотека клонов эукариотических клеток, которые экспрессируют широкий набор, состоящий по меньшей мере из  $10^3$ ,  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$ ,  $10^8$  или  $10^9$  разных связывающих молекул, причем каждая клетка содержит рекомбинантную ДНК, при этом донорная ДНК, кодирующая связывающую молекулу или субъединицу связывающей молекулы, встраивается по меньшей мере в первый и/или второй фиксированный локус в клеточной ДНК.

68. Библиотека по п. 67, отличающаяся тем, что клоны также содержат ДНК, содержащую генетический элемент для селекции клеток, которые экспрессируют кодируемую связывающую молекулу.

69. Библиотека по п. 67 или п. 68, отличающаяся тем, что связывающие молекулы представляют собой молекулы антител.

70. Библиотека по п. 69 отличающаяся тем, что молекулы антител представляют собой полноразмерные иммуноглобулины, IgG, Fab, scFv или scFv-Fc.

71. Емкость, содержащая эукариотические клетки, содержащие библиотеку по любому из пп. 66-70.

72. Емкость по п. 71, отличающаяся тем, что библиотека составляет по меньшей мере 75 %, 80 %, 85 % или 90 % эукариотических клеток в указанной емкости.

73. Емкость по п. 71 или п. 72, отличающаяся тем, что указанная емкость представляет собой культуральный флакон, содержащий клетки библиотеки, суспендированные в культуральной среде.

74. Емкость по п. 71 или п. 72, содержащая осадок или концентрированную суспензию эукариотических клеток, содержащих библиотеку.

75. Применение сайт-специфической нуклеазы для адресного расщепления клеточной ДНК при составлении библиотеки эукариотических клеток, содержащих ДНК, кодирующую набор связывающих молекул, в котором опосредованное нуклеазой расщепление ДНК улучшает сайт-специфическое встраивание генов связывающих молекул посредством эндогенных клеточных механизмов репарации ДНК.

76. Применение библиотеки по любому из пп. 66-70 в качестве дисплейной библиотеки для отбора молекул, связывающихся с требуемой мишенью.

77. Применение библиотеки по любому из пп. 66-70 для скрининга клеток, характеризующихся требуемым клеточным фенотипом, отличающееся тем, что фенотип обусловлен экспрессией клеткой связывающей молекулы.

78. Способ превращения популяций одноцепочечного антитела (scFv) в иммуноглобулин (Ig) или антигенсвязывающий фрагмент (Fab), в котором сохраняют попарное объединение оригинальных переменных доменов тяжелой (VH) и легкой (VL) цепей.

79. Способ превращения scFv в Ig или Fab по п. 78, который осуществляется с использованием нерепликативной «миникольцевой ДНК».

80. Способ превращения scFv в Ig или Fab по п. 78, для осуществления которого требуется одна трансформация *E. coli* для непосредственного получения бактериальных трансформантов, несущих плазмиды, кодирующие ДНК Ig или Fab.

81. Способ превращения scFv в Ig или Fab по любому из пп. 78-80, который может быть использован или для моноклонального, олигоклонального, или поликлонального переформатирования клонов.

82. Способ превращения scFv в Ig или Fab по пп. 78-80, который может быть использован для указанного превращения всей получаемой на выходе популяции «*en masse*» (целиком) с помощью любых широко используемых технологий дисплея, включая фаговый, дрожжевой или рибосомный дисплей.

83. Способ, который обеспечивает переформатирование любых двух связанных элементов ДНК с получением вектора, в котором элементы ДНК клонируют под контролем отдельных промоторов или разделяют альтернативными контрольными элементами, но сохраняют изначальное образование пар оснований ДНК.

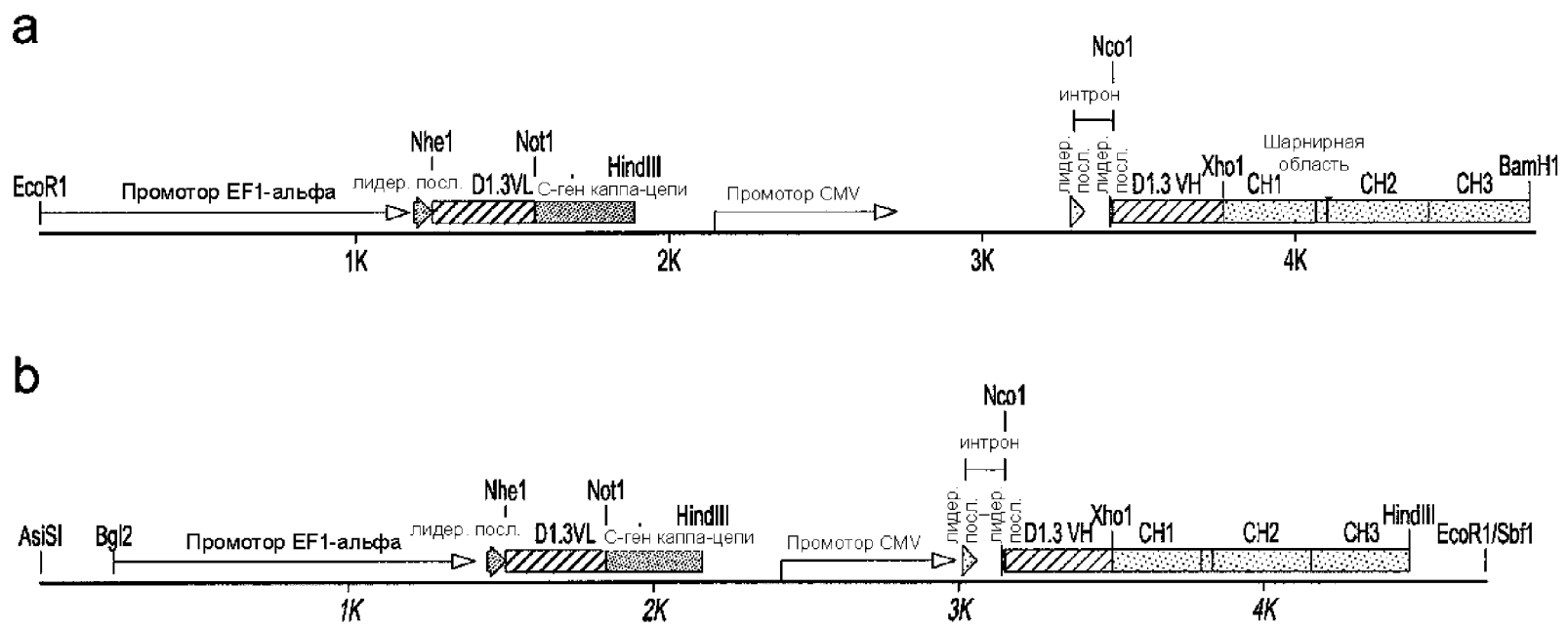
84. Способ изменения формата двух слитых элементов ДНК по п. 83, который осуществляется с использованием нерепликативной «миникольцевой ДНК».

85. Способ изменения формата двух слитых элементов ДНК по п. 83, для осуществления которого требуется одна трансформация *E. coli* для непосредственного получения бактериальных трансформантов, несущих плазмиды, кодирующие элементы ДНК под контролем отдельных промоторов или разделенные альтернативными контрольными элементами.

86. Способ изменения формата двух слитых элементов ДНК по любому из пп. 83-85, который может быть использован или для моноклонального, олигоклонального, или поликлонального переформатирования клонов.

87. Способ изменения формата двух слитых элементов ДНК по пп. 83-85, который может быть использован для превращения всей получаемой на выходе популяции

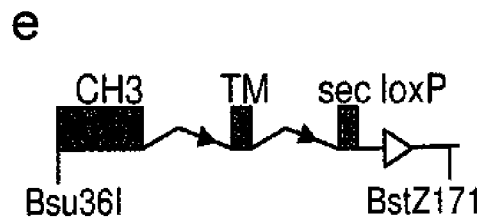
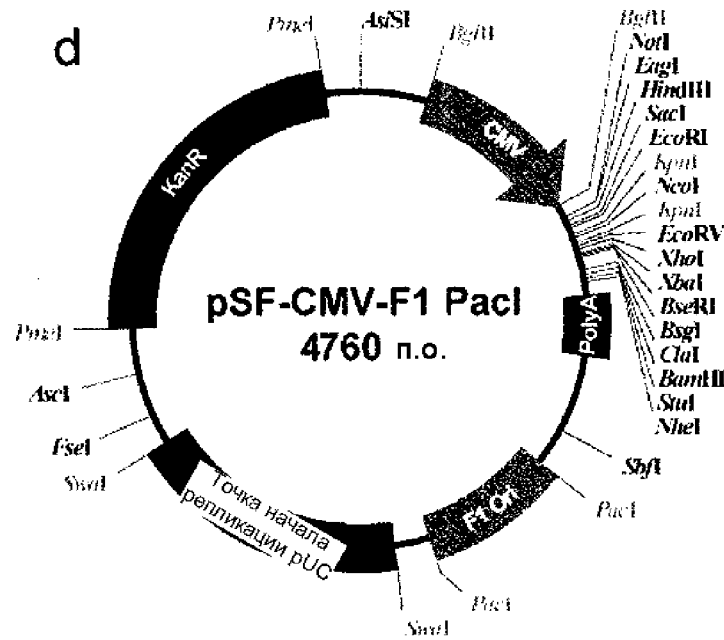
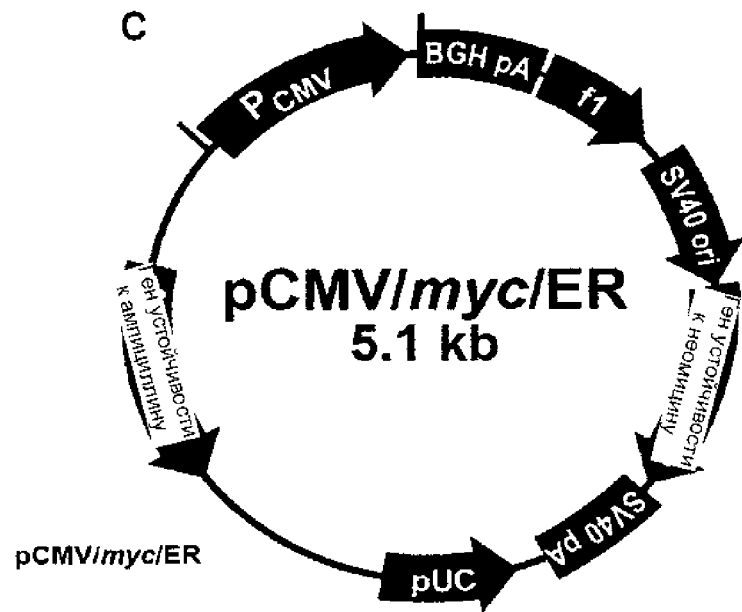
«*en masse*» с помощью любых широко используемых технологий дисплея, включая фаговый, дрожжевой или рибосомный дисплей.



1/56

Фиг. 1





**Фиг. 1**  
(продолжение)

```

                > Промотор рЕF
"Acc65I" >"EcoRI" |
|           | 10 |           |           |           |           |           |           |
GGT ACC GAA TTC CGT GAG GGT CCG GTG CCC GTC AGT GGG CAG ACC GCA CAT CGC CCA CAG
TCC CCG AGA AGT TGG GGG GAG GGG TCG GCA ATT GAA CCG GTC CCT AGA GAA GGT GGC GGG
GGG TAA ACT GGG AAA CTG ATG TCG TGT ACT GGC TCC GCC TTT TTC CCG AGG GTG GGG GAG
AAC CGT ATA TAA CTG CAG TAG TCG CCG TGA ACG TTC TTT TTC GCA ACG GGT TTG CCG CCA
GAA CAC AGG TAA GTG CCG TGT GTG GTT CCC GCG GCC CTG GCC TCT TTA CCG GTT ATG GCC
CTT GCG TGC CTT GAA TTA CTT CCA CCT GSC TCC AGT ACG TGA TTC TTG ATC CCG AGC TGG
AGC CAG GGG CCG GCC TTG CCG TTT AGS ASC CCC TTC GCC TCG TGC TTG AGT TGA GGC CTG
GCC TGG CCG CTG GGG CCG CCG CGT GCG AAT CTG GTG GCA CCT TCG CCG CTG TCT CCG TGC
TTT CGA TAA GTC TCT AGC CAT TTA AAA TTT TTG ATG ACC TGC TGC GAC GCT TTT TTT CTG
GCA AGA TAG TCT TGT AAA TGC GGG CCA GSA TCT GCA CAC TGG TAT TTC GGT TGT TGG GCC
CGC GCC CCG CGA CCG GGC CCG TGC GTC CCA GCG CAC ATG TTC GGC GAG CCG GGG CCT GCG
AGC GCG GCC ACC GAG AAT CCG ACG GGG GTA GTC TCA AGC TGG CCG GCC TGC TCT GGT GCC
TGG CCT CCG GCC CCG GTC TAT CCG CCG GCG CCG CTG GGC GGC AAG GCT GGC CCG GTC GGC ACC
AGT TGC GTG AGC GGA AAG ATG GCC GCT TCC CCG CCC TGC TCC AGG GGG CTC AAA ATG GAG
GAC GCG GCG CTC GGG ASA GCG GGC GGG TGA GTC ACC CAC ACA AAG GAA AAG GGC CTT TCC
GTC CTC AGC CGT CCG TTC ATG TGA CTC CAC GGA GTA CCG GGC GCC GTC CAG GCA CCT CGA
TTA GTT CTG GAG CTT TTG GAG TAC GTC GTC TTT AGG TTG GGG GGA GGG GTT TTA TGC GAT
GSA GTT TCC CCA CAC TGA GTG GGT GSA GAC TGA AGT TAG GCC AGC TTG GCA CTT CAT GTA
ATT CTC GTT GGA ATT TGC CCT TTT TGA GTT TGG ATC TTG GTT CAT TCT CAA GCC TCA GAC
AGT GGT TCA AAG TTT TTT TCT TCC ATT TCA GGT GTC GTG AGA CGT GGCCACC ATG AGG GCC
                                         M R A>
                                         _a_a_a_>
>"NheI" |
1210         1220         1230         1240         :         1250         1260
TGG ATC TTC TTT CTG CTT TGC CTG GCC GGG AGG GGT CTG GCA GCT AGC GAC ATC CAG AIG
W I F F L L C L A G R A L A A S> D I Q M>
_a_a_a_a_a_ "BM40 ЛИД ПОСЛ." а_а_а_а_а_а_а_а_>< "D1.3 VL" _>
1270         1280         1290         1300         1310         1320
ACC CAG AGC CCA AGC AGC CTG AGC GCC AGC GTG GGT GAC AGA GTG ACC ATC ACC TGT AGA
T Q S P S S L S A S V G D R V I I T C R>
_b_b_b_b_b_b_b_b_ "D1.3 VL" _b_b_b_b_b_b_b_b_>
1330         1340         1350         1360         1370         1380
GCC AGC GGT AAC ATC CAC AAC TAC CTG GCT TGG TAC CAG CAG AAG CCA GGT AAG GCT CCA
A S G N I H N Y L A W Y Q Q X P G K A P>
_b_b_b_b_b_b_b_b_ "D1.3 VL" _b_b_b_b_b_b_b_b_>
1390         1400         1410         1420         1430         1440
AAG CTG CTG ATC TAC TAC ACC ACC ACC CTG GCT GAC GGT GTG CCA AGC AGA TTC AGC GGT
K L L I Y Y T T T L A D G V P S R F S G>
_b_b_b_b_b_b_b_b_ "D1.3 VL" _b_b_b_b_b_b_b_b_>
1450         1460         1470         1480         1490         1500
AGC GGT AGC GGT ACC GAC TAC ACC TTC ACC ATC AGC AGC CTC CAG CCA GAG GAC ATC GCC
S G S G T D Y I F I I S S L Q P E D I A>
_b_b_b_b_b_b_b_b_ "D1.3 VL" _b_b_b_b_b_b_b_b_>
1510         1520         1530         1540         1550         1560
ACC TAC TAC TGC CAG CAC TTC TGG AGC ACC CCA AGG ACG TTC GGC CAA GGG ACC AAG GTG
T Y Y C Q H F W S T P R T F G Q G T K V>
_b_b_b_b_b_b_b_b_ "D1.3 VL" _b_b_b_b_b_b_b_b_>

```

Фиг. 2



```

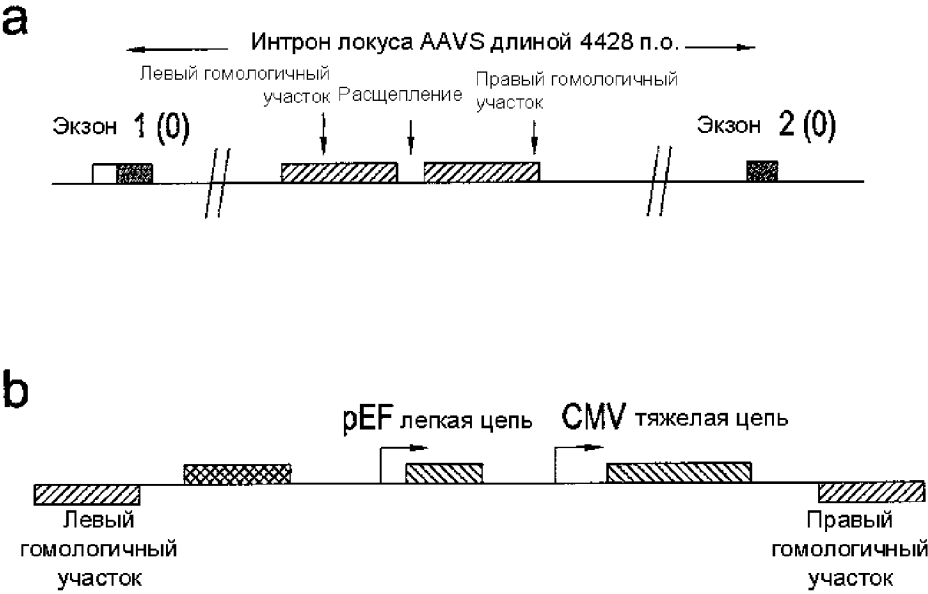
> Pme1_Pml_связка
3250 | 3260 3270 3280 3290 3300
TCC CAG GTC CAA GTT TGT GGA AAT TAA TAC GAC GTG GCC ACC ATG AGT TGG AGC IGT ATC
                                     M S W S C I>
                                     e "ЛИД. ПОСЛЕД. 1" e->
> Интрон
|
3310 3320 3330 3340 3350 3360
ATC CTC TTC TTG GTA GCA ACA GCT ACA GGT AAG GGG TTA ACA GTA GCA GGC TTG AGG ICT
I L F L V A T A T>
e e "ЛИД. ПОСЛЕД. 1" e e e e->
                                     Конец интрона > Nco1
3370 3380 3390 3400 3410 |
GGA CAT ATA TAT GGG TGA CAA TGA CAT CCA CTT TGC CTT TCT CTC CAC A GCC ATG
                                     G A M
                                     ЛИДЕРНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ 2
3420 3430 3440 3450 3460 3470
GCC CAG GTC CAA CTG CAG AGC GGT CCA GGT CTT GTG AGA CCT AGC CAG ACC CTG AGC
A <Q V Q L Q E S G P G L V R P S Q T L S>
g g g g g "D1.3 VH" g g g g g g g g g g->
3480 3490 3500 3510 3520 3530
CTG ACC TGC ACC GTG TCT GGC AGC ACC TTC AGC GCC TAT GGT GTA AAC TGG GTG AGA CAG
L T C T V S G S I F S G Y G V N W V R Q>
g g g g g "D1.3 VH" g g g g g g g g g g->
3540 3550 3560 3570 3580 3590
CCA CCT GGA CGA GGT CTT GAG TGG ATT GGA ATG ATT TGG GGT GAT GGA AAC ACA GAC TAT
P P G R G L E N I G M I W G D G N T D Y>
g g g g g "D1.3 VH" g g g g g g g g g g->
3600 3610 3620 3630 3640 3650
AAT TCA GCT CTC AAA TCC AGA GTG ACA ATG CTG GTA GAC ACC AGC AAG AAC CAG TTC AGC
N S A L K S R V T M L V D T S K N Q F S>
g g g g g "D1.3 VH" g g g g g g g g g g->
3660 3670 3680 3690 3700 3710
CTG AGA CTC AGC AGC GTG ACA GCC GCC GAC ACC GCG GTC TAT TAT TGT GCA AGA GAG AGA
L R L S S V T A A D T A V Y Y C A R E R>
g g g g g "D1.3 VH" g g g g g g g g g g->
                                     Xho1
|
3720 3730 3740 3750 3760 3770
GAT TAT AGG CTT GAC TAC TGG GGT CAA GGC AGC CTC GTC ACA GTC TCG AGT GCC AGC ACC
D Y R L D Y W G Q G S L V T V S S> A S T>
g g g g g "D1.3 VH" g g g g g g g g g g->
3780 3790 3800 3810 3820 3830
AAG GCC CCC AGC GTG TTC CCT CTG GCC CCC TGT AGC AGA AGC ACC AGC GAG AGC ACA GCC
K G P S V F P L A P C S R S T S E S T A>
h h h h h h h h "CH1" h h h h h h h h h h->

```

ФИГ. 2 (продолжение)







Фиг. 3





## ← Экспрессионная КАССЕТА АНТИТЕЛА D1.3 →

6887

Сайт loxP

Bst2171

Asc1

правый\_гомологичный  
участок\_AAVS1

TGGCCCGGGCATG ATAACCTCCGTATAATGTATGTCTATACGGAAGTTATG TATAC GGGCGGCC ACTAGGGACAGGATTGGTGACA

6960

6970

6980

6990

7000

7010

GAA AAG CCC CAT CCT TAG GCC TCC TCC TTC CTA GTC TCC TGA TAT TGG GTC TAA CCC CCA  
 CCT CCT GGT AGG CAG ATT CCT TAT CTG GTG ACA CAC CCC CAT TTC CTG GAG CCA TCT CTC  
 TCC TTG CCA GAA CCT CTA AGG TTT GCT TAC GAT GGA GCC AGA GAG GAT CCT GGG AGG GAG  
 AGC TTG GCA GGG GGT GGG AGG GAA GGG GGG GAT GCG TGA CCT GCC CGG TTC TCA GTG GCC  
 ACC CTG CGC TAC CCT CTC CCA GAA CCT GAG CTG CTC TGA CGC GGC TGT CTG GTG CGT TTC  
 ACT GAT CCT GGT GCT GCA GCT TCC TTA CAC TTC CCA AGA GGA GAA GCA GTT TGG AAA AAC  
 AAA ATC AGA ATA AGT TGG TCC TGA GTT CTA ACT TTG GCT CTT CAC CTT TCT AGT CCC CAA  
 TTT ATA TTG TTC CTC CGT GCG TCA GTT TTA CCT GTG AGA TAA GGC CAG TAG CCA GCC CCG  
 TCC TGG CAG GGC TGT GGT GAG GAG GGG GGT GTC CGT GTG GAA AAC TCC CTT TGT GAG AAT  
 GGT GCG TCC TAG GTG TTC ACC AGG TCG TGG CCG CCT CTA CTC CCT TTC TCT TTC TCC ATC  
 CTT CTT TCC TTA AAG AGT CCC CAG TGC TAT CTG GGA CAT ATT CCT CCG CCC AGA GCA GGG  
 TCC CGC TTC CCT AAG GCC CTG CTC TGG GCT TCT GGG TTT GAG TCC TTG GCA AGC CCA GGA  
 GAG GCG CTC AGG CTT CCC TGT CCC CCT TCC TCG TCC ACC ATC TCA TGC CCC TGS CTC TCC

7740

7750

Конец правого  
гомологичного участка AAVS1

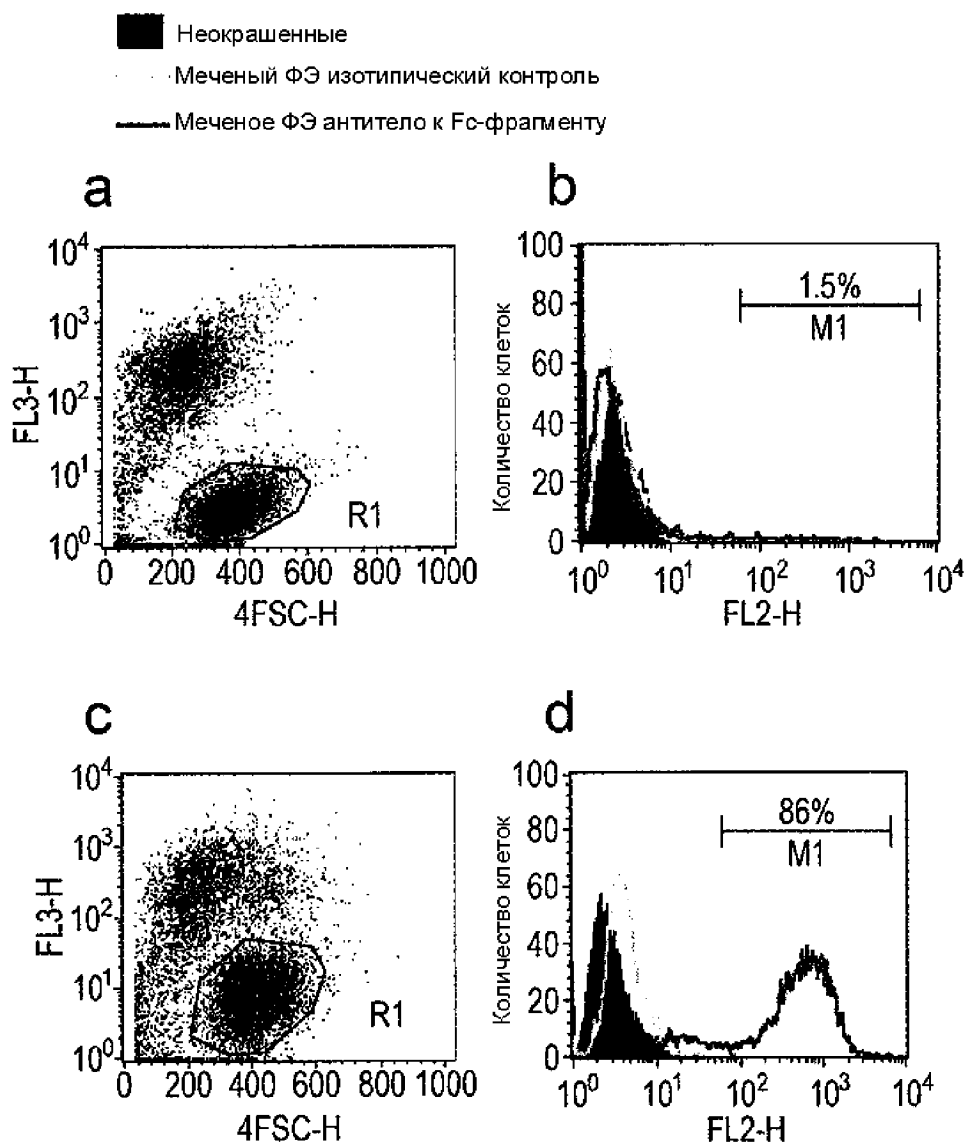
Mlu1

Bst21

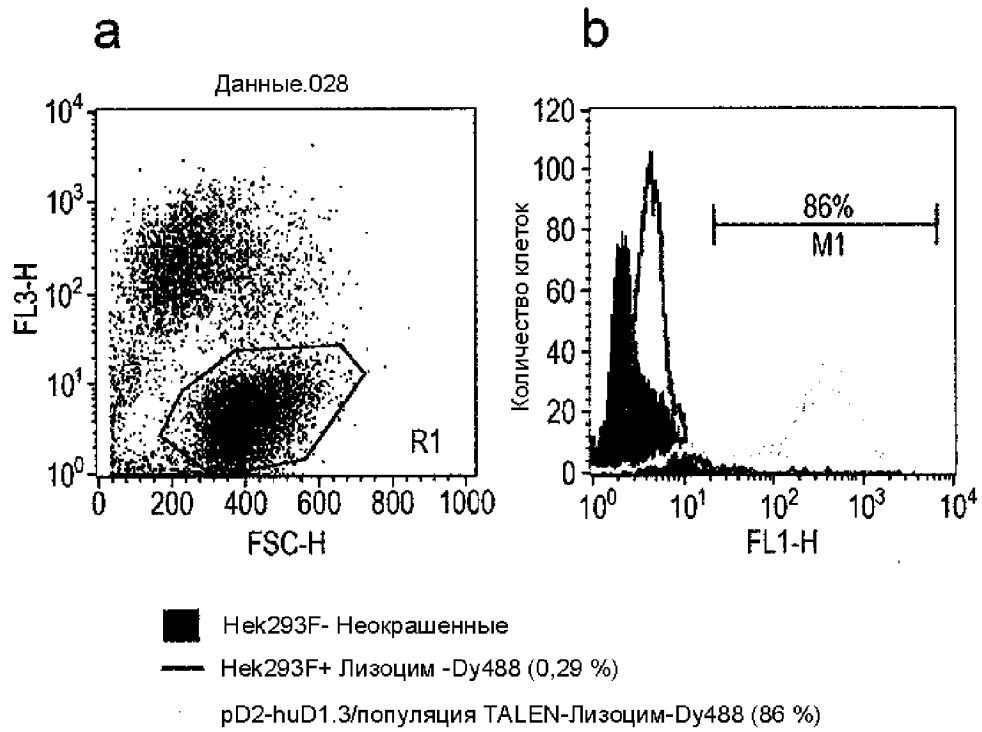
7790

TGC CCC TTC CCT ACA GGG GTT CCT GGC TCT GCT CTG ACG CGT GTATAC TCG ATC TTT CCG

Фиг. 3 (продолжение)



Фиг. 4



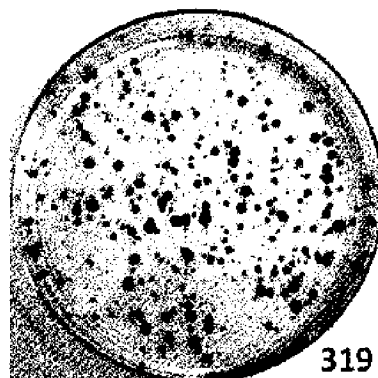
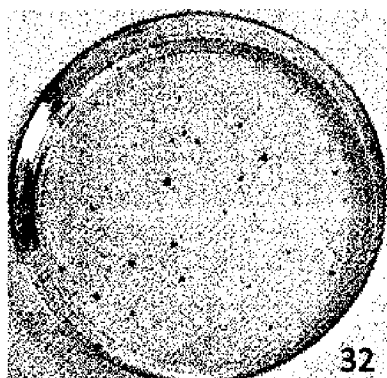
Фиг. 5

рD2 ДНК/  
10<sup>6</sup> клеток

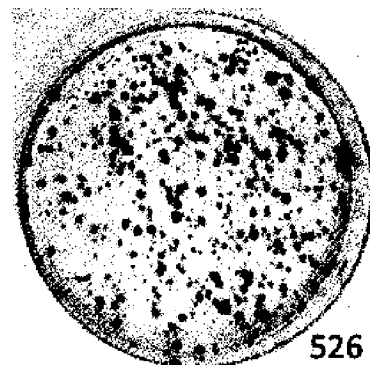
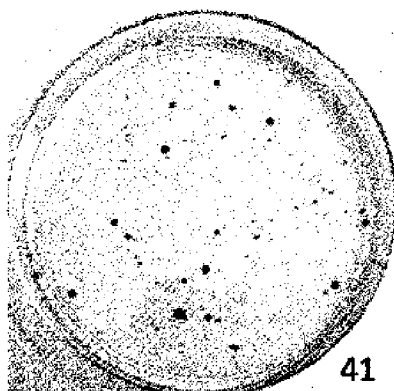
Нуклеазы Talen -AAVS1

Нуклеазы Talen +AAVS1

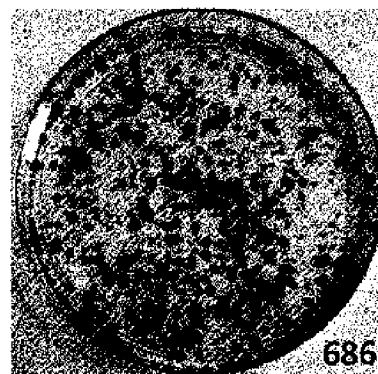
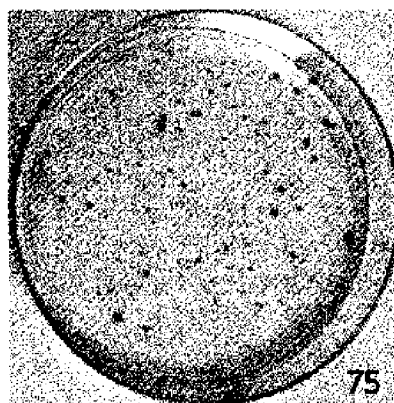
50нг/мл



200нг/мл



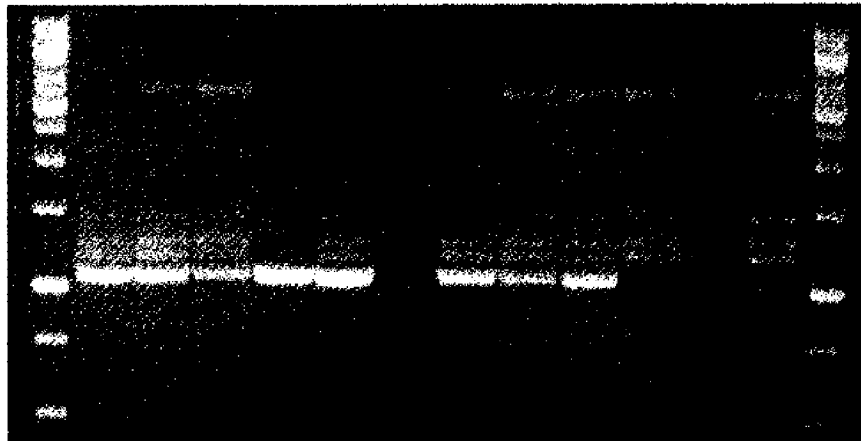
400нг/мл



Фиг. 6

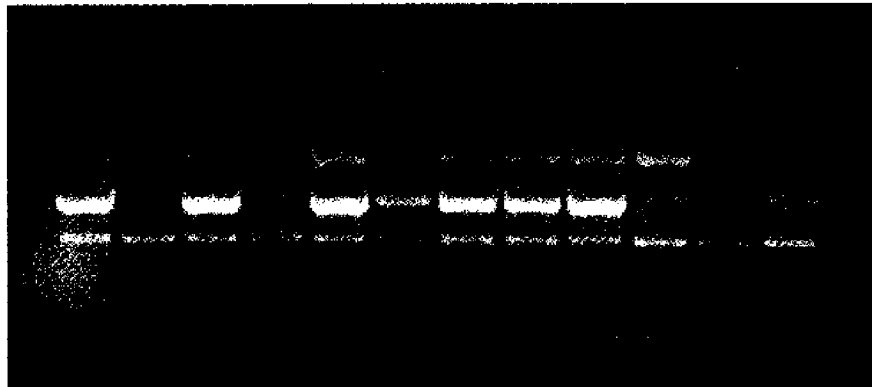
a

1 2 3 4 5 6 7 8 9 10 11 12



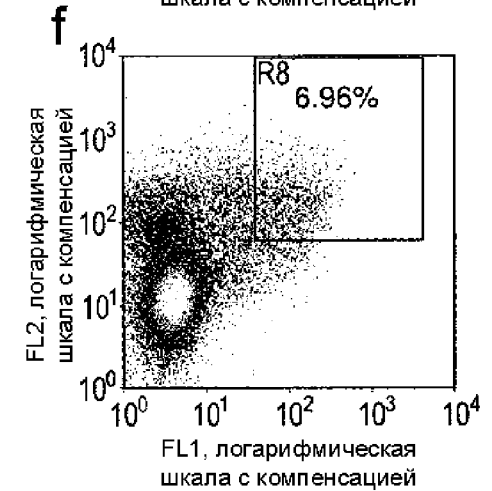
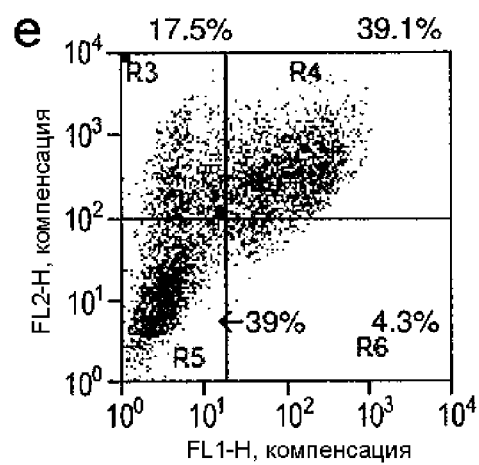
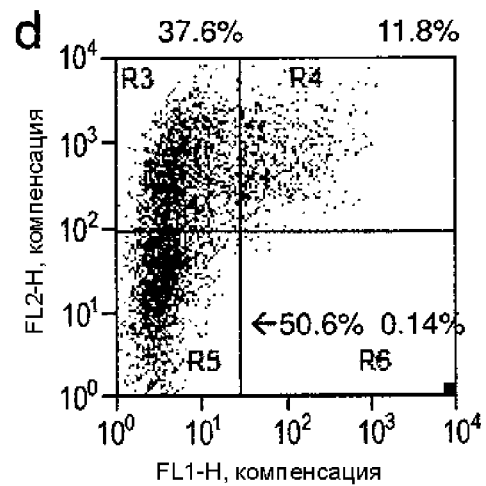
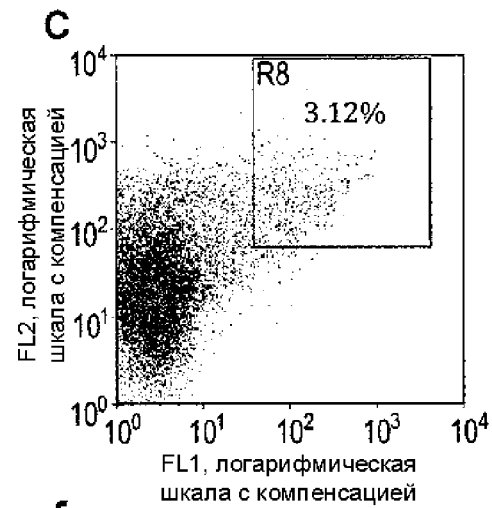
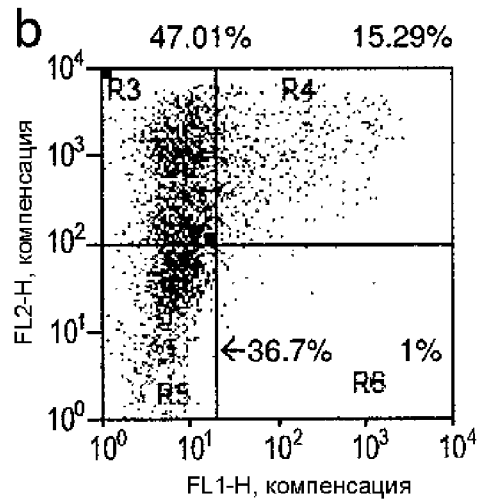
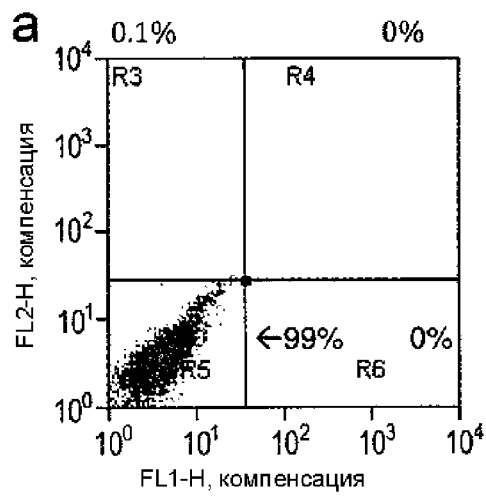
b

1 2 3 4 5 6 7 8 9 10 11 12

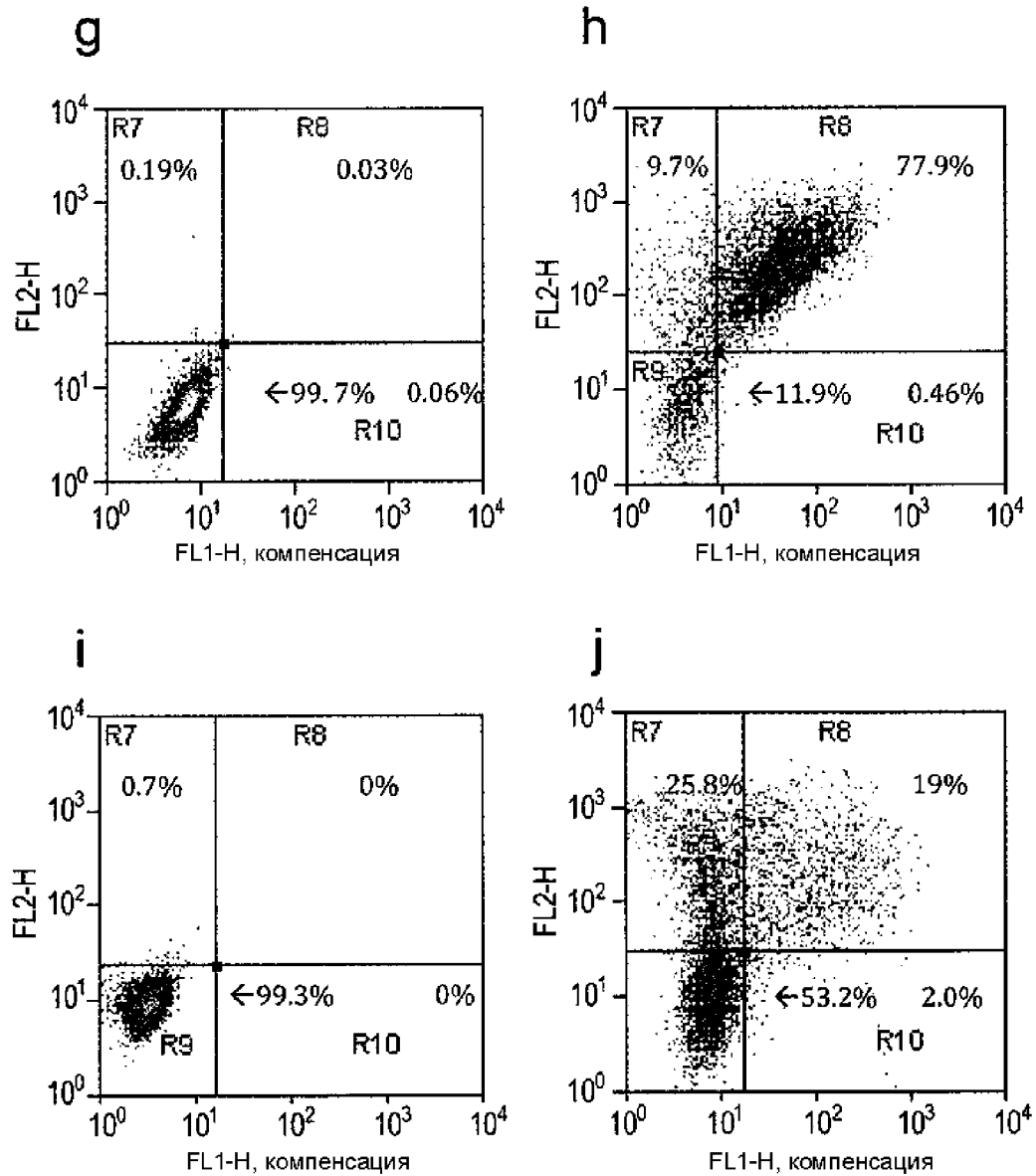


Фиг. 7



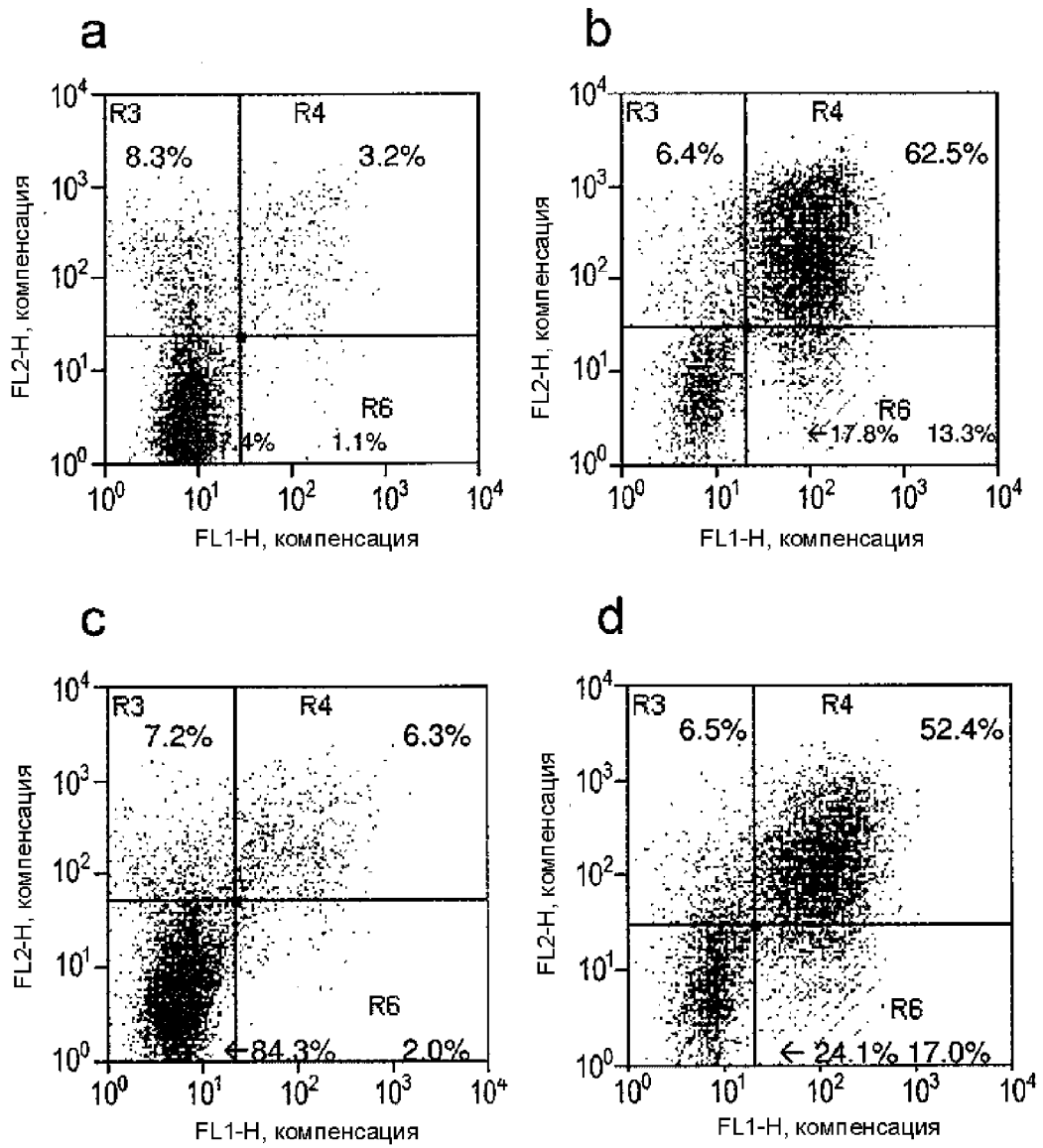


**Фиг.9**

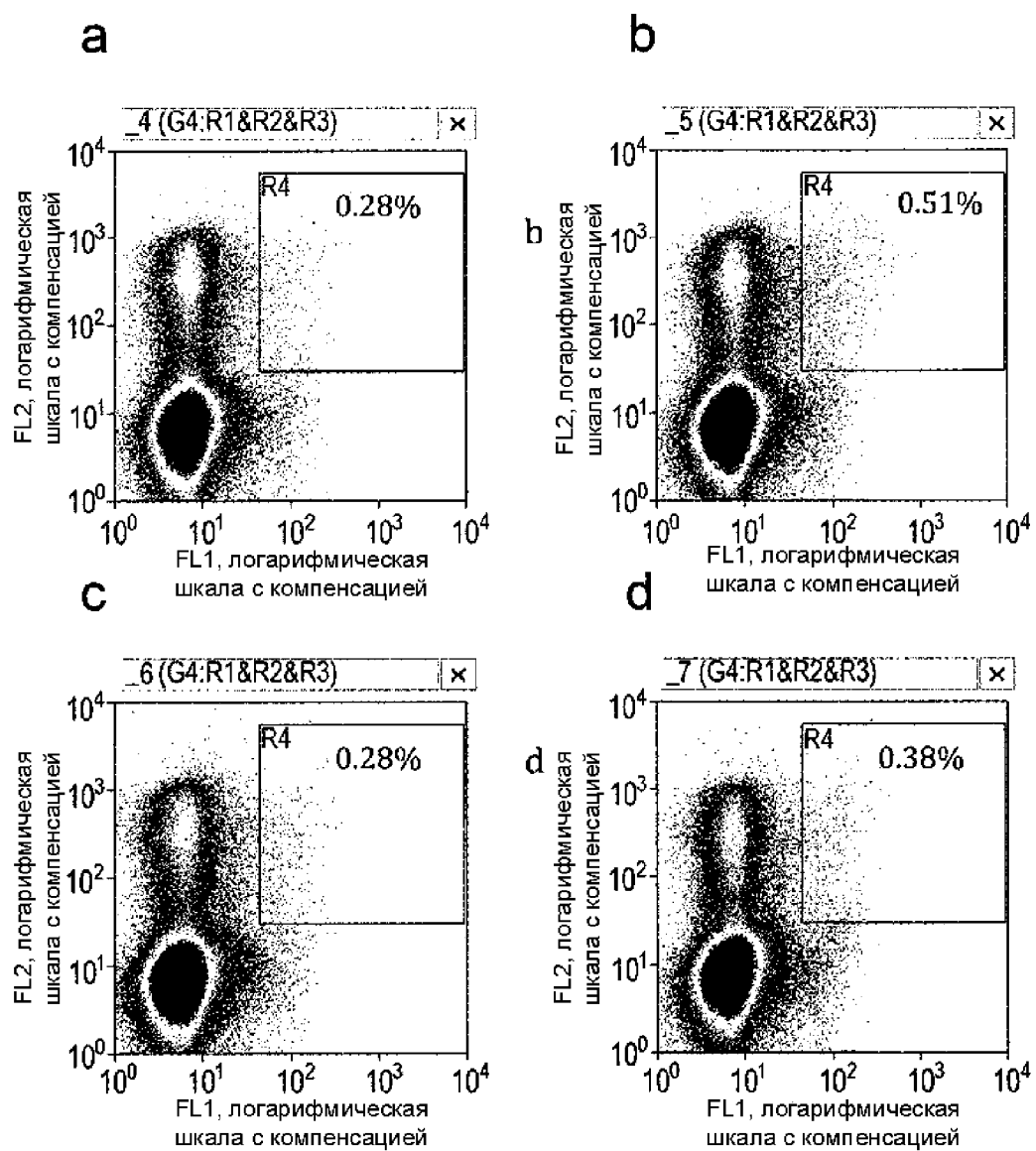


**Фиг. 9**  
(продолжение)



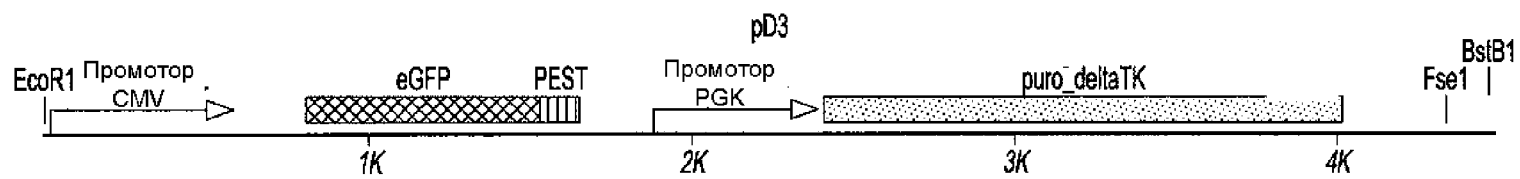


Фиг. 10

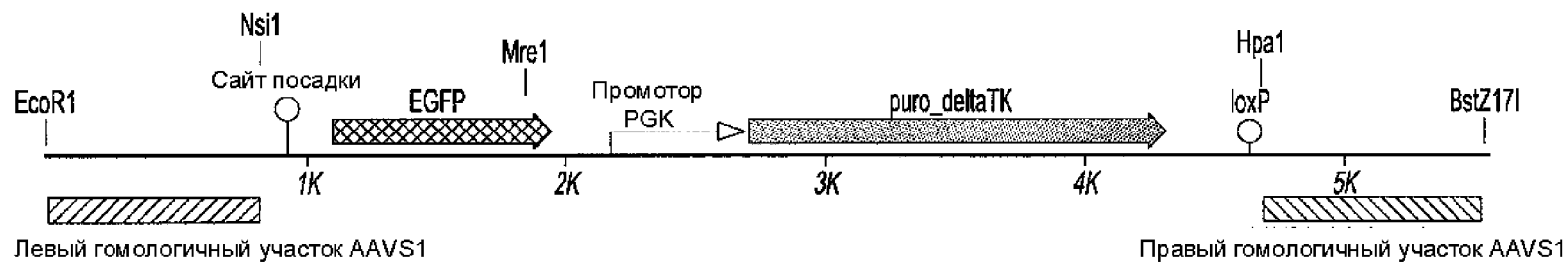


Фиг. 11

a



b



20/56

Фиг. 12

21/56

```

Acc65I"   EcoRI"   Левый гомологичный участок AAVS1 →
|         | 10      20      30      40      50      60
GGT ACC GAA TTC GGC CTT TGC TTT CTC TGA CCT GCA TTC TCT CCC CTG GGC CTG TGC CGC
                                         822
←----- Левый гомологичный ----->>>TTC TGG GTA CTT TTA TCT GTC CCC TCC ACC CCA CAG TGG GGC
участок AAVS1
      NsiI   832      FRT site
AAG ATG CAT   GAAGTTCCATCCGAAGTTCCTATTCTCTAGAAAGTATAGGAACCTC GACC
      Icx_2272      Мегануклеаза_I-SceI
ATAACITCGIATAAAAGTATCCATACGAAGTAT   GCGATCGCTCCGCGC   TAGGGATAACAGGGTAAATAAG
Левый сайт связывания eGFP для TALEN   Спейсер   Правый сайт связывания eGFP для TALEN
TCCACCGGTCCGCA CCATGCTGAGCAAGGG CGAGGAGCTCTTCA   CTTCTGACCTCTCTCTCTCC
> Акцепторный_сайт_сплайсинга
| 1030      1040|      1050      1060      1070      1080
CACAG GGC CTA GAG AGA TCT GGC AGC GGA GGC AGA GGA AGT CTT CTA ACA TGC GGT GAC GTG
      G S G E S R G S L L T C G D V>
      _b_b_b_b_b_b_b_"T2A" _b_b_b_b_b_b_b_>
      1090      1100      1110      1120      1130
GAG GAG AAT CCC GGA CCG ATG GTG AGC AAG GGA GAA GAA CTC TTC ACC GGG GTG
      E E N P G P M V S K G E E L F T G V>
      _b_b_"T2A" _b_II _e_c_c_c_c_c_"EGFP" _c_c_c_c_>
1140      1150      1160      1170      1180      1190
GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTG AAC GGC CAC AAG TTC ACC CTG TCC GGC
      V P I L V E L D G D V N G H K F S V S G>
      _c_c_c_c_c_c_c_"EGFP" _c_c_c_c_c_c_c_>
GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC TGC ACC ACC GGC
AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC CTG ACC TAC GGC CTG CAG TGC TTC
AGC CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC
TAC CTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG
GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG
GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC GTC TAT
ATC ATG GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC
GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC
CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC
1740      1750      1760      1770      1780      1790
AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CAC
      N E K R D H M V L L E F V T A A G I T H>
      _c_c_c_c_c_c_c_"EGFP" _c_c_c_c_c_c_c_>
1800      1810      1820      1830      1840 Krel 1850
GGC ATG GAC GAG CTG TAC AAG AAG CTG AGC CAC GGC TTC CCG CCG GCG GTG GCG GCG CAG
      G M D E L Y K K L S H G F P P A V A A Q>
<-----EGFP----->> II> МУТАНТНАЯ ФОРМА ПОСЛЕД. И REST ОРНИТИН-ДЕКАРБОКСИЛАЗЫ
1860      1870      1880      1890      1900      1910
GAT GAT GGC ACG CTG CCC ATG TCT TGT GCC CAG GAG AGC GGG ATG GAC CGT CAC CCT GCA
      D D G T L P M S C A Q E S G M D R H P A>
<----- МУТАНТНАЯ ФОРМА ПОСЛЕДОВАТЕЛЬНОСТИ REST ОРНИТИН-ДЕКАРБОКСИЛАЗЫ ----->
      >"bGH_polyA"
1920      1930      1940      1950 |      1960      1970      1980
GCC TGT GCT TCT GCT AGG ATC AAT GTG TAT G TGA TCA GCC TCG ACT GTG CCT TCT AGT TGC
      A C A S A R I N V *>
МУТАНТНАЯ ФОРМА ПОСЛЕДОВАТЕЛЬНОСТИ REST ОРНИТИН-ДЕКАРБОКСИЛАЗЫ>
      1990      2000      2010      2020      2030      2040
CAG CCA TGT GTT GTT TGC CCC TCC CCC GTG CCT TCC TTG ACC CTG GAA GGT GCC ACT CCC
      2050      2060      2070      2080      2090      2100
ACT GTC CTT TCC TAA TAA AAT GAG GAA ATT GCA TCG CAT TGT CTG AGT AGG TGT CAT TCT
      2110      2120      2130      2140      2150      2160
AAT CTG GGG GGT GGG CTC GGC CAG GAC AGC AAG GGG GAG GAT TGG GAA GAC AAT AGC AGG

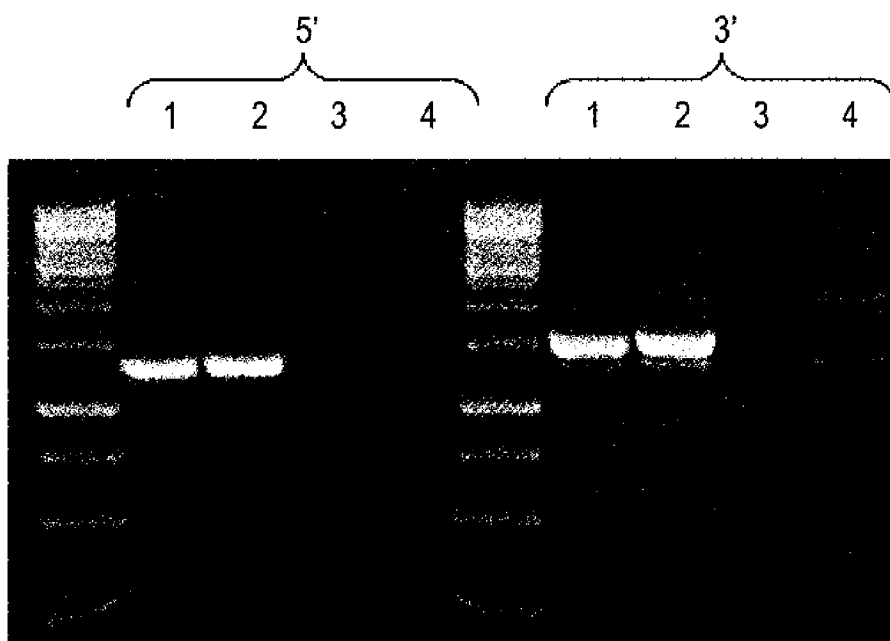
```

Фиг. 13

Промотор\_PGK →

2170 2180 2190 2200 2210 2220  
 CAT GCT GGG GAT GGC CCA ATT CTA CCG GGT AGG GGA GGC GCT TTT CCC AAG GCA GTC TGG  
 AGC ATG CGC TTT AGC AGC CCC GCT GGG CAC TTG GCG CTA CAC AAG TGG CCT CTG GCC TCG  
 CAC ACA TTC CAC ATC CAC CGG TAG GCG CCA ACC GGC TCC GTT CTT TGG TGG CCC CTT CGC  
 GCC ACC TTC TAC TCC TCC CCT AGT CAG GAA GTT CCC CCC CGC CCC GCA GCT CGC GTC GTG  
 CAG GAC GTG ACA AAT GGA AGT AGC ACG TCT CAC TAG TCT CGT GCA GAT GGA CAG CAC CGC  
 TGA GCA ATG GAA GCG GGT AGG CCT TTG GGG CAG CGG CCA ATA GCA GCT TTG CTC CTT CGC  
 TTT CTG GGC TCA GAG GCT GGG AAG GGG TGG GTC CGG GGG CGG GCT CAG GGG CGG GCT CAG  
 GGG CGG GGC GGG CGC CCG AAG GTC CTC CGG AGG CCC GGC ATT CTG CAC GCT TCA AAA GCG  
 2650 2660 12670 2680 2690 2700  
 CAC GTC TGC CGC GCT GTT CTC CTC TTC CTC ATC TCC GGG CCT TTC GAC CTG CAG CCA ACC  
 2710 2720 2730 2740 2750  
 CCA CC ATG GGG ACC GAG TAC AAG CCC ACG GTG GCG CTC GCC ACC CGC GAC GAC GTC CCC  
 M G T E Y K P T V R L A T R D D V P >  
 \_c\_c\_c\_c\_c\_c\_c\_c\_c\_c\_ "PURO\_DELTATK" \_c\_c\_c\_c\_c\_c\_c\_c\_c\_c\_>  
 2760 2770 2780 2790 2800 2810  
 CGG GCC GTA GGC ACC CTC GCC GCC GCG TTC GCC GAC TAC CCC GCC ACG CGC CAC ACC GTC  
 GAC CCG GAC CGC CAC ATC GAG CGG GTC ACC GAG CTG CAA GAA CTC TTC CTC ACG CGC GTC  
 GGG CTC GAC ATC GGC AAG GTG TGG GTC GCG GAC GAC GGC GCC GCG GTG GCG GTC TGG ACC  
 ACG CCG GAG AGC GTC GAA GCG GGG GCG GTG TTC GCC GAG ATC GGC CCG CGC ATG GCC GAG  
 TTG AGC GGT TCC CGG CTG GCC GCG CAG CAA CAG ATG GAA GGC CTC CTG GCG CCG CAC CGG  
 CCC AAG GAG CCC GCG TGG TTC CTG GCC ACC GTC GGC GTC TCG CCC GAC CAC CAG GGC AAG  
 GGT CTG GGC AGC GCC GTC GTG CTC CCC GGA GTG GAG GCG GCC GAG CGC GCC GGG GTG CCC  
 GCC TTC CTG GAG ACC TCC GCG CCC CGC AAC CTC CCC TTC TAC GAG CGC CTC GGC TTC ACC  
 GTC ACC GCC GAC GTC GAG GTG CCC GAA GGA CCG CGC ACC TGG TGC ATG ACC CGC AAG CCC  
 GGT GCC GGA TCC ATG CCC ACG CTA CTG CGG GTT TAT ATA GAC GGT CCT CAC GGG ATG GGG  
 AAA ACC ACC ACC ACG CAA CTG CTG GTG GCC CTG GGT TCG CGC GAC GAT ATC GTC TAC GTA  
 CCC GAG CCG ATG ACT TAC TGG CAG GTG CTG GGG GCT TCC GAG ACA ATC GCG AAC ATC TAC  
 ACC ACA CAA CAC CGC CTC GAC CAG GGT GAG ATA TCG GCC GGG GAC GCG GCG GTG GTA ATG  
 ACA AGC GCC CAG ATA ACA ATG GGC ATG CCT TAT GCC GTG ACC GAC GCC GTT CTG GCT CCT  
 CAT ATC GGG GGG GAG GCT GGG AGC TCA CAT GCC CCG CCC CGG GCC CTC ACC CTC ATC TTC  
 GAC CGC CAT CCC ATC GCC GCC CTC CTG TGC TAC CCG GCC GCG CGA TAC CTT ATG GGC AGC  
 ATG ACC CCC CAG GCC GTG CTG GCG TTC GTG GCC CTC ATC CCG CGC ACC TTG CCC GGC ACA  
 AAC ATC GTG TTG GGG GCC CTT CCG GAG GAC AGA CAC ATC GAC CGC CTG GCC AAA CGC CAG  
 CGC CCC GGC GAG CGG CTT GAC CTG GCT ATG CTG GCC GCG AIT CGC CGC GTT TAC GGG CTG  
 CTT GCC AAT ACG GTG CCG TAT CTG CAG GGC GGC GGG TCG TGG CCG GAG GAT TGG GGA CAG  
 CTT TCG GGG ACG GCC GTG CCG CCC CAG GGT GCC GAG CCC CAG AGC AAC CGC GGC CCA CGA  
 CCC CAT ATC GGG GAC ACG TTA TTT ACC CTG TTT CCG GCC CCC GAG TTG GTC GCC CCC AAC  
 GGC GAC CTG TAC AAC GTG TTT GCC TGG GCC TTG GAC GTC TTG GCC AAA CGC CTC GGT CCC  
 ATG CAC GTC TTT ATC CTG GAT TAC GAC CAA TCG CCC GCC GGC TGC CCG GAC GCC CTG CTG  
 CAA CTT ACC TCC GGG ATG GTC CAG ACC CAC GTC ACC ACC CCC GGC TCC ATA CCG ACG ATC  
 4260 4270 4280 4290 4300 4310 4320  
 TGC GAC CTG GCG CGC ACG TTT GCC CGG GAG ATG GGG GAG GCT AAC TGA G CTC TGG AGC TCG  
 C D L A R T F A R E M G E A N \* >  
 \_c\_c\_c\_c\_c\_c\_c\_c\_c\_c\_ "PURO\_DELTATK" \_c\_c\_c\_c\_c\_c\_c\_c\_c\_c\_>  
 >"bGH\_pA"  
 4330 4340 4350 4360 4370 4380  
 CTG ATC AGC CTC GAC TGT GCC TTC TAG TTG CCA GCC ATC TGT TGT TTG CCC CTC CCC CGT  
 4390 4400 4410 4420 4430 4440  
 GCC TTC CTT GAC CCT GGA AGG TGC CAC TCC CAC TGT CCT TTC CTA ATA AAA TGA GGA AAT  
 4450 4460 4470 4480 4490 4500  
 TGC ATC GCA TTG TCT GAG TAG GTG TCA TTC TAT TCT GGG GGG TGG GGT GGG GCA GGA CAG  
 4510 4520 4530 4540 4550 4560  
 CAA GGG GGA GGA TTG GGA AGA CAA TAG CAG GCA TGC TGG GGA TGC GGT GGG CTC TAT GGC  
 4570 4580 4590 4600 4610 4620  
 TTC TGA GGC GGA AAG AAC CAG CTG GGG CTC GAG ATC CAC TAG TTC TAG CCT CGA GGC TAG  
 HpaI  
 4630 4680  
 AGC GGC CGG CCT ATAACTTCGATATATCTATGCTATACCGAAGTTAT CAGGTAA GTT AAC AGGGCGCGGCC  
 4692 5528 Bstz171  
 ACTAGGGACAGGATTGGTGACA ← Правый\_гомологичный\_участок → TCTGCTCT GACCGGTGATA  
 AAVS1

Фиг. 13 (продолжение)



Фиг. 14

24/56

```

>"Acc65I"  >"EcoR1"  | _____ Левый_гомολогичный_участок AAVS1 _____>>
|         | 10    |         | 20    |         | 30    |         | 40    |         | 50    |         | 60
GGT ACC GAA TTC CTA GCT CTT CCA GCC CCC TGT CAT GGC ATC TTC CAG GGG TCC GAG AGC
          70          80          90          100         110         120
TCA GCT AGT CTT CTT CCT CCA ACC CGG GCC CCT ATG TCC ACT TCA GGA CAG CAT GTT TGC
          130         140         150         160         170         180
TGC CTC CAG GGA TCC TGT GTC CCC GAG CTG GGA CCA CCT TAT ATT CCC AGG GCC GGT TAA
          190         200         210         220         230         240
TGT GGC TCT GGT TCT GGG TAC TTT TAT CTG TCC CCT CCA CCC CAC AGT GGG GC
  
```

```

          NsiI           FRT_site
AAGAIGCAT    GAAGTTCCTATTCCGAAGTTCCTATTCTCTAGAAAAGTATAGGAACITC GACC
  
```

```

Lox 2272           Мегануклеаза_I-SceI
ATAACTTCGTATAAAGTATCCCTATACGAGTTAT GCGATCGCTCCGCCG TAGGGATAACAGGGTAAAT  AAG
  
```

> Левый сайт связывания eGFP TALEN | > Акцепторный сайт\_сплайсинга

```

|         |         |         |         |         |         |
| 370     | 380     | 390     | 400     | 410
TCCACCCTCGGCC ACCATGGT CT TCT GAC CTC TTC TCT TCC TCC CA CAG GGC ATG
                                G M
  
```

```

429         430         440         450         460         470
GCA AAA CCT CTG AGC CAG GAA GAA AGC ACA CTG ATT GAA AGA GCA ACC GCT ACT ATC AAC
A K P L S Q E E S T L I E R A T A T I N>
_a_a_a_a_a_a_a_a_a_ГЕН УСТОЙЧИВОСТИ_а_a_a_a_a_a_a_a_a_a_a_a_>
                        К БЛАСТИЦИДИНУ
  
```

```

480         490         500         510         520         530
AGC ATC CCC ATC TCC GAA GAC TAT TCT GTG GCT AGT GCC GCT CTG TCC AGC GAC GGG AGA
S I P I S E D Y S V A S A A L S S D G R>
_a_a_a_a_a_a_a_a_a_ГЕН УСТОЙЧИВОСТИ_а_a_a_a_a_a_a_a_a_a_a_a_>
                        К БЛАСТИЦИДИНУ
  
```

```

540         550         560         570         580         590
ATC TTC ACC GGT GTG AAC GTC TAC CAC TTT ACA GCC GGA CCA TGC GCA GAG CTG GTG GTC
I F T G V N V Y H F T G G P C A E L V V>
_a_a_a_a_a_a_a_a_a_ГЕН УСТОЙЧИВОСТИ_а_a_a_a_a_a_a_a_a_a_a_a_>
                        К БЛАСТИЦИДИНУ
  
```

```

600         610         620         630         640         650
CTG GGG ACT GCA GCC GCT GCA GCC GCT GGT AAT CTG ACC TGT ATC GTG GCC ATT GGC AAC
L G T A A A A A A A G N L T C I V A I G N>
_a_a_a_a_a_a_a_a_a_ГЕН УСТОЙЧИВОСТИ_а_a_a_a_a_a_a_a_a_a_a_a_>
                        К БЛАСТИЦИДИНУ
  
```

```

660         670         680         690         700         710
GAA AAT AGG GGC ATC CTG TCC CCA TGC GGC AGG TGT CGG CAG GTG CTG CTG GAT CTG CAT
E N R G I L S P C G R C R Q V L L D L H>
_a_a_a_a_a_a_a_a_a_ГЕН УСТОЙЧИВОСТИ_а_a_a_a_a_a_a_a_a_a_a_a_>
                        К БЛАСТИЦИДИНУ
  
```

```

720         730         740         750         760         770
CCT GGC ATC AAG GCA ATT GTC AAA GAC TCT GAT GGA CAG CCT ACC GCC GTC GGT ATC CGT
P G I K A I V K D S D G Q P T A V G I R>
_a_a_a_a_a_a_a_a_a_ГЕН УСТОЙЧИВОСТИ_а_a_a_a_a_a_a_a_a_a_a_a_>
                        К БЛАСТИЦИДИНУ
  
```

Фиг. 15

> Поли(A)-последовательность\_SV40

```

780          790          800          810          820          830 |          840
GAA CTG CTG CCT AGC GGC TAT GTC TGG GAG GGA TAA TGA A TGA GCT TGG CTT CGA AAT GAC
E   L   L   P   S   G   Y   V   W   E   G   *   * >
_a_a_ ГЕН УСТОЙЧИВОСТИ К БЛАСТИЦИДИНУ _a_a_>

          850          860          870          880          890          900
CGA CCA AGC GAC GCC CAA CCT GCC ATC ACG AGA TTT CGA TTC CAC CGC CGC CTT CTA TGA

          910          920          930          940          950          960
AAG GTT GGG CTT CCG AAT CGT TTT CCG GGA CGC CGG CTG GAT GAT CCT CCA GCG CGG GGA

          970          980          990          1000          1010          1020
TCT CAT GCT GGA GTT CTT CGC CCA CCC CAA CTT GTT TAT TGC AGC TTA TAA TGG TTA CAA

          1030          1040          1050          1060          1070          1080
ATA AAG CAA TAG CAT CAC AAA TTT CAC AAA TAA AGC ATT TTT TTC ACT GCA TTC TAG TTG
    
```

**Pac1**      → ←КАССЕТА ДЛЯ ЭКСПРЕССИИ АНТИТЕЛА D1.3→

TGG TTT AAT TAA CAA TTC

Конец Поли(A)-последовательности      6443 Рекомбинация с помощью Lox P      **Bstz171**

CAGGGCATGCTGGGGA      TGGCCCGGGCATG      АТААСТТGGTATAATGTATGCTATACGAAGTTATG TATAC

Правый сайт связывания

Asci      eGFP для TALEN      6510

GGCGGGCC GAGCAAGGG CGAGGAGCTGTTC      CTTCTGACCTCTCTCTCTCTCCGCCCTG

Акцепторный сайт_сплайсинга	>"T2A"	Начало T2a/гомологичный участок GFP	(6559-7332)
6550	6559	6570	6590      6600
<pre> AGC CTA GAG AGA TCT GGC AGC GGA GAG GGC AGA GGA AGT CTT CTA ACA TGC GGT GAC GTG GAG GAG AAT CCC GGA CCG TGA GTG AGC AAG GGA GAA GAA CTC TTC ACC GGG GTG GTG CCC ATC CTG GTC GAG CTG GAC GGC GAC GTG AAC GGC CAC AAG TTC AGC GTG TCC GGC GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTS ACC CTG AAG TTC ATC TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC CTG ACC TAC GGC GTG CAG TGC TTC AGC CGC TAC CCC GAC CAC ATG AAG CAG CAC GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG GTG AAG TTC GAG GGC CAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC ATC GAC TTC AAG GAG GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC AAC TAC AAC AGC CAC AAC CTC TAT ATC ATG GCC GAC AAG CAG AAG AAC GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC GTG CAG CTC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC ACC CAG TCC GCC CTG AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG GAG TTC GTG ACC GCC GCC GGG ATC ACT CAC GGC           7330      Mini1      <b>Bstz171</b> ATG GAC GAG CCT <u>GACGGGT</u>      <u>GTATAC</u>     </pre>			

**Фиг. 15** (продолжение)

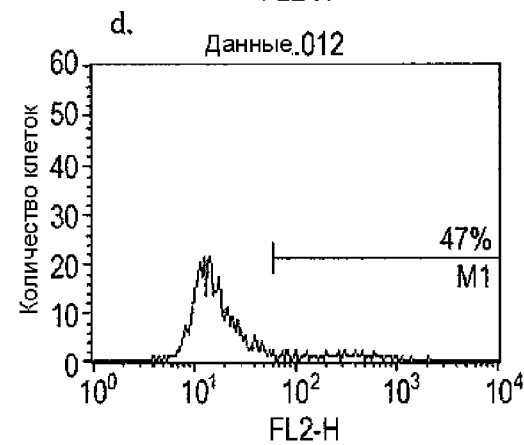
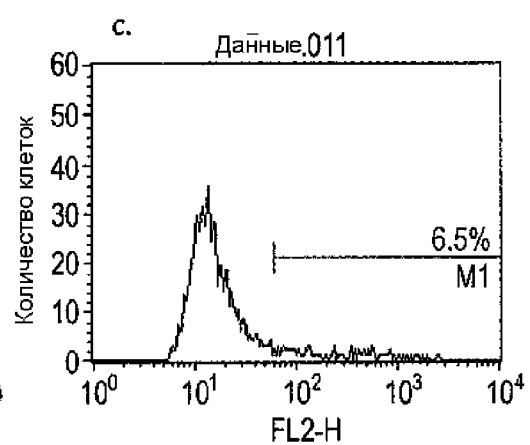
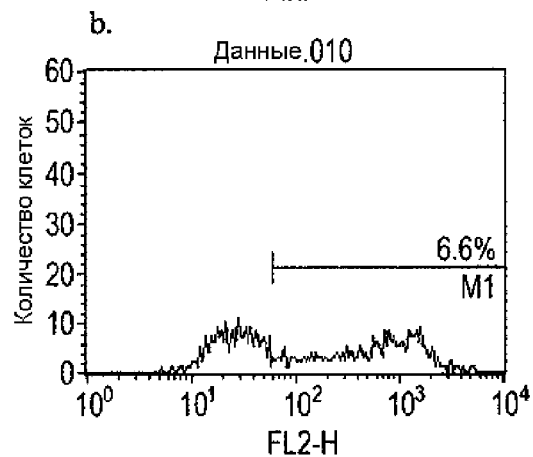
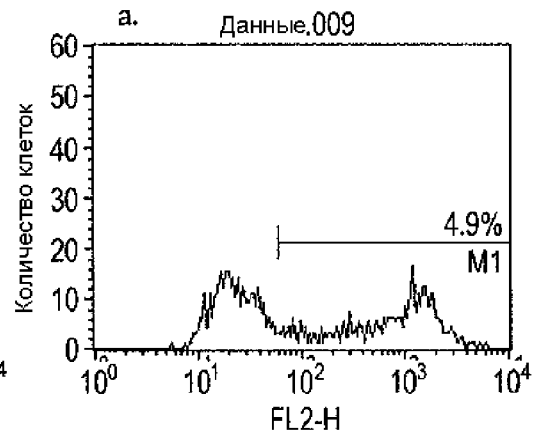
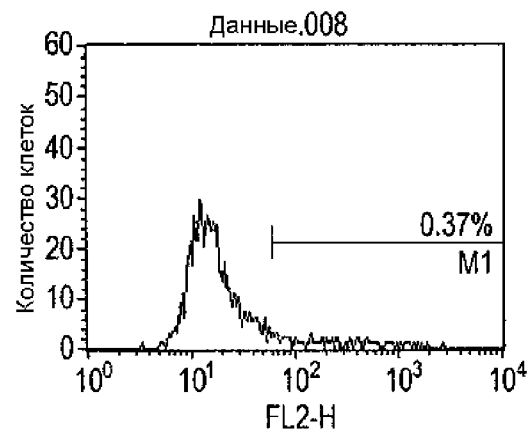
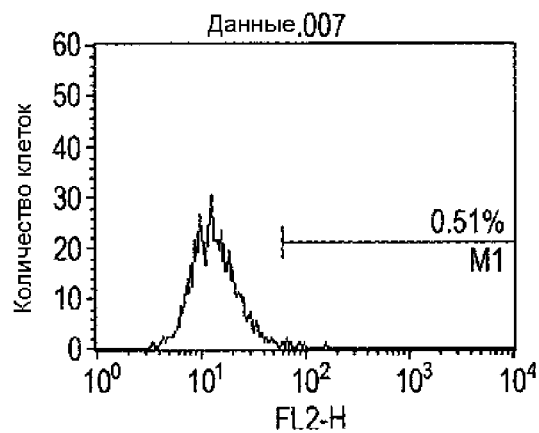


```

NcoI  10      20      30      40      50      60
CC ATG GGC TAT CCT TAC GAT GTC CCT GAT TAC GCC AAC AGT CCT GGT ATC CCT GGT ATG
M  G  Y  P  Y  D  V  P  D  Y  A  N  S  P  G  I  P  G  M>
      _b_b_b_ "HA TAG" _b_b_b_>
      70      80      90      100     110     120
GGI CCT AAA AAG AAG CGA AAA GTG GGT AGA CTG GAA CCC GGC ATG AAG AAC ATT AAG AAA
G  P  K  K  K  R  K  V  G  R  L  E  P  G  M  K  N  I  K  K>
      _c_c_ "NLS" _c_c_c_>      _d_d_ "I-SCE1" _d_d_>
      130     140     150     160     170     180
AAT CAG GTG ATG AAC CTG GGA CCT AAT TCC AAG CTG CTG AAA GAG TAC AAG TCT CAG CTG
N  Q  V  M  N  L  G  P  N  S  K  L  L  K  E  Y  K  S  Q  L>
      _d_d_d_d_d_d_d_d_ "I-SCE1" _d_d_d_d_d_d_d_d_>
      190     200     210     220     230     240
ATC GAA CTG AAC ATT GAG CAG TTT GAA GCA GGG ATC GGT CTG ATT CTG GGG GAC GCC TAC
I  E  L  N  I  E  Q  F  E  A  G  I  G  L  I  L  G  D  A  Y>
      _d_d_d_d_d_d_d_d_ "I-SCE1" _d_d_d_d_d_d_d_d_>
      250     260     270     280     290     300
ATC CGG AGC AGG GAT CAG GGC AAG ACT TAT TGC ATG CAG TTC GAA TGG AAG AAT AAG GCC
I  R  S  R  D  E  G  K  T  Y  C  M  Q  F  E  W  K  N  X  A>
      _d_d_d_d_d_d_d_d_ "I-SCE1" _d_d_d_d_d_d_d_d_>
      310     320     330     340     350     360
TAC ATG GAC CAC GTG TGT CTG CTG TAT GAT CAG TGG GTC CTG TCT CCC CCT CAC AAG AAA
Y  M  D  H  V  C  L  L  Y  D  Q  W  V  L  S  P  P  H  K  K>
      _d_d_d_d_d_d_d_d_ "I-SCE1" _d_d_d_d_d_d_d_d_>
      370     380     390     400     410     420
GAG AGA GTG AAC CAT CTG GGC AAT CTG GTC ATT ACT TGG GGA GCA CAG ACC TTC AAG CAT
E  R  V  N  H  L  G  N  L  V  I  T  W  G  A  Q  T  F  K  H>
      _d_d_d_d_d_d_d_d_ "I-SCE1" _d_d_d_d_d_d_d_d_>
      430     440     450     460     470     480
CAG GCC TTT AAC AAA CTG GCI AAC CTG TTC ATC GTG AAC AAC AAG AAA ACC ATC CCT AAC
Q  A  F  N  K  L  A  N  L  F  I  V  N  H  K  K  T  I  P  N>
      _d_d_d_d_d_d_d_d_ "I-SCE1" _d_d_d_d_d_d_d_d_>
      490     500     510     520     530     540
AAT CTG GTC GAA AAC TAC CTG ACA CCA ATG AGT CTG GCC TAT TGG TTC ATG GAC GAT GGC
N  L  V  E  N  Y  L  T  P  M  S  L  A  Y  W  F  M  D  D  G>
      _d_d_d_d_d_d_d_d_ "I-SCE1" _d_d_d_d_d_d_d_d_>
      550     560     570     580     590     600
GGA AAA TGG GAC TAC AAC AAG AAC AGC ACA AAC AAA AGC ATC GTG CTG AAT ACC CAG CCC
G  K  W  D  Y  N  K  N  S  T  N  K  S  I  V  L  N  T  Q  S>
      _d_d_d_d_d_d_d_d_ "I-SCE1" _d_d_d_d_d_d_d_d_>
      610     620     630     640     650     660
TTC ACA TTT GAG GAA GTG CAG TAT CTG GTC AAG GGC CTG CGG AAC AAA TTC CAG CTG AAC
F  T  F  E  E  V  E  Y  L  V  K  G  L  R  N  K  F  Q  L  N>
      _d_d_d_d_d_d_d_d_ "I-SCE1" _d_d_d_d_d_d_d_d_>
      670     680     690     700     710     720
IGC TAC GTG AAG ATC AAC AAG AAC AAG CCA ATC ATC TAC ATC GAT TCT ATG AGT TAC CTG
C  Y  V  K  I  N  K  N  K  P  I  I  Y  I  D  S  M  S  Y  L>
      _d_d_d_d_d_d_d_d_ "I-SCE1" _d_d_d_d_d_d_d_d_>
      730     740     750     760     770     780
ATC TTT TAT AAC CTG ATT AAG CCA TAC CTG ATC CCC CAG ATG ATG TAT AAA CTG CCC AAT
I  F  Y  N  L  I  K  P  Y  L  I  P  Q  M  M  Y  K  L  P  N>
      _d_d_d_d_d_d_d_d_ "I-SCE1" _d_d_d_d_d_d_d_d_>
      790     800
ACA ATC AGC TCC GAG ACT TTC CTG AAG GICTAGA
I  I  S  S  E  T  F  L  X  V>
      _d_d_d_ "I-SCE1" _d_d_d_>

```

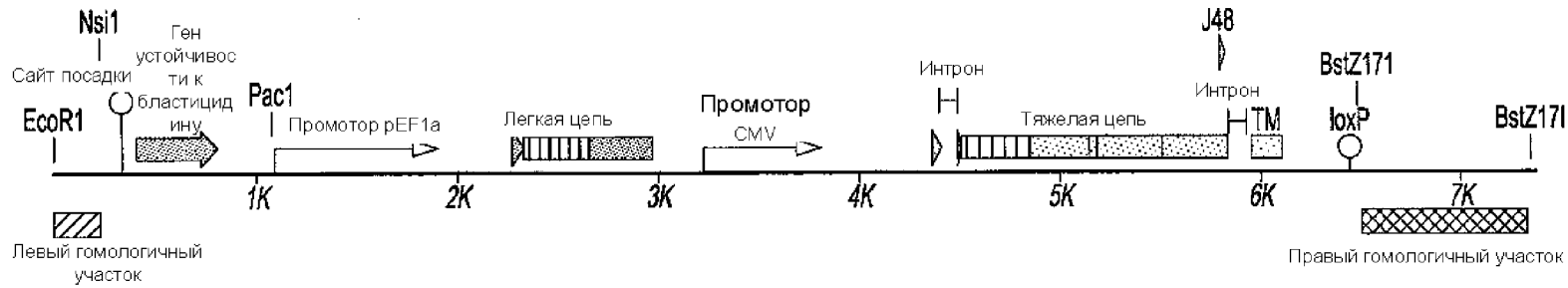
Фиг. 16



27/56

ФИГ. 17

**a**



**b**

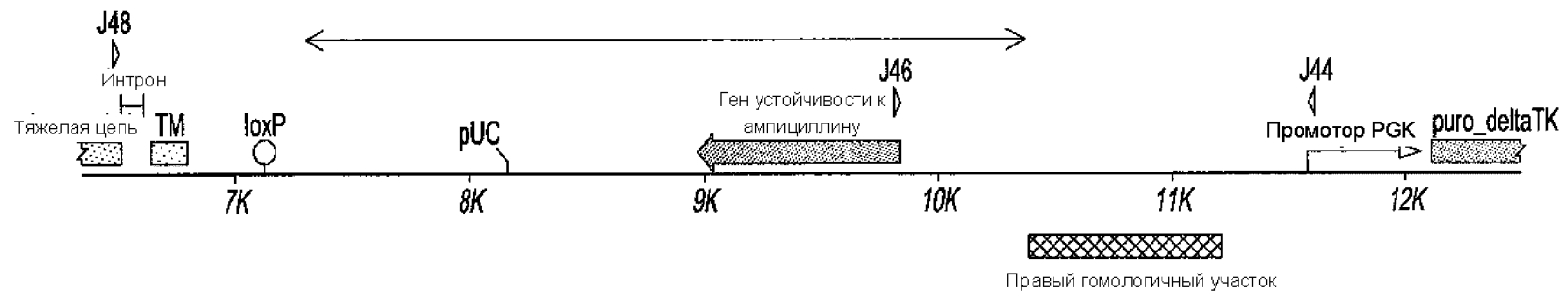


**Фиг. 18**

c



d

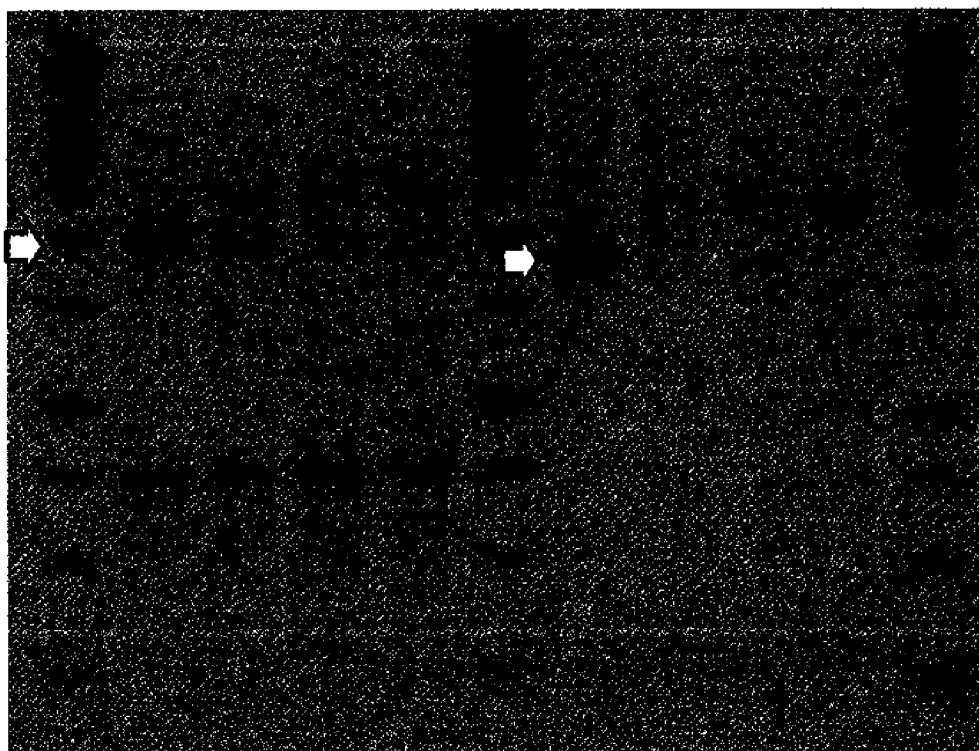


29/56

**Фиг. 18**  
(продолжение)

e

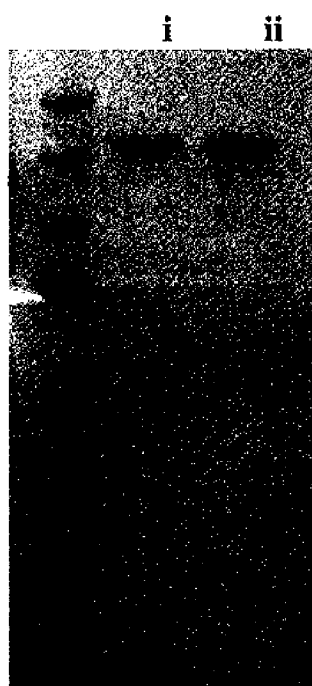
i    ii    iii    iv            v    vi    vii    viii



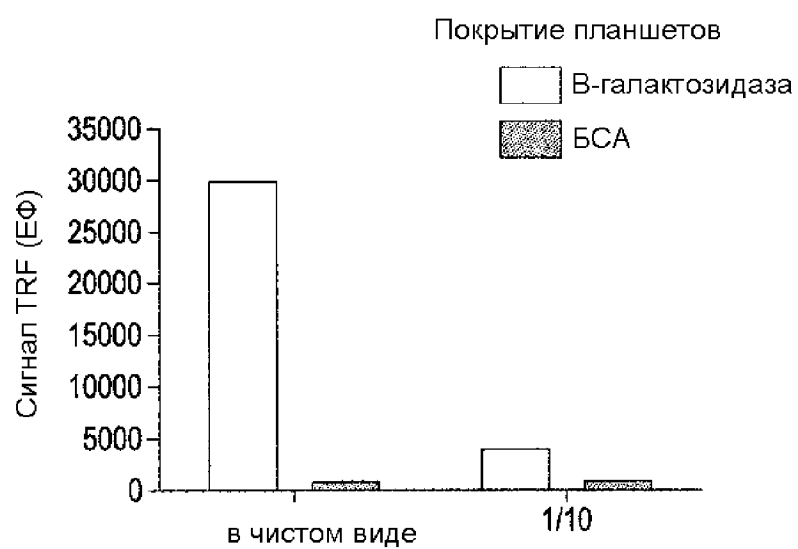
**Фиг. 18**  
(продолжение)

31/56

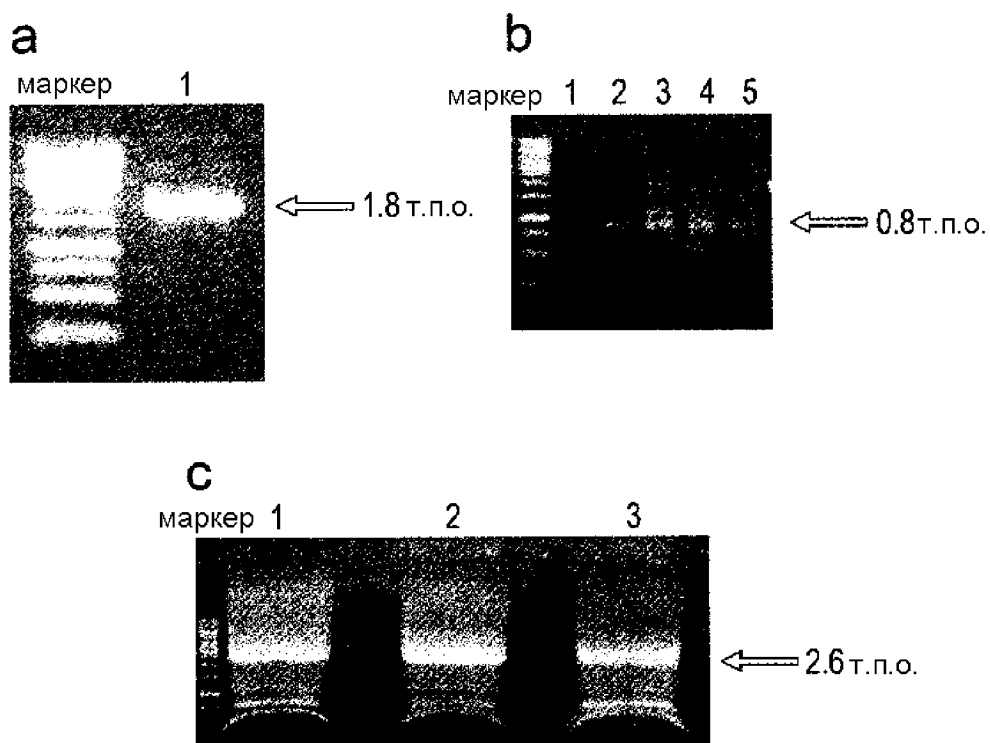
a



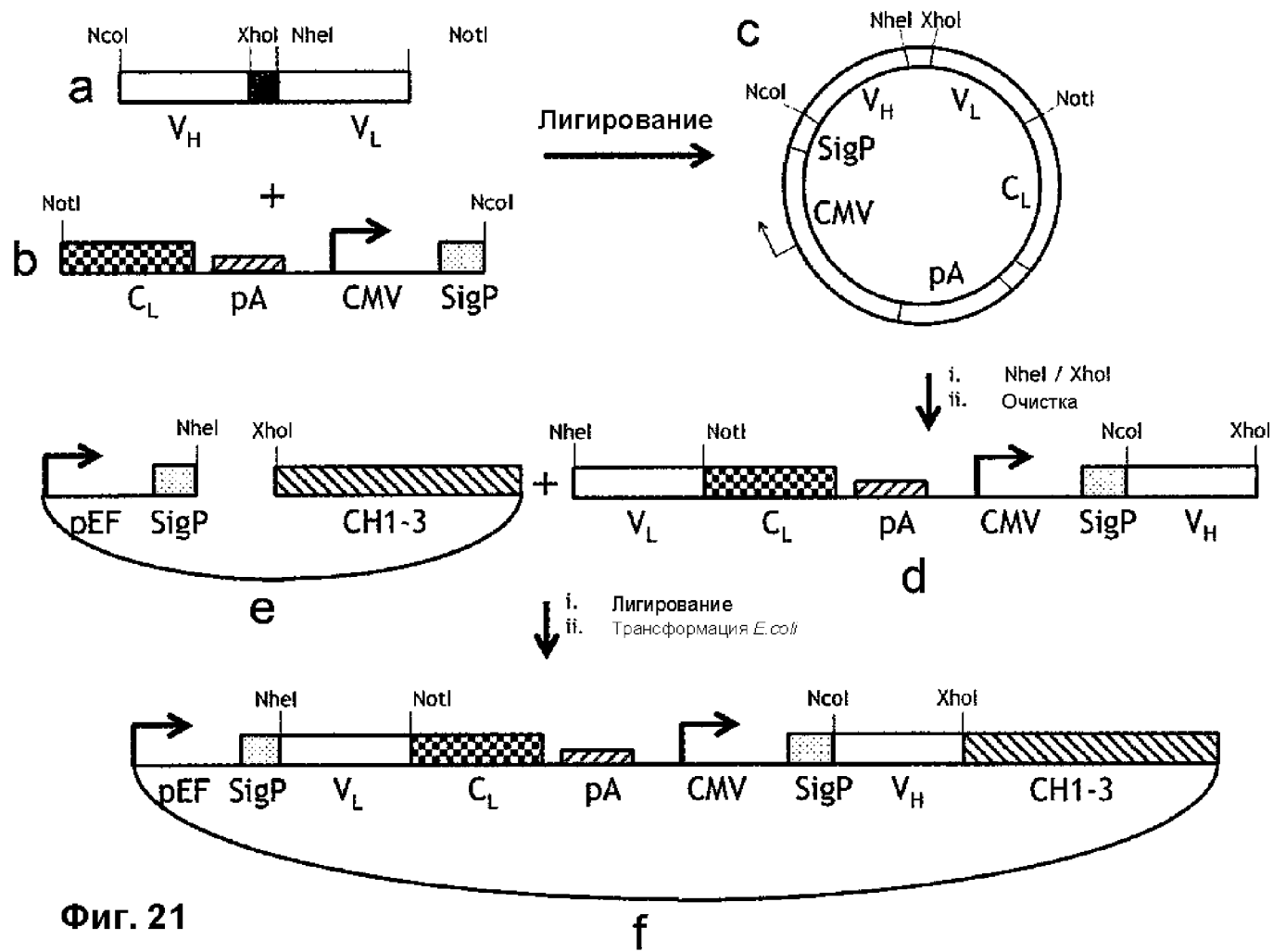
b



Фиг. 19



Фиг. 20

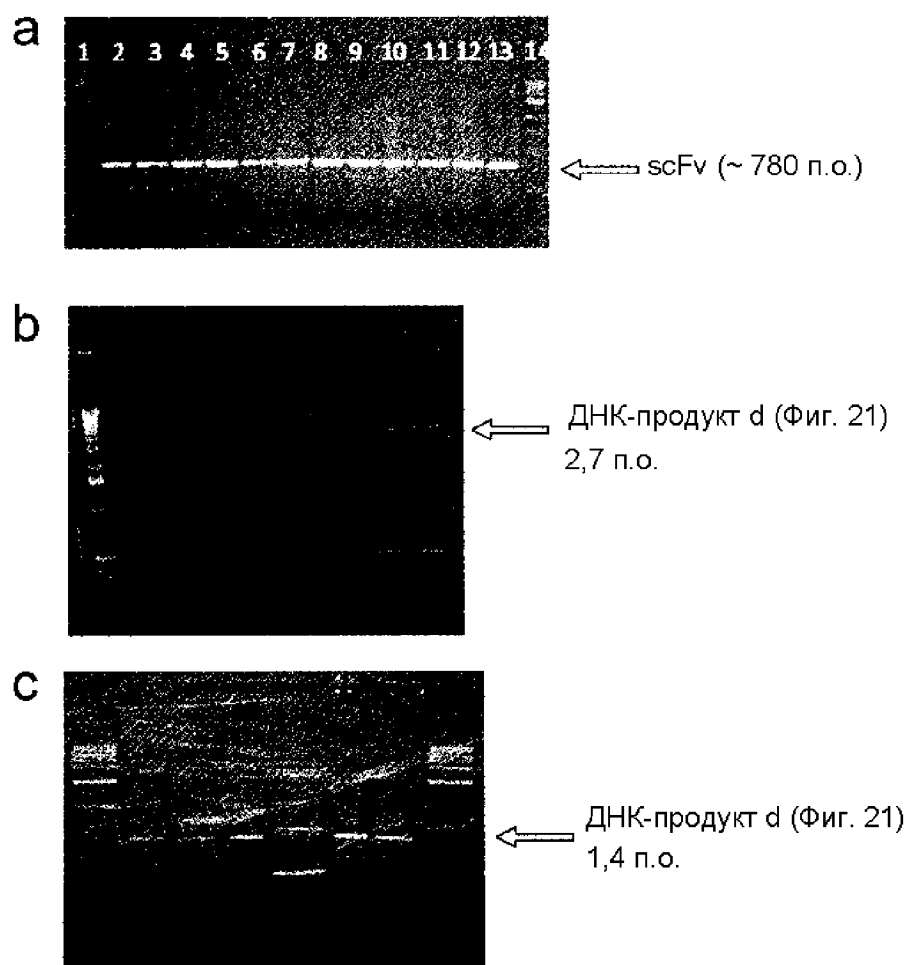


33/56

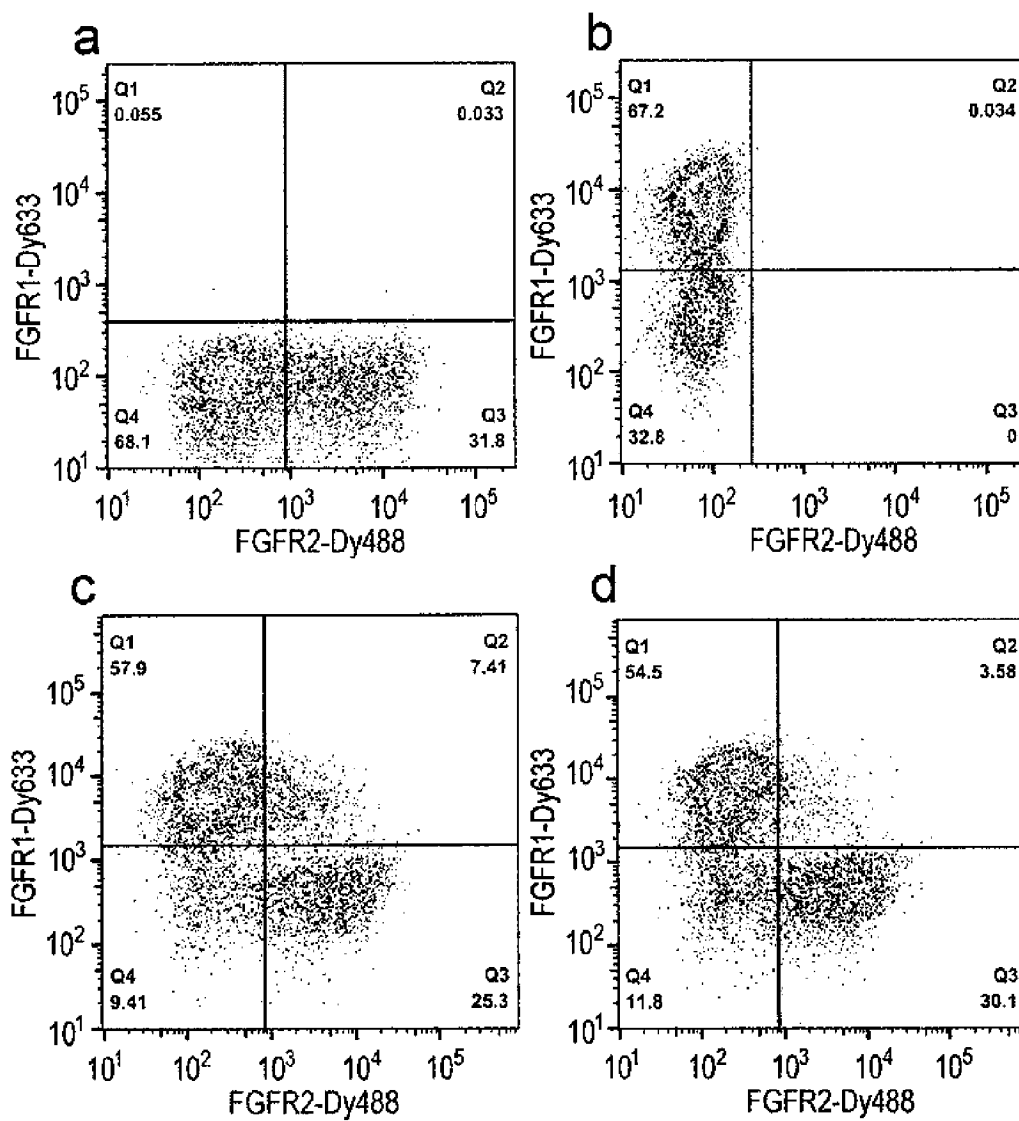
Фиг. 21

f

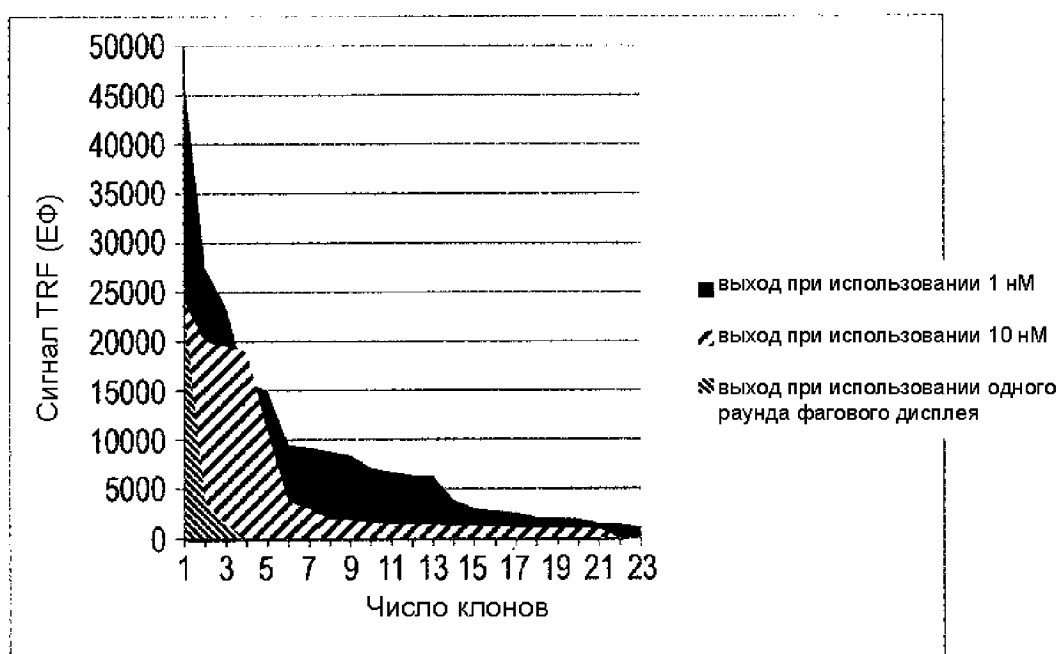




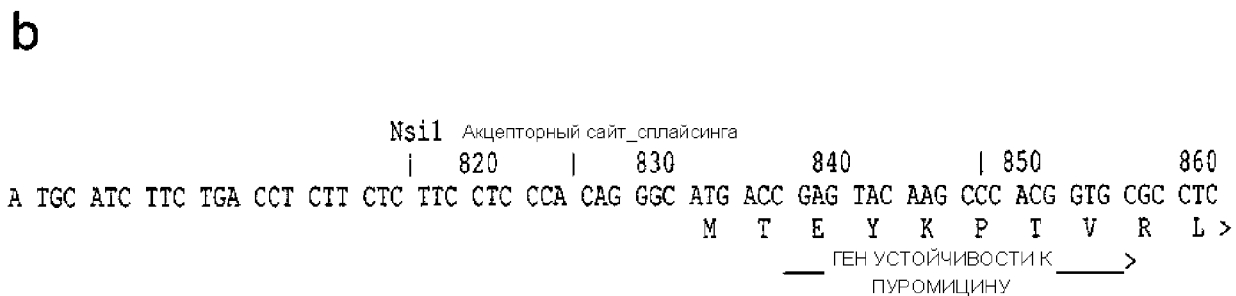
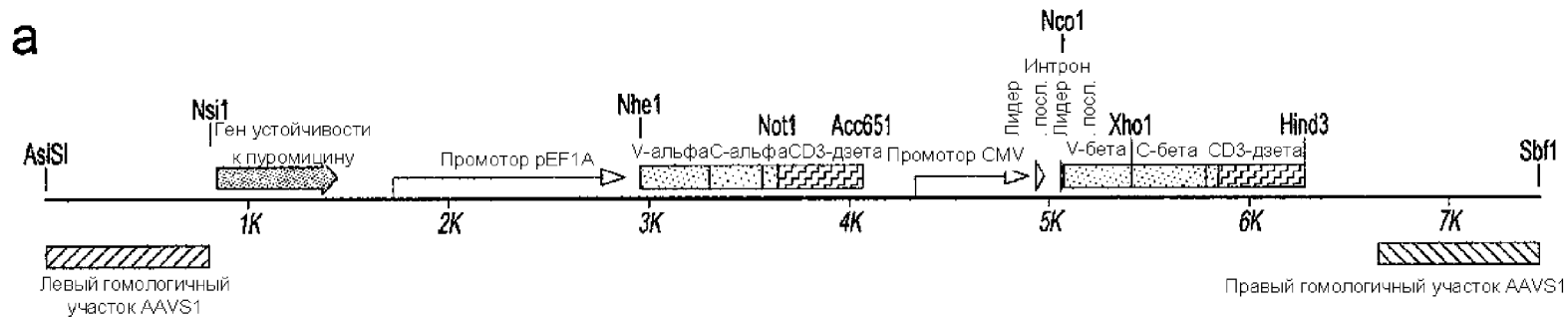
ФИГ. 22



Фиг. 23



Фиг. 24



**Фиг. 25**

C

gctagcaagcaggaagtgactcagatcccagccgctctgagcgtgcctgagggagaaaac  
**Nhe1** K Q E V T Q I P A A L S V P E G E N  
ctggtcctgaattgcagtttcaccgactcagccatctataacctgcagtggttcgccag  
L V L N C S F T D S A I Y N L Q W F R Q  
gatccaggcaagggactgacctccctgctgctgattcagagctcccagaggggaacagaca  
D P G K G L T S L L L I Q S S Q R E Q T  
tctggcagactgaatgctagtctggacaaatctagtggacggctctacctgtacatcgca  
S G R L N A S L D K S S G R S T L Y I A  
gccagccagcctggagattccgcaacatatctgtgcccgtgcgcccacttacaggcgga  
A S Q P G D S A T Y L C A V R P L T G G  
agctacattcccaccttcgggaggggtacaagcctgatcgtgcacccagacatccagaat  
S Y I P T F G R G T S L I V H P D I Q N  
ccggagcccgcctataaccagctgaaggaccccagaagccaggacagcaccctgtgcctg  
P E P A V Y Q L K D P R S Q D S T L C L  
ttcaccgacttcgacagccagatcaacgtgcccagacaatggaaagcggcaccttcac  
F T D F D S Q I N V P K T M E S G T F I  
accgacaagaccgtgctggacatgaaggctatggacagcaagagcaacggcgccattgcc  
T D K T V L D M K A M D S K S N G A I A  
tggtccaaccagaccagcttcacatgccaggacatcttcaaagagacaaacgccacctac  
W S N Q T S F T C Q D I F K E T N A T Y  
cccagcagcgacgtgccctgtgatgccaccctgaccgagaagtccttcgagacagacatg  
P S S D V P C D A T L T E K S F E T D M  
Aacctgaacttcagaacctgtccgcccgcaggcctgctggatcccaagctgtgctac

**Not1**

N L N F Q N L S A A A G L L D P K L C Y  
ctgctggacgggatcctgttcacatctacgggtgtgatcctgactgccctgttcctgcgagtc  
L L D G I L F I Y G V I L T A L F L R V  
aattttctcggagtgccgacgctcctgcataaccagcaggggcagaaccagctgtataac  
K F S R S A D A P A Y Q Q G Q N Q L Y N  
gagctgaatctgggtcggaggagggaatatgacgtgctggataagagacgcggcagggat  
E L N L G R R E E Y D V L D K R R G R D  
ccagaaatggggggcaagccccagcgacggaaaaaccctcaggaggactgtataatgaa  
P E M G G K P Q R R K N P Q E G L Y N E  
ctgcagaaggacaaaatggccgaggettactctgaaattgggatgaagggcgagaggaga  
L Q K D K M A E A Y S E I G M K G E R R  
cgcgcaaaaggacacgatggcctgtaccagggactgagcactgctaccaaggacacatat  
R G K G H D G L Y Q G L S T A T K D T Y  
gatgctctgcatatgcaggcactgcccctagataataaggctacc  
D A L H M Q A L P P R - - **Acc65I**

Фиг. 25

d

ccatggccaacgctggagtgactcagaccctaagttccaggtcctgaaaactgggcag  
**Nco** M A N A G V T Q T P K F Q V L K T G Q  
 agtatgacctgcagtgccacaggacatgaatcacgagtacatgtcatggatcggcag  
 S M T L Q C A Q D M N H E Y M S W Y R Q  
 gatccagggatgggtctgaggctgatccattacagcgtgggcgctggaactaccgaccag  
 D P G M G L R L I H Y S V G A G T T D Q  
 ggcgaggtgccaacggatataatgtctcaagaagcaccacagaagatttcccactgcca  
 G E V P N G Y N V S R S T T E D F P L R  
 ctgctgagcgcgctcctagccagacatccgtgtacttttgcgccagctccaatgtcggg  
 L L S A A P S Q T S V Y F C A S S N V G  
 aacaccggcgagctgttctttgggaagggttccgcctgacagtgctcgaggacctgaga  
 N T G E L F F G E G S R L T V L E D L R

**Xho1**

aacgtgacccccccaaggtgtccctgttcgagcctagcaaggccgagatcgccaacaag  
 N V T P P K V S L F E P S K A E I A N K  
 cagaaagccacctcgtgtgectggccagaggttcttccccgaccacgtggaactgtct  
 Q K A T L V C L A R G F F P D H V E L S  
 tgggtgggtcaacggcaaagaggtgcacagcggcgtgtccaccgatccccaggectacaaa  
 W W V N G K E V H S G V S T D P Q A Y K  
 gagagcaactacagctactgctgagcagcagactgcccgtgtccgccaccttctggcac  
 E S N Y S Y C L S S R L R V S A T F W H  
 aacccccggaaccacttcagatgccaggtgcagtttcacggcctgagcgaagaggacaag  
 N P R N H F R C Q V Q F H G L S E E D K  
 tggcccggggcagccctaagcccgtgaccagaatatctctgccgaagcctggggcaga  
 W P E G S P K P V T Q N I S A E A W G R  
 gccgactgtggcattaccagcgcagctaccagcagggcgtgtgtctgccaccatcctg  
 A D C G I T S A S Y Q Q G V L S A T I L  
 tacgaggtcgcgagcggactgctggacccaaagctgtgetacctgctggatgggatcctg  
 Y E V A S G L L D P K L C Y L L D G I L  
 ttcatctacgggtgtgattctgacagccctgttctcgcgagtcaagttcagccggagcgc  
 F I Y G V I L T A L F L R V K F S R S A  
 gagcaccagcataaccagcaggggcagaatcagctgtataacgagctgaatctgggtcgg  
 D A P A Y Q Q G Q N Q L Y N E L N L G R  
 agggaggaatacagcgtgctggataagagacgcggcagggatcccgaatgggcggaag  
 R E E Y D V L D K R R G R D P E M G G K  
 cctcagcagcggaaaaaccacaggaggactgtacaatgaactgcagaaggacaaaatg  
 P Q R R K N P Q E G L Y N E L Q K D K M  
 gctgaggcatattctgaaatcggcatgaagggagagaggagacgcggcaaggacacgat  
 A E A Y S E I G M K G E R R R G K G H D  
 gggctgtaccaggtctgagtacagccactaaggacacctatgatgccctgcatatgcag  
 G L Y Q G L S T A T K D T Y D A L H M Q  
 gctctgccaccagataataaaagctt  
 A L P P R - - **Hind3**

Фиг. 25 (продолжение)

**e**

gctagccagaaagaggtggaacagaacagcggccctctgagcgtgccagaaggcgctatc  
**NheI** Q K E V E Q N S G P L S V P E G A I  
 gccagcctgaactgcacctacagctttctgggcagccagagcttcttctggtacagacag  
 A S L N C T Y S F L G S Q S F F W Y R Q  
 tacagcggcaagagccccgagctgatcatgttcacctacagagagggcgacaaagaggac  
 Y S G K S P E L I M F T Y R E G D K E D  
 ggcagattcaccgcccagctgaacaaggccagccagcacgtgtccctgctgatcagagac  
 G R F T A Q L N K A S Q H V S L L I R D  
 agccagcctagcgcagcgcacctacctgtgcgcctgaatgatggcggcagactgacc  
 S Q P S D S A T Y L C A V N D G G R L T  
 tttggcgacggcaccacctgaccgtgaagcctgacatccagaatccggagcccgcctga  
 F G D G T T L T V K P D I Q N P E P A V  
 taccagctgaaggaccccagaagccaggacagcaccctgtgcctgttcaccgacttcgac  
 Y Q L K D P R S Q D S T L C L F T D F D  
 agccagatcaacgtgcccagaacaatggaagcggcaccttcatcaccgacaagaccgtg  
 S Q I N V P K T M E S G T F I T D K T V  
 ctggacatgaaggctatggacagcaagagcaacggcggcattgcctggtccaaccagacc  
 L D M K A M D S K S N G A I A W S N Q T  
 agcttcacatgccaggacatcttcaaagagacaaacgccacctaccccagcagcgcgtg  
 S F T C Q D I F K E T N A T Y P S S D V  
 ccctgtgatgccacctgaccgagaagtctctcgagacagacatgaacctgaactccag  
 P C D A T L T E K S F E T D M N L N F Q  
aacctgtccgcccgc  
 N L S **NotI**

**f****NcoI**

ccatggccagccagaccatccatcagtgccctgccaccctgggtgcagcctgtgggatct  
 M A S Q T I H Q W P A T L V Q P V G S  
 cctctgagcctggaatgcaccgtggaaggcaccagcaaccccaacctgtactggtacaga  
 P L S L E C T V E G T S N P N L Y W Y R  
 cagggcgtggcagaggccccagctgctgttttactggggccctttggccagatcagc  
 Q A A G R G P Q L L F Y W G P F G Q I S  
 agcgaggtgccccagaacctgagcgcagcagaccccaggaccggcagtttatcctgagc  
 S E V P Q N L S A S R P Q D R Q F I L S  
 agcaagaagctgctgctgagcgcagcggcttctacctgtgcgcttggagcgcagacaggc  
 S K K L L L S D S G F Y L C A W S E T G  
 ctgggcatggcgatggcagtttggcgagggcagcagactgacagtgctcgaq  
 L G M G G W Q F G E G S R L T V L E  
**XhoI**

Фиг. 25 (продолжение)

g

```

5' C ACG CGG CAC GCG GGT GAA NNS NNS CCT WCN ATG TAA GGG TGG AAG CCC GCTC
   C ACG CGG CAC GCG GGT GAA NNS NNS CCT WCN ATG TAA GGG TGG AAG CCC GCTC 5'
   C   A   V   R   P   L   X   X   G   S/T   Y   I   P   T   F   G   R
   -----CDR3-----

```

h

```

5' G TAC TTT TGC GCC AGC TCC NNS STC GGG NNS ACC GGC GAG CTG TTC TTTG
   C ATG AAA ACG CGG TCG AGG NNS SAG CCC NNS TGG CCG CTC GAC AAG AAAC 5'
   Y   F   C   A   S   S   N   V/L   G   N   T   G   E   L   F   F
   -----CDR3-----

```

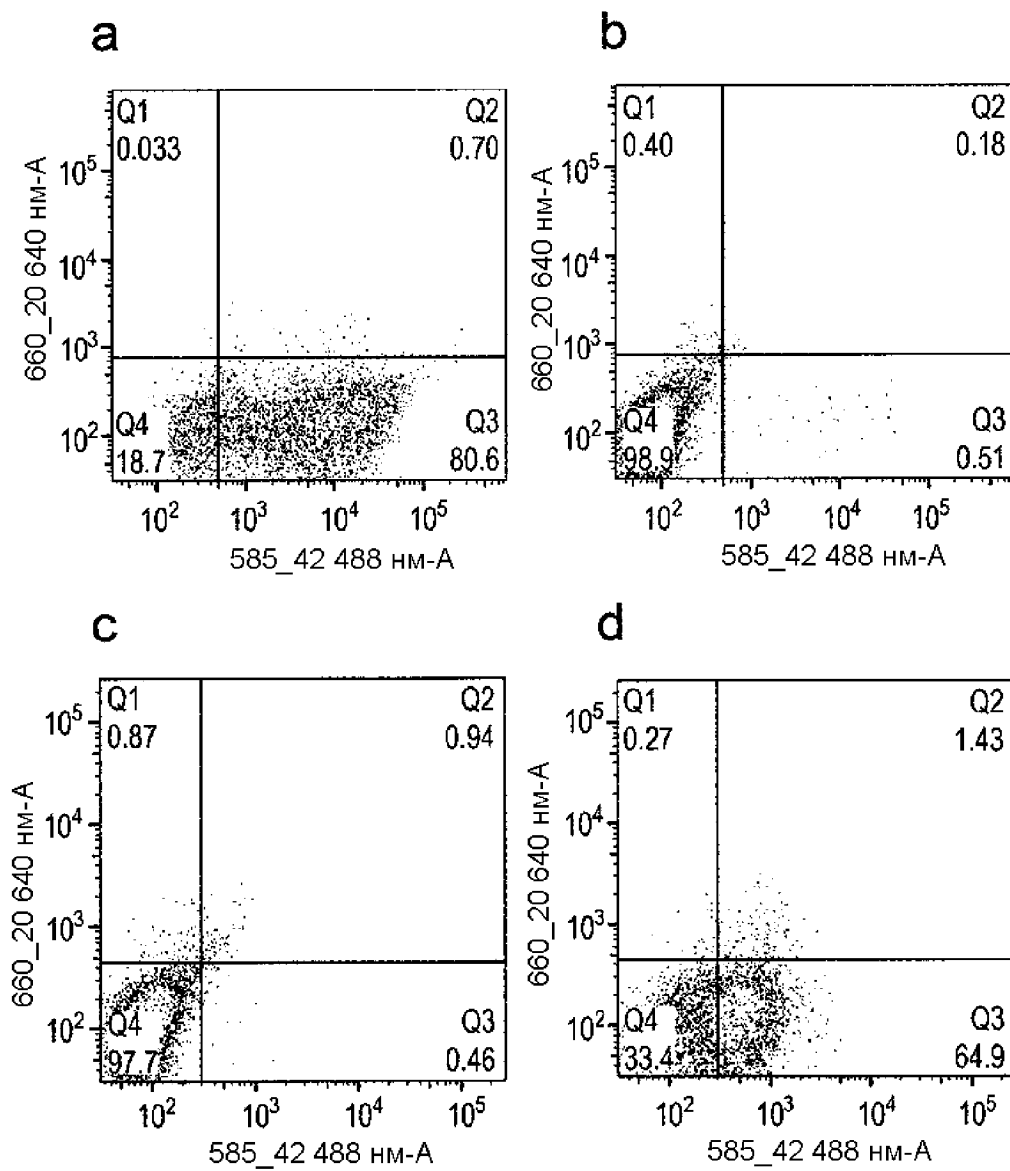
N=A, C, G или T

S= C или G

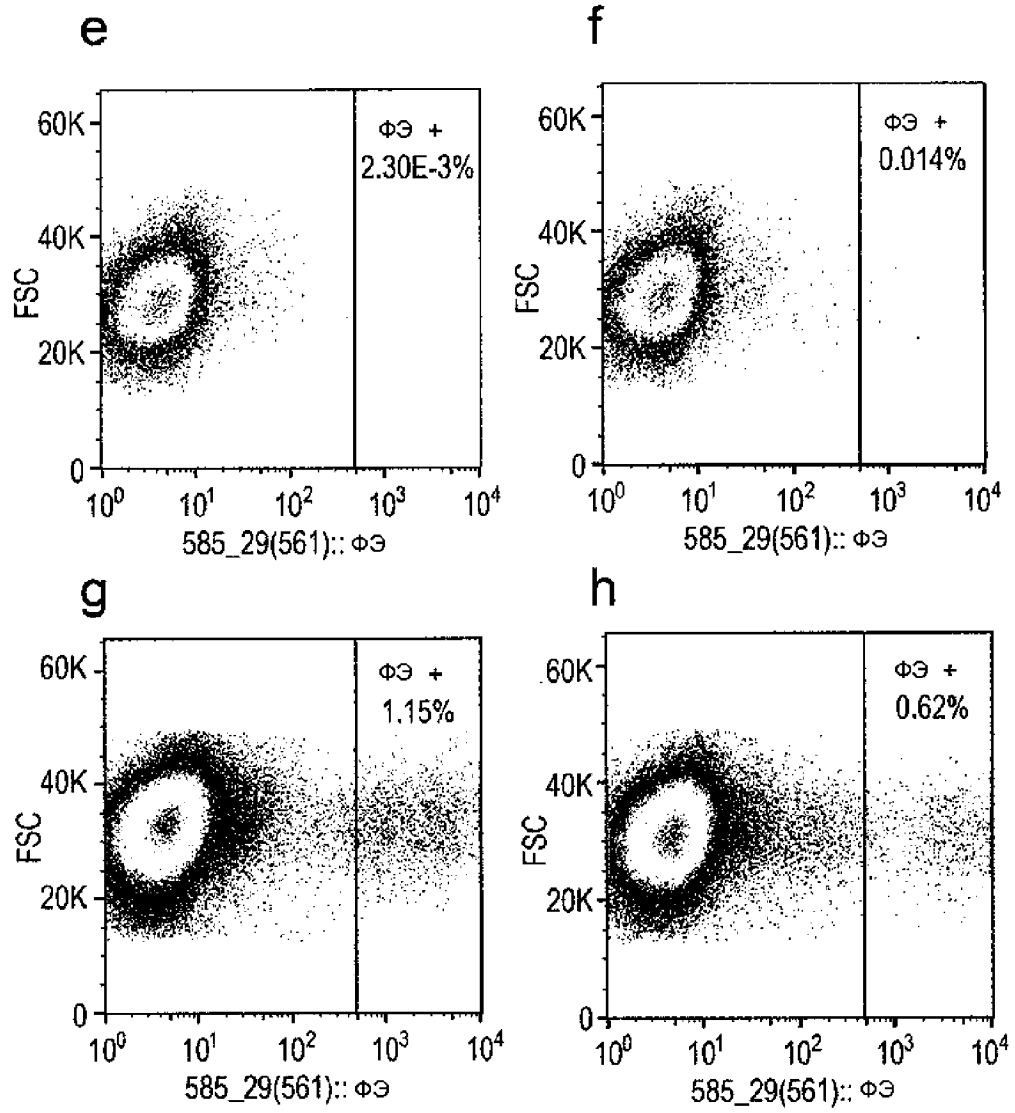
W= A или T

Фиг. 25 (продолжение)

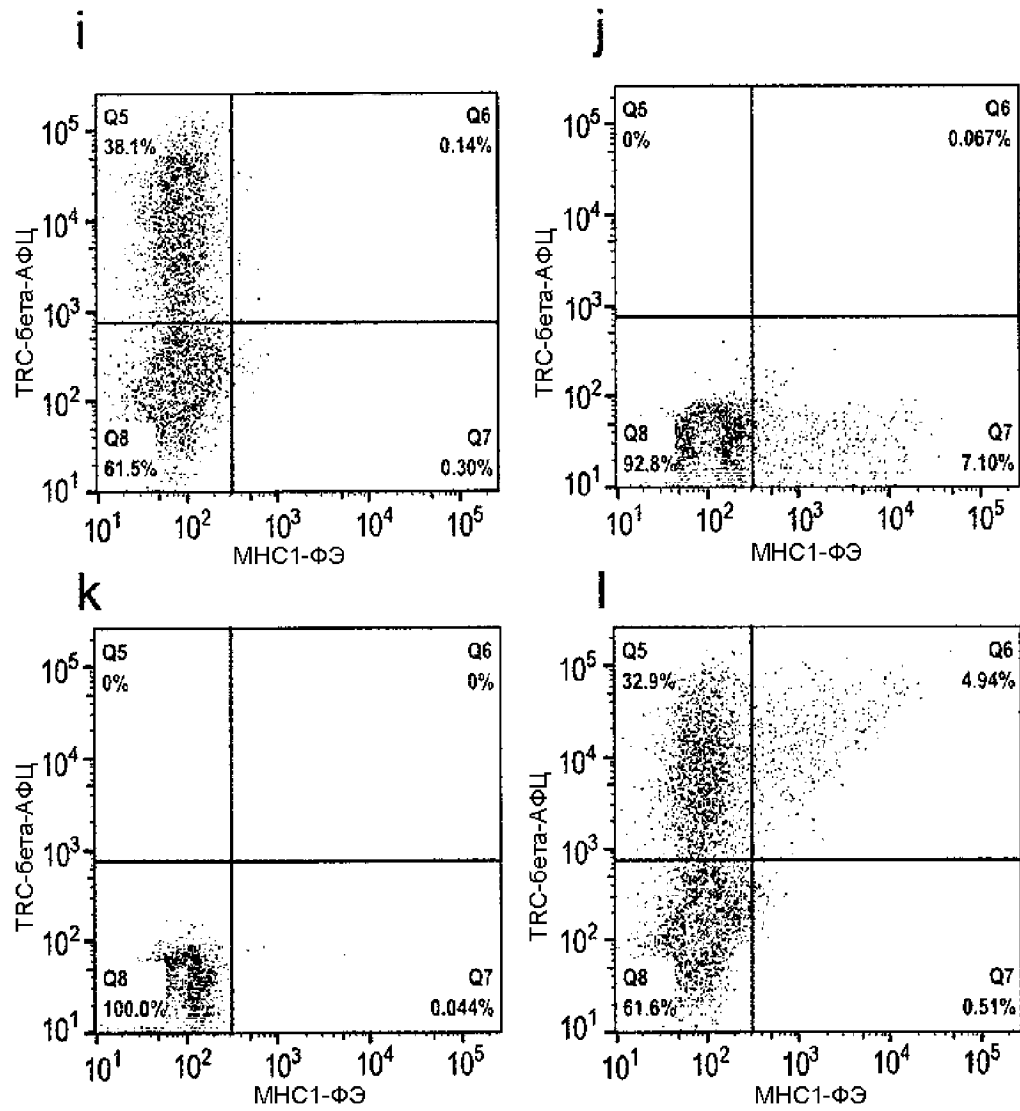




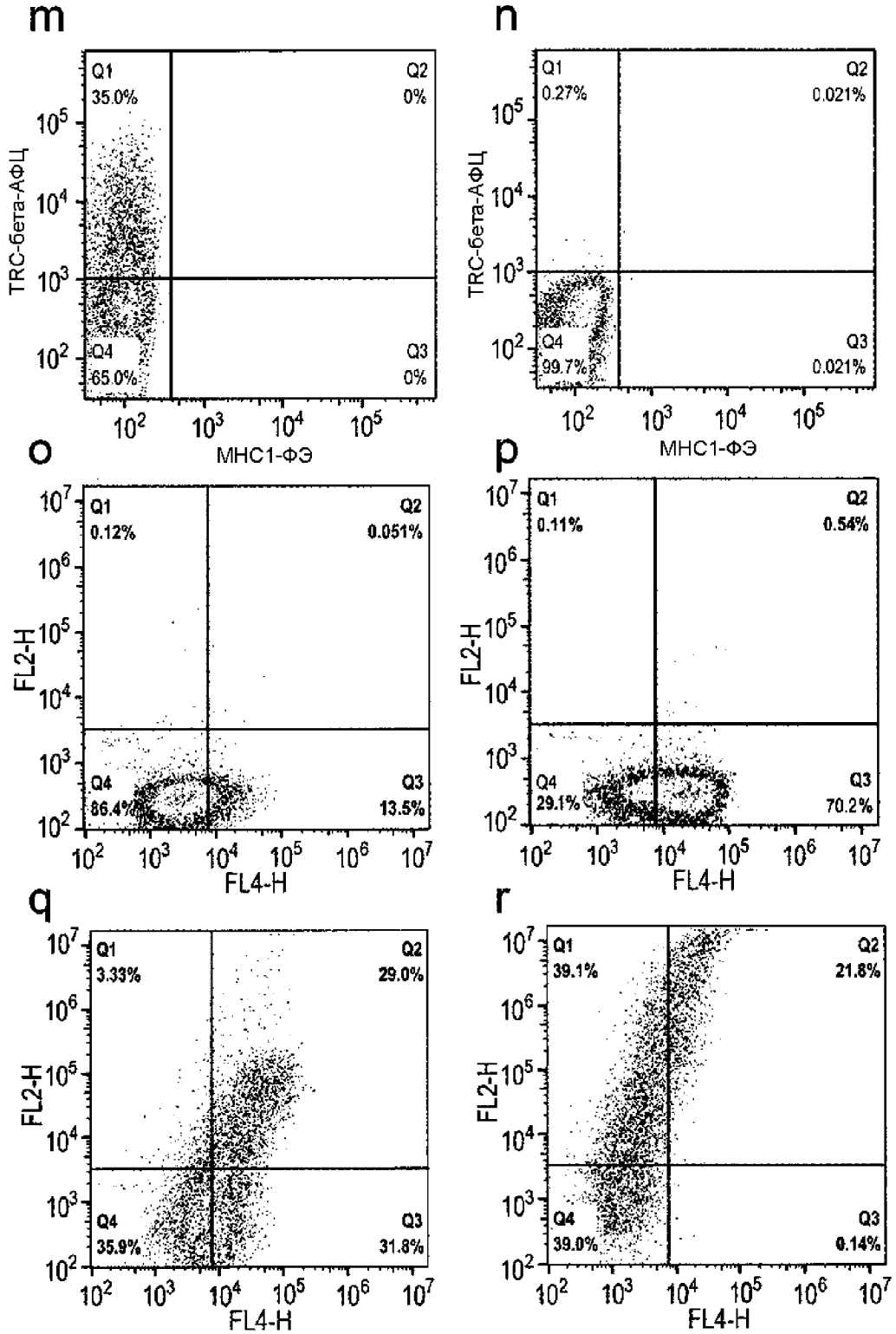
ФИГ. 26



**Фиг. 26**  
(продолжение)

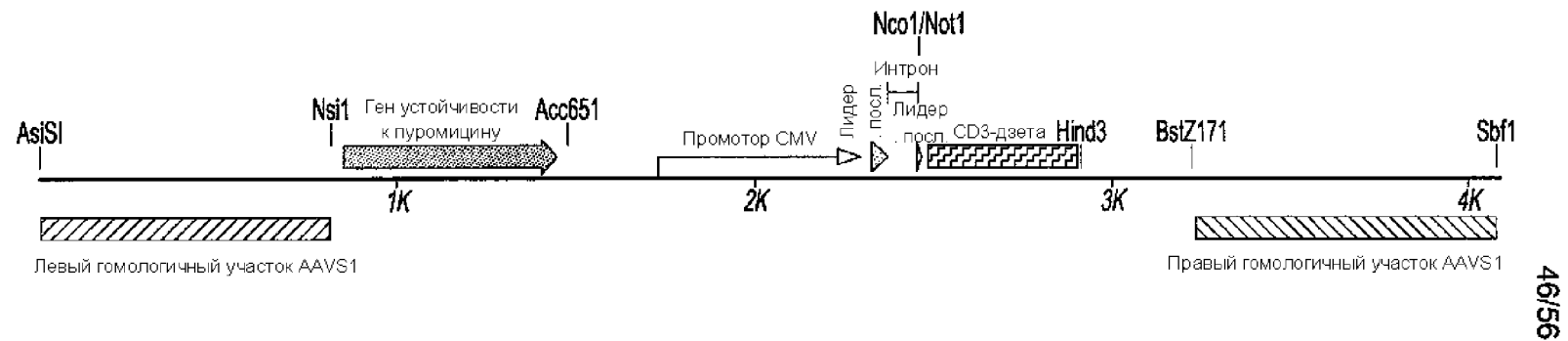


**Фиг. 26**  
(продолжение)

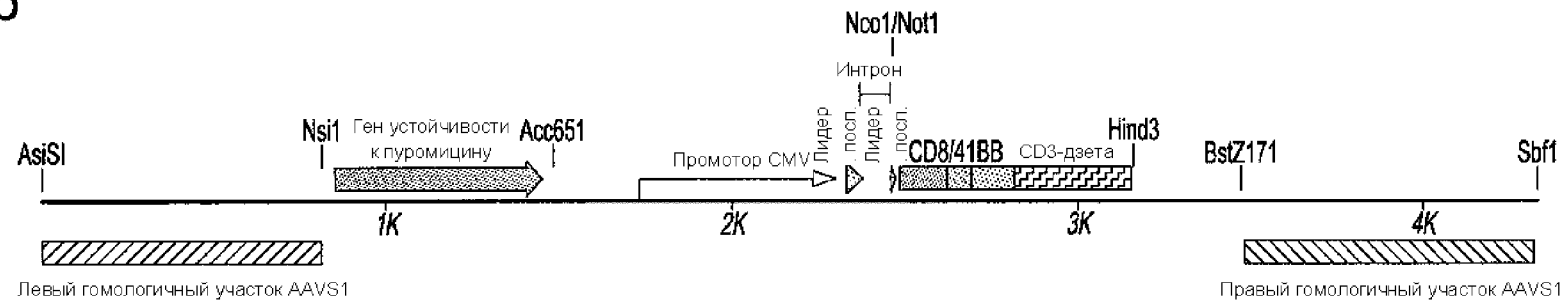


Фиг. 26 (продолжение)

a



b



Фиг. 27

**C**Лидерная  
последовательность**Nco1****Not1**

ggcgccatggcccaggtcgcgccgcaagcggactgctggacccaaagctgtgctacctg  
 G A M A S G L L D P K L C Y L  
 ctggatgggatcctgttcacatctacggtgtgattctgacagccctgttcctgcgagtcaag  
 L D G I L F I Y G V I L T A L F L R V K  
 ttcagccggagcgcggacgcaccagcataccagcaggggcagaatcagctgtataacgag  
 F S R S A D A P A Y Q Q G Q N Q L Y N E  
 ctgaatctgggtcggagggaggaatacgcagctgctggataagagacgcggcagggatccc  
 L N L G R R E E Y D V L D K R R G R D P  
 gaaatgggcccgaagcctcagcgcggaaaaacccacaggagggactgtacaatgaactg  
 E M G G K P Q R R K N P Q E G L Y N E L  
 cagaaggacaaaatggctgagggcatattctgaaatcggcatgaaggagagaggagacgc  
 Q K D K M A E A Y S E I G M K G E R R R  
 ggcaaaggacacgatgggctgtaccagggctctgagtacagccactaaggacacctatgat  
 G K G H D G L Y Q G L S T A T K D T Y D

**Hind3**

gcctgcatatgcaggctctgccaccagataataaaagcctt  
 A L H M Q A L P P R - -

**Фиг. 27** (продолжение)



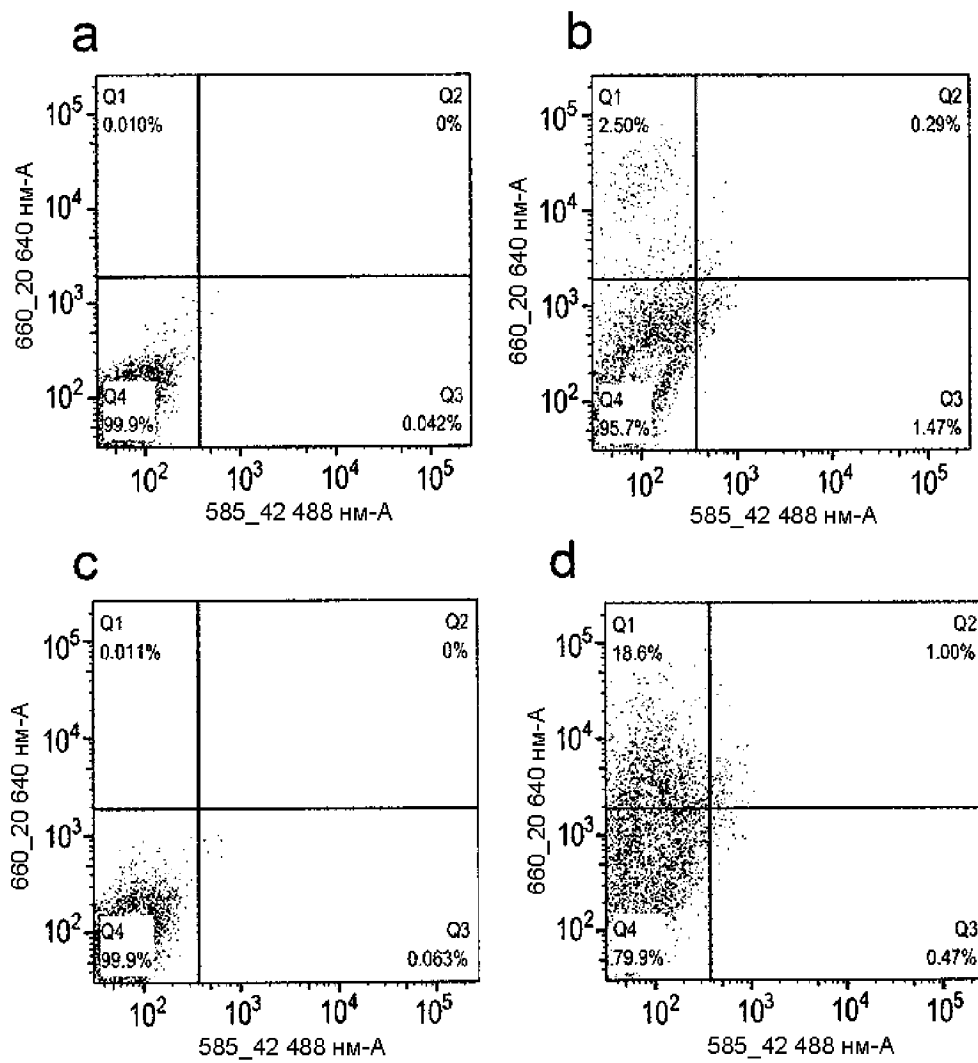
e

**Nco1**

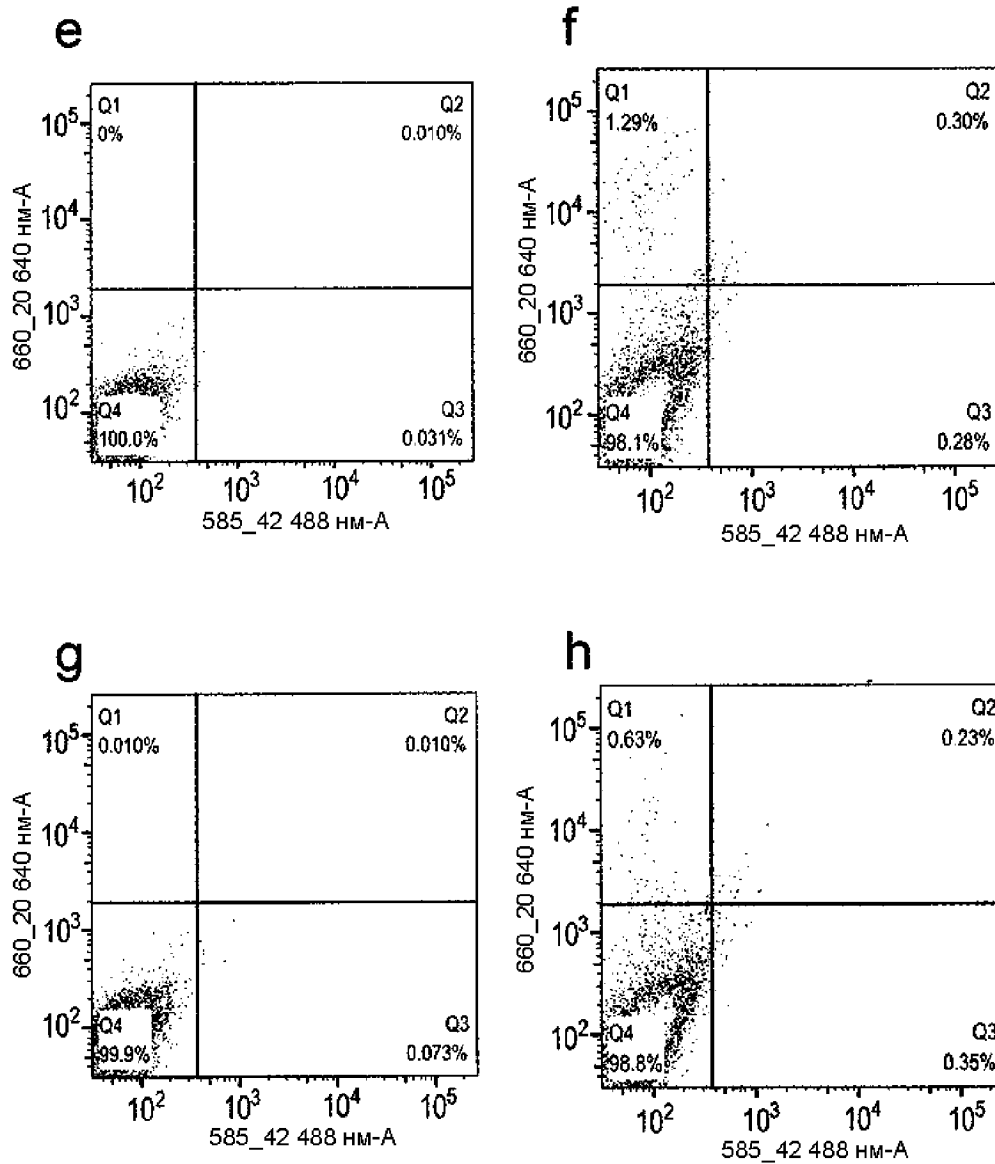
**g**ccatg**g**ccgaagtgaaactgcaggagtctggacccggcctggtggccccatctcagtct  
 A M A E V K L Q E S G P G L V A P S Q S  
 ctgagcgtgacctgtaccgtgtccggcgtgtccctgcctgactatggcgtgtcctggatc  
 L S V T C T V S G V S L P D Y G V S W I  
 agacagccccccagaaagggcctggaatggctgggagtgatctggggcagcgaaccacc  
 R Q P P R K G L E W L G V I W G S E T T  
 tactacaacagcgcctgaagtcccggtgaccatcatcaaggacaactccaagagccag  
 Y Y N S A L K S R L T I I K D N S K S Q  
 gtgttcctgaagatgaacagcctgcagaccgacacccgcatctactactgcgccaag  
 V F L K M N S L Q T D D T A I Y Y C A K  
 cactactactacggcggcagctacgctatggactactggggccagggcacctcggtcacc  
 H Y Y Y G G S Y A M D Y W G Q G T S V T  
 gtctcgagtggaggggcgttcaggcggaggtggctctggcggtggcgctagcgacatc  
 V S S G G G G S G G G G S G G G A S D I  
 cagatgaccagaccaccagcagcctgagcggcagcctgggcatagagtgaccatcagc  
 Q M T Q T T S S L S A S L G D R V T I S  
 tgcagagccagccaggacatcagcaagtacctgaactgggtatcagcagaaacccgacggc  
 C R A S Q D I S K Y L N W Y Q Q K P D G  
 accgtgaagctgctgatctaccacaccagcagactgcacagcggcgtgcccagcagattt  
 T V K L L I Y H T S R L H S G V P S R F  
 tccggctctggcagcggcaccgactacagcctgaccatctccaacctggaacaggaagat  
 S G S G S G T D Y S L T I S N L E Q E D  
 atcgctacctactctgtcagcaaggcaacaccctgcctacaccttcggcggagggacc  
 I A T Y F C Q Q G N T L P Y T F G G G T  
 aagctggagatcaaactgaccg**g**ccg**g**cc**g**ca  
 K L E I K R T A A A

**Not1****Фиг. 27** (продолжение)





Фиг. 28



Фиг. 28 (продолжение)

**a****Nco1**

gccatggctgctacagcgctgcgggctgtgcccggaatgagaacagcctggaaatcgag  
 A M A A T G V R A V P G N E N S L E I E  
 gaactggccagattcgccgtggacgagcacaacaagaaagagaacgcctgctggaattc  
 E L A R F A V D E H N K K E N A L L E F  
 gtgcgggctcgtgaaggccaagagcagtgagcgaggccgacaacgactggcacaccatg  
 V R V V K A K E Q W S E A D N D W H T M  
 tactacctgaccctggaagccaaggacggcggaagaagaagctgtacgaggccaagtg  
 Y Y L T L E A K D G G K K K L Y E A K V  
 tgggtcaagctggacctggaacctggcagcacttcaacttcaaagagctccaggaattc  
 W V K L D L E T W Q H F N F K E L Q E F  
 aagcccgtgggcgacgctgcgcccgcg  
 K P V G D A A A A

**Not1****b****Nco1**

gccatggctgctacagcgctgcgggctgtgcccggaatgagaacagcctggaaatcgag  
 A M A A T G V R A V P G N E N S L E I E  
 gaactggccagattcgccgtggacgagcacaacaagaaagagaacgcctgctggaattc  
 E L A R F A V D E H N K K E N A L L E F  
 gtgcgggctcgtgaaggccaagagcaggaacagccatcggcgagcaccctggaacgac  
 V R V V K A K E Q E Q P I G E H P V N D  
 accatgtactacctgaccctggaagccaaggacggcggaagaagaagctgtacgaggcc  
 T M Y Y L T L E A K D G G K K K L Y E A  
 aaagtgtgggtcaagcgggtgctgcggttcaccgagatctacaacttcaaagagctccag  
 K V W V K R W L R F T E I Y N F K E L Q  
 gaattcaagcccgtgggcgacgctgcgcccgcg  
 E F K P V G D A A A A

**Not1****Фиг. 29**

**C****Adhiron\_mut1**5' CAGGGTCAGGTAGTACATGGTSNNSNNSNN(SNN)<sub>n</sub> CTGCTCTTTGGCCTTCACGAC**Adhiron\_mut2**CTGGAGCTCTTTGAAGTTSNNSNNSMN(SNN)<sub>n</sub> CTIGACCCACACTTTGGC

GTC GTG AAG GCC AAA GAG CAG (NNS)<sub>n</sub> NNS NNS NNS ACC ATG TAC TAC CTG ACC CTG  
 CAG CAC TTC CGG TTT CTC GTC (NNS)<sub>n</sub> NNS NNS NNS TGG TAC ATG ATG GAC TGG GAC  
 V V K A K E Q X X X X T M Y Y L T L

GCC AAA GTG TGG GTC AAG (NNS)<sub>n</sub> NNS NNS NNS AAC TTC AAA GAG CTC CAG  
 CGG TTT CAC ACC CAG TTC (NNS)<sub>n</sub> NNS NNS NNS TTG AAG TTT CTC GAG GTC  
 A K V W V K X X X X N F K E L Q

**Фиг. 29** (продолжение)

d

Nco1

gccatggccggtgtgtgcccccaagatcttgaaaaagtgccgccgtgacagcgattgt  
 A M A G V C P K I L K K C R R D S D C  
 ccggcgcctgcatotgccgcggcaatggctattgccgagccggccgca  
 P G A C I C R G N G Y C G **Not 1**

e

GGT GTG TGC VNS VNS VNS VNS VNS VNS VNS VNS VNS VNS TGC CGC CGT  
 CCA CAC ACG BNS BNS BNS BNS BNS BNS BNS BNS BNS BNS ACG GCG GCA  
 G V C X X X X X X X X X X X X C R R

GAC AGC GAT TGT CCC GGC GCC TGC ATC TGC CGC GGC AAT GGC TAT TGC GGA  
 CTG TCG CTA ACA GGG CCG CGG ACG TAG ACG GCG CCG TTA CCG ATA ACG CCT  
 D S D C P G A C I C R G N G Y C G

**Фиг. 29** (продолжение)



**a**

Праймер J60&gt;

GCGATCGCGCTGATTGGCTTCCTTTCCCTCCCGCCGTGTGTGAAAACACAAATGGCGTGT  
 GGTGGCGTAAGGCGCCTGTCAGTTAACGGCAGCCGGAGTGCAGCCGCCGGCAGCCTCGC  
 TCTGCCACTGGGTGGGGCGGGAGGTAGGTGGGGTGGGGCAGCTGGACGTGGGGGCGCGGT  
 CGGCCTCTGGCGGGGCGGGGAGGGGAGGGTCAAGCAAAGTAGCTCCGCGCGGAGCGG  
 CCGCCACCCCTCCCCTTCCCTCTGGGGGAGTCTTTTACCCGCCCGCCGGCCGGGCTCGTCTG  
 CTGATTGGCTCTCGGGGCCAGAAAACCTGGCCCTTGCCATTGGCTCGTGTTCGTGCAAGTTG  
 AGTCCATCCGCCGGCCAGCGGGGGCGCGAGGAGGGCGCTCCCAGGTTCCGGCCCTCCCCTCG  
 GCCCGCGCCGAGAGTCTGGCCGCGCGCCCTGCGCAACGTGGCAGGAAGCGCGCGCTGGG  
 GCGGGGACGGGCAGTAGGGCTGAGCGGCTGCGGGGCGGGTGAAGCAGCTTCCGACTTGA  
 GTTGCCTCAAGAGGGGCGTGCTGAGCCAGACCTCCATCGCGCACTCCGGGGAGTGGAGGAA  
 GGAGCGAGGGCTCAGTTGGGCTGTTTTGGAGGCAGGAAGCACTTGCTCTCCCAAAGTCGCTC  
 TGAGTTGTTATCAGTAAGGGAGCTGCAGTGGAGTAGGCGGGGAGAAGGCCGACCCCTCTCC  
 GGAGGGGGAGGGGAGTGTGCAATACCTTCTGGGAGTTCTCTGCTGCCTCCTGGCTCTG  
 AGGACCGCCCTGGGCTGGGAGAATCCCTTCCCTCTTCCCTCGTGATCTGCAACTCCAGT  
CTTTCTAGAAATGCATTAAGGGATCTGTAGGGCGCAGTAGTCCAGGGTTTCCTTGATGATGC  
 < 2706 NsiI 2709→

ATACTTATCCTGTCCCTTTTTTTTCCACAGCTCGCGGTTGAGGACAAACTCTTCGCGGTCTT  
 TCCAGTGGGATCGACGGTATCGTAGAGTCGAGGCGCTCTAGGAATTCAGCCGCCACC

← Праймер 2710

ATG ACC GAG

M T E

Ген устойчивости →  
к пурамицину

**b**

BstZ171 Праймер J61→

GTATACGGGAATTGAACAGGTGTAATAATTGGAGGGACAAGACTTCCCACAGATTTTCGGTTT  
 TGTCGGGAAGTTTTTAATAGGGGCAAATAAGGAAAATGGGAGGATAGGTAGTCATCTGGGG  
 TTTTATGCAGCAAACTACAGGTATTATTGCTTGATCCGCCTCGGAGTATTTCCATCG  
 AGGTAGATTAAAGACATGCTCACCCGAGTTTTATACTCTCCTGCTTGAGATCCTTACTACAG  
 TAGAAATTACAGTGTGCGGAGTTAGACTATGTAAGCAGAATTTAATCATTTTTAAAGAGC  
 CCAGTACTTTCATATCCATTTCTCCCGCTCCTTCTGCAGCCTTATCAAAAGGTATTTTAGAAC  
 ACTCATTTTTAGCCCATTTTCATTTATTATACTGGCTTATCCAACCCCTAGACAGAGCATTG  
 GCATTTTCCCTTTCCCTGATCTTAGAAGTCTGATGACTCATGAAACCAGACAGATTAGTTACA  
 TACACCACAATCGAGGCTGTAGCTGGGGCTCAACACTGCAGTTCCTTTATAACTCCTTAG  
 TACACTTTTTGTTGATCCTTTGCCCTGATCCTTAATTTTCAGTGTCTATCACCTCTCCCGTC  
 AGGTGGTGTTCACATTTGGCCCTATTCTCAGTCCAGGGAGTTTTACAACAATAGATGTATT  
GAGAATCCAACCTCCTGCAGG

← Праймер J62 SbfI