

(19)



Евразийское
патентное
ведомство

(21) 202290477 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2022.05.11

(51) Int. Cl. C07K 16/00 (2006.01)

(22) Дата подачи заявки
2020.08.04

(54) МОДИФИЦИРОВАННЫЕ ВАРИАБЕЛЬНЫЕ ДОМЕНЫ ЧЕЛОВЕКА

(31) 19190578.5

(32) 2019.08.07

(33) EP

(86) PCT/NL2020/050496

(87) WO 2021/025555 2021.02.11

(71) Заявитель:
МЕРУС Н.В. (NL)

(72) Изобретатель:

Силвермен Питер Брайан, Крамер
Роберт Арьен, Баккер Александер
Бертолд Хенрик (NL)

(74) Представитель:

Веселицкий М.Б., Веселицкая И.А.,
Кузенкова Н.В., Каксис Р.А., Белоусов
Ю.В., Куликов А.В., Кузнецова Е.В.,
Соколов Р.А., Кузнецова Т.В. (RU)

(57) В заявке представлены полипептиды, содержащие переменные домены тяжелой цепи модифицированного человеческого, гуманизированного или химерного иммуноглобулина. Также предусмотрены соответствующие антитела, варианты, фрагменты, нуклеиновые кислоты, векторы, фаги, библиотеки, способы и наборы.

202290477

A1

A1

202290477

МОДИФИЦИРОВАННЫЕ ВАРИАБЕЛЬНЫЕ ДОМЕНЫ ЧЕЛОВЕКА

5 Область техники

Настоящее изобретение относится к полипептидам, содержащим модифицированные, гуманизированные или химерные переменные домены тяжелой цепи иммуноглобулина человека. Также в изобретении предусматривают соответствующие антитела, варианты, фрагменты, нуклеиновые кислоты, векторы, фаги, библиотеки, способы и наборы.

Предшествующий уровень техники

Терапевтические белки, такие как антитела, содержат ряд посттрансляционных модификаций, некоторые из которых могут оказывать нежелательное воздействие на белок. Например, пироглутамат (обычно обозначаемый аббревиатурой pE, pyroE или pyroGlu) может образовываться на N-конце полипептидной цепи *in vitro* и *in vivo*. Образование пироглутамата происходит путем перегруппировки первоначально синтезированных остатков глутамата или глутамина в этом положении (см. фиг. 1).

Первый остаток на N-конце переменного домена тяжелой цепи (VH) человеческого, гуманизированного или химерного антитела, кодируемый в результате перегруппировки сегмента гена VH зародышевой линии, обычно представляет собой либо глутамин, либо глутамат. Было показано, что как глутамин, так и глутамат на N-концах антител спонтанно циклизируются в пироглутамат *in vitro*. Когда пироглутамат образуется путем циклизации глутамина, полученное антитело становится более кислым. И наоборот, когда пироглутамат образуется путем циклизации глутамата, полученное антитело становится менее кислым. Со временем это может привести к неоднородности заряда в препаратах антител, что может быть нежелательно в различных контекстах. Может оказаться полезным уменьшение такой изменчивости в препаратах антител.

Однако присутствие глутамина или глутамата на N-конце домена VH может быть важным. Действительно, при секреции антител из прокариотических и эукариотических клеток сигнальный пептид (SP – signal peptide; также известный как лидерный пептид) удаляется с N-конца незрелой тяжелой цепи

путем расщепления между сигнальным пептидом и вариабельным доменом тяжелой цепи. Эффективность отщепления сигнального пептида зависит от последовательности сигнального пептида, а также от последовательности домена VH. Пептидный сегмент, распознаваемый ферментами сигнальных пептидаз (такими как SPaseI), простирается до начала зрелого белка (Choo с соавт., 2008). Следовательно, фланкирующие остатки в домене VH могут влиять на процессинг сигнальной пептидазы и вносить вклад в неканонические сайты расщепления.

Краткое описание изобретения

Неожиданно было обнаружено, что аминокислотные вариации могут быть введены в N-конец человеческого, гуманизированного или химерного домена VH, сохраняя при этом требуемую аффинность, специфичность и/или структурные взаимодействия соответствующих немодифицированных человеческих, гуманизированных или химерных доменов VH. Преимущественно это снижает вариабельность (например, гетерогенность заряда) в белковых препаратах (например, препаратах антител), которые включают модифицированный человеческий, гуманизированный или химерный домен VH.

Таким образом, впервые представленные в настоящем документе новые антитела, вариабельные области тяжелых цепей, варианты, фрагменты, нуклеиновые кислоты, векторы, фаги, библиотеки, способы и наборы, содержащие модифицированные N-концевые остатки, утратившие глутамин или глутамат, при этом не оказывая вредного воздействия на отщепление сигнального пептида или выработку белка. Кроме того, описан способ получения антител, в котором целые панели человеческих, гуманизированных или химерных антител, вариабельных областей тяжелой цепи и нуклеиновых кислот, которые их кодируют, могут быть модифицированы для универсального удаления N-концевого глутамина или глутамата, чтобы можно было отобразить, исследовать и оценить полные панели таких антител человека, вариабельных областей тяжелой цепи с начала, без такого остатка, избегая задачи выполнения последующих модификаций после идентификации ведущего кандидата, экономя время и усилия, и предпочтительно, хотя и не причиняя вреда, воздействуя на отщепление сигнального пептида или экспрессию белка.

Кроме того, при получении панелей тяжелых цепей или антител для скрининга активности с помощью технологии дисплея или функциональных анализов желательно тестировать такие белки в их окончательной форме, чтобы избежать изменчивости результатов такого анализа и избегать проведения

5 нескольких раундов одного и того же анализа на основе вариантов, которые могут быть использованы при последующей наработке. Однако общепринятой практикой является анализ тяжелых цепей или антител на предмет удаления остатков, которые могут привести к посттрансляционным модификациям, только после идентификации лидерного элемента, такого как антитело, и после

10 выяснения функции. Затем определенные остатки могут быть изменены с помощью различных методов рекомбинации, известных в данной области, включая мутагенез нуклеиновых кислот или синтез нуклеиновых кислот, приводящий к новой последовательности, кодирующей вариации, которые изменяют остатки, в противном случае вызывая потенциальные изменения

15 лидерного элемента во время хранения. Необходимость специфически модифицировать каждый белок, полученный в результате образования антител, например, после иммунизации и сбора В-клеток иммунизированного животного (трансгенного или иного) или перед получением гибридомы или библиотеки для анализа полученных антител, является трудоемким и время затратным

20 процессом. Соответственно, в настоящем изобретении разработаны новые средства модификации всей панели антител, тяжелых цепей или переменных областей, содержащих домен человеческой, гуманизированной или химерной VH, посредством изменения N-концевого глутамина и глутамата.

Еще одно преимущество состоит в том, что можно использовать

25 стандартный вектор, содержащий нуклеиновую кислоту, кодирующую сигнальный пептид, и содержащий первую или две первые аминокислоты домена VH, причем первая аминокислота домена VH не является глутамином или глутаматом, и по меньшей мере первая или первые две аминокислоты отличаются от немодифицированного домена VH, также называемого в

30 настоящем описании контрольной родительской переменной областью, или обычного человеческого, гуманизированного или химерного домена VH. Такой вектор можно использовать в клетке-хозяине для создания панели антител или тяжелых цепей, которые содержат модифицированный домен VH согласно

описанию настоящего изобретения, и дополнительного повышения надежности и эффективности выработки антител.

Другая сложность модификации каждого антитела или тяжелой цепи, полученных после иммунизации или создания панели и перед созданием библиотеки дисплея, заключается в том, что при этом могут быть затруднены сигнальные пептиды, ответственные за дисплей антител или тяжелых цепей на клеточной поверхности, что приводит к ухудшению широты и надежности библиотек дисплея.

Тот факт, что вариации аминокислот могут быть введены в N-конец человеческого, гуманизированного или химерного домена VH, позволяет использовать более эффективный метод получения антител, поскольку модификация может быть осуществлена во время создания панелей антител, после того, как выбраны один или несколько лидерных элементов.

В настоящем описании показано, что устранение глутамина или глутамата на N-конце человеческого, гуманизированного или химерного домена VH и замена такого остатка другой аминокислотой (такой как аланин) устраняет спонтанное образование пироглутамата на N-конце любого белка, такого как антитело, включающего модифицированный переменный домен.

Преимущественно, успешное получение Fab-фрагментов и библиотек фагового дисплея, содержащих модифицированный человеческий, гуманизированный или химерный домен VH, демонстрирует, что модифицированные человеческие, гуманизированные или химерные домены VH сохраняют требуемую аффинность, специфичность и/или структурные взаимодействия соответствующего немодифицированного человеческого, гуманизированного или химерного домена VH. Более того, применяемая методология с использованием универсальных праймеров позволяет создавать целые панели переменных областей тяжелой цепи человека или антител, у которых глутамин или глутамат удален на N-конце, независимо от того, какой сегмент V гена обусловил переменную область, что позволяет проводить единообразное тестирование таких панелей и включать их в высокопроизводительный скрининг.

Некоторые преимущества могут быть связаны с устранением образования пироглутамата. Согласно указанному выше, присутствие пироглутамата может изменить кислотность антител, что может привести к разрушению антител или

повлиять на срок годности. Также сообщалось, что антитела с пониженным рН, превышающим определенный пороговый уровень, вызывают боль у пациентов при инфузии. Устранение образования пироглутамата устраняет такие потенциально вредные свойства. Кроме того, устранение пироглутамата также полезно для целей регулирования, поскольку, например, основополагающие принципы Европейского медицинского агентства требуют от заявителей биологических препаратов подвергать мониторингу несколько структурных характеристик, включая состояние агрегации, N- и C-концы (пироглутаминовая кислота на N-конце и лизин на C-конце тяжелой цепи). Устранение N-концевого пироглутамата будет легче оценивать за счет исключения этой переменной для мониторинга. Отсутствие N-концевого пироглутамата также может привести к увеличению срока годности антител. Устранение образования пироглутамата также обеспечивает больший контроль процесса и может иметь преимущества против заболеваний, связанных с таким N-концевым остатком глутамина или глутамата. Отсутствие N-концевого пироглутамата также обеспечивает снижение неоднородности заряда в антителах и, следовательно, более эффективную очистку и разделение на основе заряда. Добавление N-концевых модификаций и последующий анализ также могут быть проще в отсутствие пироглутамата.

В настоящем исследовании неожиданно было установлено, что такая аминокислотная вариация (вариации) может способствовать эффективному отщеплению сигнального пептида от N-конца незрелого человеческого, гуманизированного или химерного домена VH. Например, введение остатка аланина (необязательно вместе с дополнительным вторым остатком, например, аланин-пролин, аланин-аспартат, аланин-глутамат, аланин-треонин, аланин-валин, аланин-серин или аланин-лейцин) на **N-конце человеческого, гуманизированного или химерного домена VH** может усиливать расщепление между доменом VH и расположенным выше по цепи сигнальным пептидом, таким образом, не оказывая неблагоприятного воздействия на экспрессию и секрецию, и предпочтительно повышать эффективную экспрессию и секрецию модифицированного человеческого, гуманизированного или химерного домена VH, а также антител или фрагментов антител, например фрагментов Fab, в состав которых входят такие домены.

В настоящем изобретении предусмотрено несколько универсальных праймеров, которые можно использовать для амплификации и модификации нуклеиновой кислоты, кодирующей человеческий, гуманизированный или химерный домен VH (в частности, для модификации первой (и необязательно второй) аминокислоты, кодируемой на N-конце домена VH). Преимущественно, универсальные праймеры, представленные в настоящем изобретении, могут в определенных комбинациях амплифицировать и модифицировать весь репертуар человеческих, гуманизированных или химерных тяжелых цепей, полученных из любого функционального генного сегмента VH, присутствующего в каждом семействе генов VH образца генома человека, без предвзятости. Эти универсальные праймеры способны генерировать такие модификации в любой панели выработанных человеческих, гуманизированных или химерных переменных областей, независимо от сегмента гена VH, который был рекомбинирован или перегруппирован с образованием переменной области. Эти универсальные праймеры способны генерировать такие модификации с из группы, состоящей из: использованием общего репертуара тяжелых цепей человека.

Настоящее изобретение предусматривает полипептид, содержащий переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, где переменный домен содержит N-концевую аминокислоту, выбранную аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и валина.

Соответственно, переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина может содержать N-концевую аминокислоту, выбранную из группы, состоящей из: аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана и тирозина.

Соответственно, переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина может содержать N-концевую аминокислоту, выбранную из группы, состоящей из: аргинина, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, серина, треонина, триптофана и тирозина.

Соответственно, переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина может содержать N-концевой аланин.

5 Соответственно, переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина может содержать N-концевую последовательность, выбранную из группы, состоящей из последовательностей: аланин-пролин, аланин-аспартат, аланин-глутамат, аланин-треонин, аланин-валин, аланин-серин и аланин-лейцин.

10 Соответственно, переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина может содержать N-концевую последовательность аланин-пролин.

Соответственно, полипептид может содержать сигнальный пептид выше по цепи N-концевой аминокислоты переменного домена тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина.

15 Соответственно, сигнальный пептид может содержать аминокислотную последовательность AQPAMA (SEQ ID NO: 5).

Настоящее изобретение также предусматривает антитело, вариант антитела или фрагмент антитела, содержащие описанный в настоящем изобретении полипептид.

20 Настоящее изобретение также предусматривает нуклеиновую кислоту, кодирующую полипептид, антитело, вариант антитела или фрагмент антитела, описанные в настоящем изобретении.

Кроме того, настоящее изобретение предусматривает вектор, включающий нуклеиновую кислоту, описанную в настоящем изобретении.

25 Настоящее изобретение также предусматривает матричный или стандартный вектор, содержащий последовательность нуклеиновой кислоты, кодирующую сигнальный пептид, и первую аминокислоту или первые две аминокислоты тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина согласно описанию в настоящем изобретении. Такой
30 матричный вектор не содержит нуклеиновой кислоты, кодирующей тяжелую цепь человеческого, гуманизированного или химерного иммуноглобулина.

Соответственно, вектор содержит последовательность нуклеиновой кислоты, кодирующую сигнальный пептид, и N-концевую аминокислоту, выбранную из группы, состоящей из: аланина, аргинина, аспарагина, аспартата,

глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и валина.

Соответственно, вектор содержит последовательность нуклеиновой кислоты, кодирующей сигнальный пептид и N-концевую последовательность, выбранную из группы, состоящей из последовательностей: аланин-пролин, аланин-аспартат, аланин-глутамат, аланин-треонин, аланин-валин, аланин-серин и аланин-лейцин.

Соответственно, вектор содержит последовательность нуклеиновой кислоты, кодирующей сигнальный пептид и N-концевую последовательность аланин-пролин.

Соответственно, сигнальный пептид, кодируемый последовательностью нуклеиновой кислоты в векторе, содержит аминокислотную последовательность AQRAMA (SEQ ID NO: 5).

Соответственно вектор может представлять собой фагмиду или плазмиду. Кроме того, настоящее изобретение предусматривает фаг, содержащий нуклеиновую кислоту, согласно описанному в настоящем изобретении.

Также в настоящем изобретении предусматривают библиотеку, содержащую по меньшей мере примерно 10^6 различных нуклеиновых кислот, векторов или фагов согласно описанному в настоящем изобретении.

Настоящее изобретение предусматривает способ одновременной амплификации и модификации нуклеиновой кислоты, кодирующей человеческий, гуманизированный или химерный переменный домен тяжелой цепи иммуноглобулина, причем способ включает:

(а) получение нуклеиновой кислоты, кодирующей человеческий, гуманизированный или химерный переменный домен тяжелой цепи иммуноглобулина; и

(б) осуществление полимеразной цепной реакции по меньшей мере с одним 5'-праймером, по меньшей мере с одним 3'-праймером и нуклеиновой кислотой для получения амплифицированной нуклеиновой кислоты,

причем по меньшей мере один 5'-праймер содержит нуклеиновую кислоту с сайтом модификации, который вносит модификацию в амплифицированную нуклеиновую кислоту, при этом амплифицированная нуклеиновая кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевую аминокислоту,

выбранную из группы, состоящей из: аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и валина.

Соответственно, каждая амплифицированная нуклеиновая кислота может кодировать переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевой аланин.

Соответственно, по меньшей мере один 5'-праймер может содержать нуклеиновую кислоту, которая вводит по меньшей мере две модификации в каждую амплифицированную нуклеиновую кислоту таким образом, что каждая амплифицированная нуклеиновая кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевую последовательность, выбранную из группы, состоящей из последовательностей: аланин-пролин, аланин-аспартат, аланин-глутамат, аланин-треонин, аланин-валин, аланин-серин и аланин-лейцин.

Соответственно, каждая амплифицированная нуклеиновая кислота может кодировать переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевую последовательность аланин-пролин.

Соответственно, по меньшей мере один 5'-праймер может кодировать сигнальный пептид или часть сигнального пептида выше по цепи от сайта модификации.

Соответственно, сигнальный пептид может содержать аминокислотную последовательность AQPAMA (SEQ ID NO:5).

Соответственно, нуклеиновая кислота (кислоты) на стадии (а) может быть кДНК.

Соответственно, способ может включать предварительную стадию выделения нуклеиновых кислот из В-клеток животного и получения кДНК из нуклеиновых кислот для получения нуклеиновой кислоты (кислот), предусмотренных на стадии (а).

Соответственно, животное может быть иммунизировано антигеном интереса, и нуклеиновые кислоты из В-клеток могут кодировать тяжелые цепи, обладающие специфичностью и аффинностью к антигену интереса.

Соответственно, животное может быть животным, отличным от человека, включая, но не ограничиваясь ими, мышь, крысу, кролика и курицу.

Соответственно, животное может быть животным, отличным от человека, причем трансгенным животным.

5 Соответственно, животное может быть трансгенным грызуном, содержащим локус тяжелой цепи человеческого или химерного иммуноглобулина.

Соответственно, животным может быть трансгенный грызун, кролик или курица, содержащие общую легкую цепь.

10 Соответственно, все нуклеиновые кислоты, которые кодируют переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина из В-клеток, могут быть модифицированы для кодирования:

15 - переменных доменов тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащих N-концевую аминокислоту, выбранную из группы, состоящей из: аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серин, треонин, триптофан, тирозин и валин; или

20 - переменных доменов тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащих N-концевую аминокислоту, выбранную из группы, состоящей из последовательностей: аланин-пролин, аланин-аспартат, аланин-глутамат, аланин-треонин, аланин-валин, аланин-серин и аланин-лейцин.

25 Соответственно, стадия (а) может включать предоставление множества различных нуклеиновых кислот, кодируемых или основанных на по меньшей мере одном сегменте рекомбинированного гена человека, выбранного из каждого из следующих семейств генов человека: IGHV1, IGHV2, IGHV3, IGHV4, IGHV5, IGHV6. и IGHV7.

Соответственно, способ может включать:

30 (а) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV1, с использованием 5'-праймера, выбранного из 1308AP, 1308AP2, 2020AP2, 2018AP или 2018AP2; и/или

(б) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV2, с использованием 5'-праймера, выбранного из 1310AP2, 1310AP3, 1310AP4 или 1310AP5; и/или

5 (в) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV3, с использованием 5'-праймера, выбранного из 0508AP, 0508AP2, 2018AP, 2018AP2, 2021AP, 2021AP2, 2021AP3, 2021AP4 или 2021AP5; и/или

10 (г) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV4, с использованием 5'-праймера, который представляет собой 1312AP; и/или

(д) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV5, с использованием 5'-праймера, выбранного из 1313AP или 1313AP2;

15 (е) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV6, с использованием 5'-праймера, выбранного из 1310AP2, 1310AP3, 1310AP4, 1310AP5 или 1312AP2; и/или

(ж) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV7, с использованием 5'-праймера, выбранного из 1314AP или 1314AP2.

20 Соответственно, способ может включать:

(а) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV1, с использованием 5'-праймера, выбранного из 1308AP2, 2020AP2 или 2018AP2; и/или

25 (б) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV2, с использованием 5'-праймера, который представляет собой 1310AP5; и/или

(в) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV3, с использованием 5'-праймера, выбранного из 0508AP; 2021AP2 или 2018AP2; и/или

30 (г) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV4, с использованием 5'-праймера, выбранного из 1312AP2 или 2019AP2; и/или

(д) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV5, с использованием 5'-праймера, который представляет собой 1313AP2;

5 (е) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV6, с использованием 5'-праймера, выбранного из 1310AP2, 1310AP3, 1310AP4, 1310AP5 или 1312AP2; и/или

(ж) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV7, с использованием 5'-праймера, который представляет собой 1314AP2.

10 Соответственно, модифицированные человеческие, гуманизированные или химерные переменные домены тяжелой цепи иммуноглобулина могут быть подвергнуты частотному анализу для идентификации лидерной последовательности.

15 Соответственно, способ может дополнительно включать введение каждой амплифицированной и модифицированной нуклеиновой кислоты в вектор.

Соответственно, вектор содержит последовательность нуклеиновой кислоты, кодирующую сигнальный пептид.

20 Соответственно, сигнальный пептид, кодируемый последовательностью нуклеиновой кислоты в векторе, содержит аминокислотную последовательность AQRAMA (SEQ ID NO: 5).

Соответственно, применимым вектором может быть фагида или плазида.

Соответственно, способ может дополнительно включать трансформацию или трансфекцию каждого вектора в клетку для создания библиотеки.

25 Соответственно, применимой клеткой может быть клетка, компетентная к фагу.

Соответственно, модифицированный человеческий, гуманизированный или химерный переменный домен тяжелой цепи иммуноглобулина может быть интегрирован в фаг для скрининга специфичности связывания и/или аффинности.

30 Соответственно, способ можно применять для снижения образования пироглутамата в переменном домене тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина.

Таким образом, настоящее изобретение также относится к способу снижения образования пироглутамата в переменном домене тяжелой цепи

человеческого, гуманизированного или химерного иммуноглобулина, причем способ включает: модификацию нуклеиновой кислоты, кодирующей переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, таким образом, чтобы модифицированная нуклеиновая кислота кодировала переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевую аминокислоту, выбранную из группы, состоящую из: аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и валина.

Соответственно, модифицированная нуклеиновая кислота может кодировать человеческий, гуманизированный или химерный переменный домен тяжелой цепи иммуноглобулина, содержащий N-концевой аланин.

Соответственно, модифицированная нуклеиновая кислота кодирует человеческий, гуманизированный или химерный переменный домен тяжелой цепи иммуноглобулина, содержащий N-концевую последовательность, выбранную из группы, состоящей из последовательностей: аланин-пролин, аланин-аспартат, аланин-глутамат, аланин-треонин, аланин-валин, аланин-серин и аланин-лейцин.

Соответственно, модифицированная нуклеиновая кислота может кодировать человеческий, гуманизированный или химерный переменный домен тяжелой цепи иммуноглобулина, содержащий N-концевую последовательность аланин-пролин.

Настоящее изобретение дополнительно относится к 5'-праймеру для амплификации и модификации любой нуклеиновой кислоты, которая кодирует человеческий, гуманизированный или химерный переменный домен тяжелой цепи иммуноглобулина, выбранный из одного или нескольких из следующих семейств генов VH человека или основанный на них: IGHV1, IGHV2, IGHV3, IGHV4, IGHV5, IGHV6 и IGHV7, где праймер содержит сайт модификации, который вводит модификацию в амплифицированную нуклеиновую кислоту таким образом, что амплифицированная нуклеиновая кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевую аминокислоту, выбранную из группы, состоящей из: аланина, аргинина, аспарагина, аспартата,

глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и валина.

Соответственно, сайт модификации может быть таким, что амплифицированная нуклеиновая кислота кодирует переменный домен

5 тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевой аланин.

Соответственно, сайт модификации может быть таким, что амплифицированная нуклеиновая кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного

10 иммуноглобулина, содержащий N-концевую последовательность, выбранную из группы, состоящей из последовательностей: аланин-пролин, аланин-аспартат, аланин-глутамат, аланин-треонин, аланин-валин, аланин-серин и аланин-лейцин.

Соответственно, сайт модификации может быть таким, что амплифицированная нуклеиновая кислота кодирует переменный домен

15 тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевую последовательность аланин-пролин.

Соответственно, праймер может кодировать сигнальный пептид или часть сигнального пептида выше по цепи сайта модификации.

Настоящее изобретение дополнительно предусматривает набор, содержащий по меньшей мере один 5'-праймер, выбранный из каждой из

20 следующих групп:

(а) 1308AP, 1308AP2, 2020AP2, 2018AP или 2018AP2;

(б) 1310AP2, 1310AP3, 1310AP4 или 1310AP5;

(в) 0508AP, 0508AP2, 2018AP, 2018AP2, 2021AP, 2021AP2, 2021AP3,

25 2021AP4 или 2021AP5;

(г) 1312AP2;

(д) 1313AP or 1313AP2;

(е) 1310AP2, 1310AP3, 1310AP4, 1310AP5 или 1312AP2;

(ж) 1314AP2 или 1314AP.

30 Соответственно, набор может содержать по меньшей мере один 5'-праймер, выбранный из каждой из следующих групп:

(а) 1308AP2, 2020AP2 или 2018AP2;

(б) 1310AP5;

(в) 0508AP, 2021AP2 или 2018AP2;

(г) 1312AP2 или 2019AP2;

(д) 1313AP2;

(е) 1310AP2, 1310AP3, 1310AP4, 1310AP5 или 1312AP2;

(ж) 1314AP2.

5 Настоящее изобретение дополнительно предусматривает способ получения антитела, варианта антитела или фрагмента антитела, включающий:

- модификацию нуклеиновой кислоты, которая кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина таким образом, что модифицированная нуклеиновая кислота
10 кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевую аминокислоту, выбранную из группы, состоящую из: аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и валина;

15 - последующее использование технологии скрининга антител для идентификации переменного домена тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащего N-концевую аминокислоту, выбранную из аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина,
20 серина, треонина, триптофана, тирозина и валина для связывания с антигеном-мишенью;

- отбор переменного домена тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, который связывает антиген-мишень; и

25 - применение указанного переменного домена тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина для разработки терапевтического антитела, варианта антитела или фрагмента антитела без дополнительной модификации N-концевой аминокислоты.

Соответственно, модифицированная нуклеиновая кислота может
30 кодировать переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий на N-конце аланин.

Соответственно, модифицированная нуклеиновая кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или

химерного иммуноглобулина, содержащий N-концевую последовательность, выбранную из группы, состоящей из последовательностей: аланин-пролин, аланин-аспартат, аланин-глутамат, аланин-треонин, аланин-валин, аланин-серин и аланин-лейцин.

5 Соответственно, модифицированная нуклеиновая кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевую последовательность аланин-пролин.

10 Соответственно, методика скрининга антител включает скрининг переменного домена тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина в соединении с легкой цепью.

15 В настоящем описании и формуле изобретения повсеместно слова «включать» и «содержать», а также их варианты, означают понятие «включая, но этим не ограничиваясь», и они не предназначены (и не исключают) другие фрагменты, добавки, компоненты, целые числа или стадии.

20 В настоящем описании и формуле изобретения повсеместно единственное число охватывает множественное число, если контекст не требует иного. В частности, когда используют неопределенный артикль, следует понимать как предполагаемую множественность, а также и единственность, если контекст не требует иного.

25 Признаки, целые числа, характеристики, соединения, химические фрагменты или группы, описанные в связи с конкретным объектом и вариантом осуществления настоящего изобретения или примером настоящего изобретения, следует понимать как применимые к любому другому объекту, варианту осуществления настоящего изобретения или примеру, приведенному в настоящем описании, при условии их совместимости.

Различные объекты настоящего изобретения подробнее описаны ниже.

Краткое описание фигур

30 Варианты осуществления настоящего изобретения далее описаны со ссылкой на прилагаемые фигуры:

Фигура 1. Образование пироглутаминовой кислоты (pyroGlu) с N-концов белков.

Фигура 2. Результаты анализа по прогнозированию расщепления сигнальной последовательности (SP) с использованием SignalP. Для каждой из

360 проанализированных последовательностей указана оценка D, а также средняя оценка (сред.) D для аминокислоты в VH в положении 1. Данные для последовательностей с прокариотической (P) SP приведены слева; данные для последовательностей с эукариотической (E) SP приведены справа. Справа от каждого набора данных в столбце «nat. частота (%)» перечислены частоты аминокислот в положении 1 для панели грамотрицательных и эукариотических SP-содержащих белков [Choo, Ranganathan, 2008].

Фигура 3. Результаты анализа по прогнозированию расщепления сигнальной последовательности (SP) с использованием SignalP для 1170 последовательностей. Для анализа 1170 последовательностей приводят оценку D, а также среднюю оценку (сред.) D для аминокислоты в VH в положении 2. Данные для последовательностей с прокариотической (P) SP приведены слева; данные для последовательностей с эукариотической (E) SP приведены справа. Справа от каждого набора данных в столбце «nat. частота (%)» перечислены частоты аминокислот в положении 1 для панели грамотрицательных и эукариотических SP-содержащих белков [Choo, Ranganathan, 2008].

Фигура 4. Результаты анализа по прогнозированию расщепления сигнальной последовательности (SP) с использованием SignalP. Для каждой из 18 комбинаций SP и начала VH даны оценки D для подмножества 1170 последовательностей с вариациями в положениях 1 и 2. Данные для последовательностей с прокариотической (P) SP приведены слева; данные для последовательностей с эукариотической (E) SP приведены справа. Для обеих SP также указана средняя (сред.) оценка D. Верхние строки показывают результаты для последовательностей дикого типа (WT). В средних строках показаны результаты для лучших вариантов, т. е. последовательностей с самыми высокими оценками D. Нижние строки показывают результаты для консенсуса лучших вариантов, т. е. последовательностей с оценкой D, которая выше, чем у последовательностей WT, как в комбинации с бактериальной SP, так и в комбинации с эукариотической SP; также приведены различия в оценках D по сравнению с WT (последняя строка).

Фигура 5. Выравнивание новых 5'-праймеров AP (первые два кодона VH выделены жирным шрифтом и подчеркнуты).

Фигура 6. Белковая трансляция новых разработанных 5'-праймеров, отжигаемых до начала сегментов гена VH, экспрессируемых двумя линиями

Merus Mouse (MeMo®). Первые два остатка областей VH, которые были заменены на AP, выделены жирным шрифтом и подчеркнуты.

Фигура 7. Эффективность амплификации для 5'-праймеров AP (сравнение выхода продуктов ПЦР на агарозном геле).

5 Фигура 8. Выравнивание новых праймеров AP и AP2.

Фигура 9. Эффективность амплификации для 5'-праймеров AP2 (сравнение выхода продуктов ПЦР на агарозном геле).

Фигура 10. Эффективность амплификации для всех новых праймеров (AP и AP2), тестируемых параллельно.

10 Фигура 11. Эффективность амплификации для 5 различных вариантов праймера 1310AP вместе с пятью различными вариантами праймера 2021AP.

Фигура 12. Выравнивание нового праймера 1308AP2 с функциональными последовательностями IGHV1.

15 Фигура 13. Выравнивание нового праймера 2020AP2 с функциональными последовательностями IGHV1.

Фигура 14. Выравнивание нового праймера 2018AP2 с функциональными последовательностями IGHV1.

Фигура 15. Выравнивание новых праймеров 1308AP2, 2018AP2 и 2020AP2 для IGHV1.

20 Фигура 16. Выравнивание нового праймера 1310AP5 со всеми функциональными последовательностями IGHV2.

Фигура 17. Выравнивание нового праймера 0508AP со всеми функциональными последовательностями IGHV3.

25 Фигура 18. Выравнивание нового праймера 2018AP2 со всеми функциональными последовательностями IGHV3.

Фигура 20. Выравнивание трех новых праймеров, специфичных в отношении IGHV3.

Фигура 21. Выравнивание нового праймера 1312AP2 со всеми функциональными последовательностями IGHV4.

30 Фигура 22. Выравнивание нового праймера 2019AP2 со всеми функциональными последовательностями IGHV4.

Фигура 23. Выравнивание нового праймера 1312AP2 со всеми функциональными последовательностями IGHV5.

Фигура 24. Выравнивание новых праймеров с IGHV6-1.

Фигура 25. Выравнивание нового праймера 1312AP2 с функциональным генным сегментом IGHV7.

Представленные в настоящем изобретении фигуры показывают выравнивание последовательностей конкретных генных сегментов VH в каждом семействе с соответствующими праймерами. Специалисту в данной области известно, что генные сегменты VH могут различаться по последовательности из-за аллельной изменчивости, и соответствующие праймеры для различных последовательностей генных сегментов VH в каждом семействе также относятся к настоящему изобретению и описаны в нем.

Патентная, научная и техническая литература, ссылки на которую приведены в описании настоящего изобретения, представляет знания, которые были доступны специалистам в данной области техники на момент подачи заявки. Полное раскрытие выданных патентов, опубликованных и находящихся на рассмотрении патентных заявок и других публикаций, которые цитируются в настоящем описании, включены в него в виде ссылок в той же степени, как если бы каждая из них была конкретно и отдельно указана для включения посредством ссылки. В случае каких-либо несоответствий настоящее описание будет иметь преимущественную силу.

Различные аспекты настоящего изобретения более подробно описаны ниже.

Подробное описание изобретения

Настоящее изобретение относится к полипептидам, содержащим модифицированные человеческие, гуманизированные или химерные переменные домены тяжелой цепи (VH) иммуноглобулина.

Термины «переменный домен тяжелой цепи иммуноглобулина» и «домен VH» в настоящем изобретении используют взаимозаменяемо. Термины обычно применяют в настоящем изобретении для обозначения человеческих, гуманизированных или химерных доменов VH (если в контексте конкретно не указано иное).

Термины «пептид», «белок» и «полипептид» в настоящем изобретении используют взаимозаменяемо. N-конец белковой последовательности (также известный как амино-конец, NH₂-конец или N-конец) является началом данной белковой последовательности. Традиционно пептидные последовательности записывают от N-конца к С-концу (слева направо). С-конец (также известный как карбоксильный конец, карбоксиконец, С-концевой хвост, С-конец или

СООН-конец) представляет собой конец аминокислотной цепи (белка или полипептида), оканчивающийся свободной карбоксильной группой (-СООН).

Модифицированный человеческий, гуманизированный или химерный переменный домен (также называемый в настоящем изобретении модифицированным доменом VH), описанный в настоящем изобретении, содержит аминокислотную модификацию по сравнению с обычным человеческим или родительским гуманизированным или химерным доменом VH, а именно, что N-концевая аминокислота модифицированного домена VH выбрана из группы, состоящей из: аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и валина. Это отличается от обычного человеческого или родительского гуманизированного или химерного домена VH, также называемого в настоящем изобретении немодифицированным доменом VH, поскольку обычные человеческие и гуманизированные или химерные родительские домены VH имеют остаток глутамин или глутамата на N-конце.

Используемый в настоящем описании термин «N-конец» домена VH (или «N-концевая аминокислота» домена VH) относится к началу аминокислотной последовательности домена VH (т.е. к первой аминокислоте (слева направо) домена VH), независимо от того, какие другие пептидные домены и последовательности могут присутствовать в составе полипептида. Таким образом несомненно «N-конец» домена VH (или «N-концевая аминокислота» домена VH) относится к первой аминокислоте аминокислотной последовательности зрелого домена VH, причем не следует учитывать любые расположенные выше по цепи аминокислоты, которые могут присутствовать в полипептиде как часть, например, последовательность сигнального пептида. Следовательно, «N-конец» домена VH (или «N-концевая аминокислота» домена VH) на самом деле может не находиться в начале полипептидной цепи (он может иметь другие аминокислотные остатки (остаток) впереди по цепи). Другими словами, для полипептидов, которые содержат сигнальный пептид непосредственно выше по цепи и рядом с аминокислотной последовательностью домена VH, «N-конец» домена VH относится к первой аминокислоте домена VH (т.е. первой аминокислоте после последовательности сигнального пептида). В качестве примера, не ограничивающего рамки охвата настоящего изобретения,

когда полипептид содержит сигнальный пептид
(MKYLLPTAAAGLLLLAAQPAMA (SEQ ID NO:1) + домен VH (QVQLVQSG
(SEQ ID NO:2)..... (согласно IGHV1-3*01_X62109.1_Homo)) как показано в
последовательности MKYLLPTAAAGLLLLAAQPAMAQVQLVQSG (SEQ ID
5 NO:3), N-концевая аминокислота домена VH подчеркнута.

Домены VH состоят из четырех каркасных областей и трех
гипервариабельных областей (также известных как CDR), имеющих сборку FR1-
CDR1-FR2-CDR2-FR3-CDR3-FR4 (от N-конца к С-концу). Каркасные области
составляют около 85% вариабельной области и действуют как каркас для CDR
10 домена VH. Каркасные области имеют меньшую вариабельность
аминокислотных последовательностей по сравнению с CDR. Первая
аминокислота области FR1 также является N-концевой аминокислотой домена
VH. Схема нумерации по Kabat широко используют для нумерации остатков в
последовательностях антител. (Кабат Е.А. с соавт., публикация NIH № 91-3242
15 (1991)).

Хотя полагают, что N-концевой остаток глутамина или глутамата
человеческого, гуманизированного или химерного домена VH играет важную
(или критическую) роль в антигенной аффинности, антигенной специфичности
и/или структурных взаимодействиях антитела, в настоящее время неожиданно
20 установили, что вариации аминокислот в положении возможны.
Преимущественно это позволяет модифицировать глутамин или глутамат другой
(предпочтительной) аминокислотой, такой как аланин, аргинин, аспарагин,
аспартат, глицин, гистидин, изолейцин, лейцин, лизин, метионин, фенилаланин,
пролин, серин, треонин, триптофан, тирозин или валин. Модификация N-
25 концевой аминокислоты человеческого, гуманизированного или химерного
домена VH таким образом устраняет образование пироглутамата (и, таким
образом, позволяет избежать потенциальных вредных эффектов, связанных с
образованием пироглутамата, которые подробно обсуждаются в настоящем
описании).

30 Аминокислоты могут быть сгруппированы в соответствии с их
биохимическими свойствами (например, зарядом, гидрофобностью, размером и
т. д.). Например, кислотные остатки содержат аспартат и глутамат. К примерам
некислотных остатков с полярными боковыми цепями относятся аспарагин и
глутамин. Таким образом, в одном из примеров N-концевой остаток глутамина

или глутамата человеческого, гуманизированного или химерного домена VH заменен кислотным или полярным остатком, таким как аспартат или аспарагин. Эти аминокислоты имеют сходные биохимические свойства с глутамином или глутаматом и, таким образом, могут быть полезным выбором, поскольку такие

5 изменения сохраняют сходные биохимические свойства, устраняя потенциальную изменчивость нижестоящих по цепи переменных областей тяжелой цепи с течением времени, которая может возникнуть при переходе от глутамина или глутамата к пироглутамату.

Возможное воздействие каждой аминокислотной вариации (т.е. N-концевого аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина или валина) на распознавание сайта отщепления

10 сигнального пептида выше по цепи от человеческого, гуманизированного или химерного домена VH проанализирован в настоящем описании. Обнаружено, что модификация N-концевого остатка глутамата или глутамина в человеческих, гуманизированных или химерных доменах VH с помощью аланина особенно полезна, поскольку она устраняет образование пироглутамата, а также поддерживает (например, улучшает) распознавание сайта расщепления

15 сигнального пептида у прокариот и эукариот. Другие аминокислоты, которые устраняют образование пироглутамата, поддерживая при этом относительное распознавание сайта отщепления сигнального пептида у прокариот на основании анализа *in silico* среднего показателя D, представляют собой глицин, метионин, аспарагин, серин, треонин, валин и тирозин. Другие аминокислоты, которые

20 устраняют образование пироглутамата, сохраняя при этом относительное распознавание сайта расщепления сигнального пептида у эукариот, на основании анализа среднего показателя D *in silico*, включают фенилаланин, изолейцин, лейцин, валин и триптофан. Таким образом, каждая из этих аминокислот также применима для модификации N-концевого глутамата или остатка глутамина.

25 Оценка устранения образования пироглутамата при сохранении относительного распознавания сайта отщепления сигнального пептида, как у прокариот, так и у эукариот на основе частоты немодифицированных остатков показывает, что предпочтительными остатками являются аланин, аспарагиновая кислота и серин.

30

Аланин является алифатическим остатком. Соответственно, в другом примере N-концевой остаток глутамин или глутамат человеческого,

гуманизированного или химерного домена VH заменен алифатическим остатком, таким как аланин, глицин, валин, лейцин или изолейцин.

В настоящем изобретении также исследуют результаты замены второй аминокислоты на N-конце домена VH другой (предпочтительной)

5 аминокислотой. Без сомнений, «вторая аминокислота на N-конце домена VH» относится к аминокислоте, непосредственно примыкающей (в направлении от N-конца к С-концу) к N-концевой аминокислоте домена VH (другими словами, аминокислота в положении два в аминокислотной последовательности домена VH, где N-концевая аминокислота (такая как Q или E в немодифицированных

10 человеческих, гуманизированных или химерных доменах VH) находится в положении один).

Неожиданно было обнаружено, что замена второй аминокислоты на N-конце домена VH (в направлении от N-конца к С-концу) на один пролин, аспартат, глутамат, треонин, валин, серин или лейцин способствует отщеплению

15 сайта расщепления сигнального пептида выше по цепи домена VH на основе немодифицированной частоты таких остатков в зрелых белках грамотрицательных бактерий, содержащих сигнальный пептид (Choo, Ranganathan, 2008). Это особенно значимо, когда пролин выбран в качестве второй аминокислоты на N-конце домена VH.

20 Соответственно, модифицированный человеческий, гуманизированный или химерный домен VH, описанный в настоящем изобретении, может содержать первую N-концевую аминокислоту, выбранную из группы, состоящей из: аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина,

25 триптофана, тирозина и валина; и вторую аминокислоту, выбранную из группы, состоящей из пролина, аспартата, глутамата, серина, треонина, валина или лейцина. В одном конкретном примере второй аминокислотой на N-конце домена VH выбран пролин, причем предпочтительное первое и второе положения включают «аланин-пролин» или «АР». Предпочтительно в настоящем

30 изобретении избегают остатка цистеина, поскольку он вводит высокореакционноспособную группу на N-конце, что может вызвать проблемы при получении и хранении.

Как показано в представленных ниже данных, особенно полезными комбинациями аминокислот на N-конце модифицированного домена VH

являются аланин-пролин, аланин-аспартат, аланин-глутамат, аланин-серин, аланин-треонин, аланин-валин и аланин-лейцин. Форматирование, используемое в настоящем изобретении для «аланин-пролин» или «AP» и т. д., относится к двум соседним аминокислотам на N-конце модифицированного домена VH человека (т.е. «первая-вторая» аминокислоты в N -конец домена VH (в направлении от N-конца к C-концу)).

Иные предпочтительные комбинации выведены из анализа прогнозирования расщепления SP с использованием signalP. На основании данных, представленных на фиг. 3, другие предпочтительные комбинации включают каждую из следующих: глицин-пролин, глицин-аспартат, глицин-глутамат, глицин-серин, глицин-треонин, глицин-лейцин, глицин-валин, метионин-пролин, метионин-аспартат, метионин-глутамат, метионин-серин, метионин-треонин, метионин-лейцин, метионин-валин, аспарагин-пролин, аспарагин-аспартат, аспарагин-глутамат, аспарагин-серин, аспарагин-треонин, аспарагин-лейцин, аспарагин-валин, серин- пролин, серин-аспартат, серин-глутамат, серин-серин, серин-треонин, серин-лейцин, серин-валин, треонин-пролин, треонин-аспартат, треонин-глутамат, треонин-серин, треонин-треонин, треонин-лейцин, треонин-валин, валин-пролин, валин-аспартат, валин-глутамат, валин-серин, валин-треонин, валин-лейцин, валин-валин, тирозин-пролин, тирозин-аспартат, тирозин-глутамат, тирозин-серин, тирозин- лейцин, тирозин-валин и тирозин-треонин.

В настоящем изобретении исследованы эффекты изменения первых двух аминокислот на N-конце человеческого, гуманизированного или химерного домена VH на аланин-пролин. Обнаружено, что эта комбинация особенно выгодна для повышения эффективности отщепления сигнального пептида. Таким образом, описанное настоящее изобретение особенно выгодно при использовании с модифицированным человеческим, гуманизированным или химерным доменом VH, который экспрессируется вместе с сигнальным пептидом, расположенным выше по цепи в клетке. В этом контексте «выше по цепи» относится к расположению сигнального пептида в полипептиде относительно модифицированного домена VH. Другими словами, при рассмотрении полипептида в целом сигнальный пептид находится выше модифицированного домена VH, когда он ближе к N-концу полипептида, чем аминокислотная последовательность модифицированного домена VH.

Специалистам в данной области известно несколько различных сигнальных пептидов. Используемый в настоящем изобретении термин «сигнальный пептид» относится к лидерной последовательности, обеспечивающей вход в секреторный метаболический путь. Сигнальный пептид представляет собой относительно короткий пептид, расположенный на N-конце секреторных белков, которые направляют белок в просвет эндоплазматического ретикулума для последующего экспорта из клетки. Например, у эукариот сигнальный пептид, содержащий 5-30 аминокислот, присутствующих на N-конце образующихся белков, распознается частицей распознавания сигнала (SRP – signal recognition particle) в цитозоле, пока белок все еще синтезируется на рибосоме. Затем SRP доставляет комплекс SRP-рибосома-зарождающаяся цепь (SRP-RNC) к SRP-рецептору (SR) в мембране эндоплазматического ретикулума (ER). Затем GTP-зависимые механизмы доставляют комплекс RNC к мембраносвязанному транслокону, который допускает транслокацию растущей полипептидной цепи в просвет ER. После пересечения мембраны ER сигнальный пептид отщепляется пептидазой сигнального пептида (SPP – signal peptide peptidase). Ряд сигнальных пептидов, подходящих для экспрессии белков в эукариотических клетках, известны специалистам в данной области. N-концы тяжелой и легкой цепей иммуноглобулина имеют нативные сигнальные пептиды или лидерные последовательности (*Nucl. Acids Res.* (2005), 33, D256-D261). Другие последовательности сигнальных пептидов хорошо известны в данной области, они перечислены в *Nucleic Acid Research* (1984) 12, 5145 – 5164. Примеры сигнальных пептидов мембранных белков и секреторных белков включают сигнальные пептиды, полученные из *Saccharomyces cerevisiae* или вирусов дрожжей, которые используют в обычных системах экспрессии мембранных и секреторных белков в дрожжах, включая секреторные сигнальные пептиды, полученные из α -фактора, рецептора α -фактора, препротоксина, белков SUC2 и белков PHO5, белков BGL2 и белков AGA2. Были предоставлены компьютерные программы, которые прогнозируют последовательности секреторных сигнальных пептидов. SignalP (<http://www.cbs.dtu.dk/services/SignalP/>), PSORT (<http://psort.nibb.ac.jp/>) и Phobius (<http://phobius.cgb.ki.se>). Использование этих компьютерных программ позволяет прогнозировать последовательность секреторного сигнала.

Конкретным примером сигнального пептида, который можно использовать для экспрессии домена VH в эукариотических клетках (который используется в разделе примеров ниже), является MGWSCILFLVLLLAQPAMA (SEQ ID NO:4). Следует отметить, что этот пример не ограничивает рамок охвата настоящего изобретения, поскольку SRP избирательно связывают сигнальные пептиды на основе их общих признаков, несмотря на вариабельность их первичной последовательности. Соответственно, в настоящем изобретении также могут быть использованы другие соответствующие сигнальные пептиды.

Большинство белков, которые транспортируются во внецитоплазматическую среду у бактерий, используют для нацеливания так называемый путь Sec. Этот путь инициируется, когда сигнальный пептид на вновь синтезированном белке-предшественнике распознается SecA, белком, который встречается только в прокариотах и органеллах прокариотического происхождения, таких как митохондрии.

Конкретным примером сигнального пептида, который можно использовать для экспрессии домена VH в прокариотических клетках (который используют в разделе примеров ниже), является MKYLLPTAAAGLLLLAAQPAMA (SEQ ID NO:1). Следует отметить, что этот пример не ограничивает рамок охвата настоящего изобретения, поскольку SecA избирательно связывает сигнальные пептиды на основе их общих признаков, несмотря на вариабельность их первичной последовательности. Соответственно, здесь также могут быть использованы другие подходящие сигнальные пептиды. Репрезентативные сигнальные последовательности включают сигнальные последовательности из PelB, OmpA, PhoA, эндоксилаказы и StII (*Appl. Microbiol. Biotechnol* (2004) 64:625-635).

Следует отметить, что два (не ограничивающих рамок охвата настоящего изобретения) сигнальных пептида, использованных в приведенном ниже разделе примеров, имеют последовательность AQPAMA (SEQ ID NO:5) на С-конце сигнального пептида (другими словами, аминокислотной последовательностью, расположенной непосредственно выше по цепи N-конца домена VH, является ...AQPAMA (SEQ ID NO:5)). Соответственно, в одном примере сигнальный пептид содержит аминокислотную последовательность AQPAMA (SEQ ID NO:5).

Выше указано, что эффективность расщепления сигнального пептида зависит от последовательности сигнального пептида в месте расщепления сигнального пептида, а также от последовательности VH. Таким образом, остатки домена VH и сигнального пептида, граничащие с сайтом расщепления сигнального пептида, могут влиять на процессинг сигнальной пептидазы и вносить вклад в неканонические сайты расщепления. Таким образом, предпочтительно эти остатки представляют собой **AQPMAA** (SEQ ID NO:6) или **AQPMAA**P (SEQ ID NO:7) (где последовательность на С-конце сигнального пептида выделена жирным шрифтом, а N-концевая аминокислота модифицированного человеческого, гуманизированного или химерного домена VH подчеркнута).

В другом примере, специально для экспрессии в прокариотических клетках-хозяевах, аминокислотными остатками домена VH и сигнального пептида, которые фланкируют сайт расщепления сигнального пептида, могут быть **MKYLLPTAAAGLLLLAAQPMAA** (SEQ ID NO:8) или **MKYLLPTAAAGLLLLAAQPMAA**P (SEQ ID NO:9) (где последовательность на С-конце сигнального пептида выделена жирным шрифтом, а N-концевая аминокислота модифицированного человеческого, гуманизированного или химерного домена VH подчеркнута).

В другом примере, специально для экспрессии в эукариотических клетках-хозяевах, аминокислотными остатками домена VH и сигнального пептида, которые фланкируют сайт расщепления сигнального пептида, могут быть **MGWSCILFLVLLLAQPMAA** (SEQ ID NO:10) или **MGWSCILFLVLLLAQPMAA**P (SEQ ID NO:11) (где последовательность на С-конце сигнального пептида выделена жирным шрифтом, а N-концевая аминокислота модифицированного человеческого, гуманизированного или химерного домена VH подчеркнута).

Описанные в настоящем изобретении полипептиды содержат модифицированный домен VH. Полипептид может быть любым белком, который включает описанный здесь модифицированный домен VH. Например, это может быть антитело, вариант антитела или фрагмент антитела, который включает описанный здесь модифицированный домен VH.

Антитела, варианты антител и фрагменты антител

Термин «антитело» означает белковую молекулу, принадлежащую к классу иммуноглобулинов, содержащих один или несколько доменов, которые связывают эпитоп на антигене, причем такие домены происходят из

5 вариабельной области антитела или их последовательности гомологичны вариабельной области антитела. Связывание антител проявляется по-разному, включая специфичность и аффинность. Специфичность определяет, какой антиген или его эпитоп специфически связывается со связывающим доменом. Аффинность представляет собой силу связывания с определенным антигеном или эпитопом. Следует отметить, что «специфичность» антитела относится к его

10 селективности в отношении определенного антигена, тогда как «аффинность» относится к силе взаимодействия между антигенсвязывающим сайтом антитела и эпитопом, который оно связывает. Таким образом, термин «специфичность связывания» в настоящем изобретении относится к способности сайта связывания отдельного антитела реагировать с антигенной детерминантой.

15 Обычно сайт связывания антитела по настоящему изобретению расположен в участках Fab и построен из гипервариабельных областей тяжелой и легкой цепей.

В настоящем описании термин «фрагмент антитела» относится к белковому фрагменту, содержащему функциональную часть антитела (в данном случае, по

20 меньшей мере, модифицированный домен VH, описанный в настоящем изобретении). Фрагмент антитела может быть любым связывающим агентом, включая, но ими не ограничиваясь, одноцепочечные Fv, одноцепочечные или тандемные диатела (TandAb®), VHH, Anticalins®, Nanobodies®, BiTE®, Fab, белки с анкириновыми повторами или DARPINs®, Avimers®, DART, TCR-

25 подобное антитело, Adnectins®, Affilins®, Trans-bodies®, Affibodies®, TrimerX®, MicroProteins, Fynomers®, Centyrins® или KALBITOR®.

Термин «аффинность» означает силу взаимодействия между одним сайтом связывания антигена и его антигеном. Один антигенсвязывающий сайт антитела по настоящему изобретению для антигена может быть выражен в терминах

30 константы равновесной диссоциации (K_d), также известной как константа аффинности. Обычно антитела для терапевтического применения могут иметь аффинность со значениями K_d в диапазоне от микромолярного (10^{-6} М; низкое сродство) до пикомолярного (10^{-12} М; высокое сродство).

Термин «антиген» означает молекулу, способную индуцировать иммунный ответ (вырабатывать антитело) в организме-хозяине и/или являться мишенью для антитела. На молекулярном уровне антиген отличается способностью связываться с антигенсвязывающим участком антитела. Также смеси антигенов могут рассматриваться как «антиген», но специалисту в данной области очевидно, что иногда лизат опухолевых клеток или вирусные частицы могут быть обозначены как «антиген», хотя такой лизат опухолевых клеток или препарат вирусных частиц состоит из многих антигенных детерминант. Антиген содержит по меньшей мере один, но часто несколько эпитопов.

Понятие «эпитоп» или «антигенная детерминанта» означает участок на антигене, с которым специфически связывается иммуноглобулин или антитело. Эпитопы могут быть образованы из смежных аминокислот или несмежных аминокислот, но расположенных рядом за счет третичной укладки белка (так называемые линейные и конформационные эпитопы, соответственно). Эпитопы, образованные из смежных линейных аминокислот, обычно сохраняются при воздействии денатурирующих растворителей, тогда как конформация эпитопов, образованных третичной укладкой, обычно теряется при обработке денатурирующими растворителями. Эпитоп обычно может включать 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 или 15 аминокислот в уникальной пространственной конформации.

Термин «тяжелая цепь» или «тяжелая цепь иммуноглобулина» означает последовательность константной области тяжелой цепи иммуноглобулина из какого-либо организма и, если не указано иное, включает переменный домен тяжелой цепи (VH). Переменные домены тяжелой цепи включают три области CDR тяжелой цепи и четыре каркасных (FR) области, если не указано иное. Фрагменты тяжелых цепей включают CDR и FR и их комбинации. Типичная тяжелая цепь имеет после переменного домена (от N-конца к С-концу) домен CH1, шарнир, домен CH2 и домен CH3. Функциональный фрагмент тяжелой цепи включает фрагмент, способный специфически распознавать антиген и содержащий по меньшей мере одну CDR.

Термин «легкая цепь» включает переменный домен легкой цепи иммуноглобулина или VL (или его функциональный фрагмент); и последовательность константного домена иммуноглобулина или CL (или его функционального фрагмента) из какого-либо организма. Если не указано иное,

термин «легкая цепь» может включать легкую цепь, выбранную из каппа- или лямбда-цепи человека и их комбинации. Вариабельные домены легкой цепи (VL) обычно включают три CDR легкой цепи и четыре области FR, если не указано иное. Обычно полноразмерная легкая цепь включает от N-конца к С-концу домен VL, который содержит FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4, и константный домен легкой цепи. Легкие цепи, которые можно использовать в настоящем изобретении, включают такие, например, которые селективно не связывают эпитоп, селективно связываемый тяжелыми цепями.

Соответствующие легкие цепи для использования в антителе по настоящему изобретению включают общую легкую цепь (cLC – common light chain), из числа таких, которые можно идентифицировать путем скрининга наиболее часто используемых легких цепей в существующих библиотеках антител (влажных библиотеках или *in silico*), где легкие цепи существенно не влияют на аффинность и/или селективность эпитоп-связывающих доменов тяжелых цепей, но также подходят для сопряжения с набором тяжелых цепей. Например, соответствующей легкой цепью является цепь из трансгенных животных, такая как MeMo®, имеющая общую легкую цепь, интегрированную в ее геном, которую можно использовать для создания больших панелей антител с общей легкой цепью, которые различны по тяжелой цепи и способны к специфическому связыванию антигена при воздействии указанного антигена.

Термин «общая легкая цепь» относится к легким цепям, которые могут быть идентичными или иметь некоторые отличия в аминокислотной последовательности, при этом специфичность связывания антитела по настоящему изобретению не затрагивается, т.е. различия не оказывают существенного влияния на образование функциональных участков связывания. Например, в рамках определения представления об общих цепях, используемые в настоящем изобретении, для получения или выявления вариабельных цепей, которые не идентичны, но все же функционально эквивалентны, например, путем введения и тестирования консервативных аминокислотных замен, замен аминокислот в областях, которые не способствуют или только частично способствуют специфичности связывания с родственной цепью, и т.п. Таким образом, такие варианты также способны связывать различные родственные цепи и образовывать функциональные антигенсвязывающие домены. Термин «общая легкая цепь», используемый в настоящем описании, таким образом,

относится к легким цепям, которые могут быть идентичными или иметь некоторые отличия в аминокислотной последовательности, сохраняя при этом специфичность связывания полученного антитела после спаривания с тяжелой цепью. Комбинация определенной общей легкой цепи и таких функционально эквивалентных вариантов охватывается термином «общая легкая цепь».

Подробнее описание применения и соответствующих распространенных легких цепей приведено в WO 2004/009618, WO 2009/157771 и WO 2020/141973.

Обозначение «Fab» означает связывающий домен, содержащий переменную область, обычно связывающий домен, содержащий парный переменный домен тяжелой цепи и переменный домен легкой цепи. Fab может содержать домены константной области, включая домены CH1 и VH, соединенные с константным доменом легкой цепи (CL) и доменом VL. Такое спаривание может иметь место, например, как ковалентная связь через дисульфидный мостик в доменах CH1 и CL.

«Одноцепочечный переменный фрагмент» (scFv – single-chain variable fragment) означает связывающий домен, содержащий домен VH и домен VL, которые соединены через линкер, например пептидный линкер, например, длиной от примерно 10 до примерно 25 аминокислот. В настоящем описании термин «связанные» относится к доменам, которые соединены друг с другом посредством своей первичной аминокислотной последовательности. Например, базовая часть антитела может быть соединена с дополнительным связывающим доменом (или дополнительный связывающий домен может быть соединен с дополнительным связывающим доменом) через линкер. Точно так же домен CH1 может быть связан с переменной областью тяжелой цепи, а домен CL может быть связан с переменной областью легкой цепи. Кроме того, понятие «спаривание» относится к взаимодействиям между полипептидами по настоящему изобретению таким образом, что они могут формировать мультимеры. Например, дополнительный связывающий домен может содержать участок тяжелой цепи (CH1-VH), соединенный с участком легкой цепи (CL-VL), где CH1 (участок тяжелой цепи) и CL (участок легкой цепи) спариваются обычно за счет образования дисульфидной связи. Домены цепей антител или полипептиды, такие как домен смешанного связывания, могут дополнительно взаимодействовать и образовывать пары для формирования области контакта (интерфейса) посредством ковалентных или нековалентных взаимодействий,

например, посредством сил Ван-дер-Ваальса, водородных связей, опосредованных водой водородных связей, солевых мостиков или других электростатических сил, притяжения между ароматическими боковыми цепями, образования дисульфидных связей или других сил, известных специалистам в данной области.

Термин «процент (%) идентичности» применительно к последовательностям нуклеиновых кислот или аминокислот в настоящем изобретении означает процент остатков в последовательности-кандидате, которые идентичны остаткам в выбранной последовательности, после выравнивания последовательностей для целей оптимального сравнения. Процент идентичности последовательности при сравнении нуклеиновых кислот определяют с помощью приложения AlignX программного обеспечения Vector NTI Program Advance 11.5.2 с настройками по умолчанию, в которых применяется модифицированный алгоритм ClustalW (Thompson JD, Higgins DG, Gibson TJ (1994) Nuc Acid Res.22: 4673-4680), матрицы оценок swgapdnamt, штраф за открытие гэпа 15 и штраф за продолжение гэпа 6,66. Аминокислотные последовательности выравнивают с помощью приложения AlignX программного обеспечения Vector NTI Program Advance 11.5.2 с использованием настроек по умолчанию, в которых используется модифицированный алгоритм ClustalW (Thompson JD, Higgins DG, Gibson TJ, 1994), матрица оценок blosum62mt2, штраф за открытие гэпа 10 и штраф за продолжение гэпа 0,1.

Термин «множество» означает два или более.

Термин «вариант» антитела по настоящему изобретению может включать функциональную часть, функциональное производное, производное и/или аналог антитела. К нему относятся миметики антител, моноклоналы и аптамеры. Вариант обычно сохраняет специфичность связывания антитела, например специфичность биспецифического антитела. Функциональное производное антитела по настоящему изобретению представляет собой белок, содержащий переменный домен, который связывает одну мишень, и переменный домен, который связывает вторую мишень, соединенные связывающей областью. Переменные домены могут быть переменными доменами как таковыми, или Fab-фрагментами, или молекулами, подобными таким переменным доменам, как одноцепочечные Fv-фрагменты (scFv), содержащие VH и VL, связанные друг с другом через линкер. Другими примерами молекул, подобных переменным

доменам, являются так называемые однодоменные фрагменты антител. Фрагмент однодоменного антитела (sdAb) представляет собой фрагмент антитела с одной мономерной переменной областью антитела. Подобно целому антителу, оно способно избирательно связываться со специфическим антигеном. При молекулярной массе всего 12–15 кДа, фрагменты однодоменных антител

намного меньше, чем обычные антитела (150–160 кДа), состоящие из двух тяжелых белковых цепей и двух легких цепей, и даже меньше, чем Fab-фрагменты (~50 кДа, одна легкая цепь и половина тяжелой цепи), и одноцепочечные переменные фрагменты 25 кДа, две переменные области, одна из легкой и одна из тяжелой цепи). Однодоменные антитела сами по себе не намного меньше нормальных антител (обычно 90–100 кДа). Фрагменты однодоменных антител в основном создаются из антител с тяжелой цепью, обнаруженных у верблюдовых; они называются фрагментами VHH (Nanobody®). У некоторых рыб также есть антитела только с тяжелой цепью (IgNAR, «иммуноглобулиновый новый антигенный рецептор»), из которых можно получить однодоменные фрагменты антител, называемые фрагментами VNAR. Альтернативный подход заключается в разделении димерных переменных доменов общего иммуноглобулина G (IgG) из человека или мыши на мономеры. Хотя большинство исследований однодоменных антител в настоящее время основано на переменных доменах тяжелой цепи, также было показано, что нанотела, полученные из легких цепей, специфически связываются с эпитопами-мишенями. Другими неограничительными примерами молекул, подобных переменным доменам, являются VHH, антитела к человеческому домену (dAbs) и Unibodies. Предпочтительными функциональными частями являются части, содержащие переменные домены, содержащие переменную область тяжелой цепи и переменную область легкой цепи. Неограничительными примерами таких переменных доменов являются F(ab)-фрагменты и одноцепочечные Fv-фрагменты. Биспецифические форматы для связывания (подобного переменной области домену) представляют собой, например, человеческий сывороточный альбумин (HSA), связанный с двумя разными scFv; биспецифические мини-антитела, содержащие два различных фрагмента scFv, связанных друг с другом с помощью мотивов димеризации или самоассоциирующихся вторичных структур, таких как спиральные пучки или спиральные клубки, для обеспечения димеризации фрагментов scFv (Morrison.

Nat. Biotechnol. 2007, 25:1233-34). Примеры подходящих линкеров HSA и способ связывания scFv с линкером описаны в WO 2009/126920.

Функциональное производное может представлять собой миметик антитела, полипептид, аптамер или их комбинацию. Эти белки или аптамеры обычно связываются с одной мишенью. Белок по изобретению связывается с двумя или более мишенями. Следует учитывать, что любая комбинация этих антител, миметиков антител, полипептидов и аптамеров может быть связана вместе способами, известными в данной области. Например, в некоторых вариантах осуществления настоящего изобретения связывающая молекула по настоящему изобретению представляет собой конъюгат или слитый белок. Миметик антитела представляет собой полипептид, который, как и антитела, может специфически связываться с антигеном, но структурно не родственен антителам. Миметики антител обычно представляют собой искусственные пептиды или белки с молекулярной массой примерно от 3 до 20 кДа. Общими преимуществами по сравнению с антителами являются лучшая растворимость, проникновение в ткани, устойчивость к нагреванию и ферментам и сравнительно низкие затраты на получение. Неограничивающими примерами миметиков антител являются молекулы аффитела (обычно основанные на домене Z белка A); аффилины (обычно на основе кристаллического гамма-В или убиквитина); аффирмеры (обычно на основе цистатина); аффитины (обычно основанные на Sac7d из *Sulfolobus acidocaldarius*); альфа-тела (обычно на основе спиральной катушки с тройной спиралью); антикарины (обычно на основе липокалинов); авимеры (обычно на основе А-доменов различных мембранных рецепторов); DARPins (обычно на основе мотива анкиринового повтора); финомеры (обычно основанные на домене SH3 Fyn 7); пептиды домена Кунитца (обычно основанные на доменах Кунитца различных ингибиторов протеазы); и монотела (обычно основанные на домене типа III фибронектина).

Монотела представляют собой синтетические связывающие белки, которые сконструированы с использованием домена фибронектина III типа (FN3) в качестве молекулярного каркаса. Монотела представляют собой простую и надежную альтернативу антителам для создания белков, связывающихся с мишенью. Термин «монотело» был введен в 1998 году группой Koide, опубликовавшей первую статью, демонстрирующую концепцию монотела с использованием десятого домена FN3 человеческого фибронектина. Монотела и

другие миметики антител обычно создаются из комбинаторных библиотек, в которых части каркаса диверсифицированы с использованием технологий молекулярного дисплея и направленной эволюции, таких как фаговый дисплей, дисплей мРНК и дисплей на поверхности дрожжей. Большое количество миметиков антител обладают высокой аффинностью и высокой специфичностью в отношении своих соответствующих мишеней. Аптамеры представляют собой молекулы олигонуклеотидов или пептидов, которые связываются с определенной молекулой-мишенью. Аптамеры обычно создаются путем их выбора из большого пула случайных последовательностей, но природные аптамеры также существуют в рибопереключателе. Аптамеры можно использовать как для фундаментальных исследований, так и для клинических целей в виде макромолекул. Взаимодействия «без связи» действуют между атомами, не связанными ковалентными связями. Соответственно, это связи, которые не включают в себя совместное использование электронов, а скорее включают более рассредоточенные вариации электромагнитных взаимодействий между молекулами или внутри молекулы. Взаимодействия «без связи» включают электростатические взаимодействия, такие как водородные связи, ионные взаимодействия и галогенные связи. Силы Ван-дер-Ваальса представляют собой подмножество электростатических взаимодействий с участием постоянных или индуцированных диполей (или мультиполей). К ним относятся следующие: постоянные диполь-дипольные взаимодействия, дипольно-индуцированные дипольные взаимодействия и индуцированные диполь-индуцированные дипольные взаимодействия. Солевые мостики представляют собой комбинацию двух нековалентных взаимодействий: водородной связи и ионной связи. Гидрофобные взаимодействия представляют собой взаимодействие неполярных (неионизируемых) молекул углеводов, вынужденных быть вместе из-за более сильного взаимодействия вода:вода.

Нуклеиновые кислоты и векторы

Настоящее изобретение также предусматривает нуклеиновые кислоты, кодирующие полипептид, вариант антитела или фрагмент антитела по настоящему изобретению.

Описанные здесь нуклеиновые кислоты можно использовать для получения полипептида, антитела, варианта антитела или фрагмента антитела по

настоящему изобретению. Соответственно, также предусматривают векторы (например, векторы экспрессии), содержащие такие нуклеиновые кислоты, которые можно использовать для получения полипептида, антитела, варианта антитела или фрагмента антитела по настоящему изобретению.

5 Антитела обычно вырабатываются клетками, содержащими нуклеиновые кислоты, которые кодируют полипептиды; полипептиды объединяются с образованием антитела. Нуклеиновые кислоты, используемые для получения полипептидов антитела, могут быть помещены в любой подходящий вектор экспрессии, и при соответствующих обстоятельствах два или более вектора
10 могут быть помещены в одну клетку-хозяина. Обычно нуклеиновые кислоты, кодирующие модифицированные домены VH, могут быть клонированы с соответствующими линкерами и/или константными областями, а последовательности расположены в оперативной связи с промотором в соответствующей экспрессирующей конструкции в соответствующей линии
15 клеток, предназначенной для экспрессии.

 Векторная ДНК, в которую может быть введена нуклеиновая кислота, кодирующая модифицированный домен VH, например, клонированием или путем синтеза, предпочтительно содержит нуклеиновую кислоту, кодирующую
20 сигнальный пептид, и первую или первые две аминокислоты домена VH. Таким образом, векторная ДНК может использоваться в качестве стандартного вектора для получения модифицированных антител или тяжелых цепей согласно описанию настоящего изобретения, тем самым исключая необходимость изменения первой или первых двух аминокислот каждого отдельного домена VH. Таким образом, эта векторная ДНК еще не содержит нуклеиновую кислоту,
25 кодирующую модифицированный домен VH. Специалист в данной области знает, как ввести нуклеиновую кислоту, кодирующую модифицированный домен VH, в такую векторную ДНК, чтобы получить функциональное антитело или тяжелую цепь, например, путем исключения кодонов, кодирующих одну первую или первые две аминокислоты нуклеиновой кислоты, кодирующей
30 модифицированный домен VH, которые уже присутствуют в векторной ДНК.

 Вектор может быть любым соответствующим вектором, например фагмидой (для экспрессии в фаге) или плазмидой (для экспрессии в бактериальной или эукариотической клетке).

Фаганда представляет собой вектор клонирования на основе ДНК, обладающий свойствами, как бактериофага, так и плазмиды. Фагмиды несут ориджин репликации плазмиды и ориджин репликации, полученный из бактериофага. Фагмиды могут использоваться в качестве типа вектора клонирования в комбинации с нитчатым фагом M13 и могут быть упакованы в капсид бактериофага. Фагмиды используют в различных методах биотехнологии; например, их можно использовать в фаговом дисплее (подробности которого приведены в настоящем изобретении в другом месте). Несколько различных фагмид являются коммерчески доступными и могут использоваться в контексте настоящего изобретения.

Настоящее изобретение также предусматривает фаг, содержащий нуклеиновую кислоту, или вектор по настоящему изобретению. Такой фаг может быть частью библиотеки, например, библиотека фагового дисплея.

Плазмиды также хорошо известны. Плазмиды могут быть сконструированы для экспрессии генов тяжелой и легкой цепей иммуноглобулина в бактериях или млекопитающих, переменные области которых вырабатываются (и модифицируются) с помощью полимеразной цепной реакции (ПЦР), как описано в настоящем изобретении. Несколько различных плазмид являются коммерчески доступными и могут быть использованы в контексте настоящего изобретения.

Скрининг, клетки-хозяева и способы получения полипептидов по настоящему изобретению

Настоящее изобретение также относится к способам получения полипептида, содержащего модифицированный домен VH, описанный здесь в настоящем изобретении. Нуклеиновые кислоты VH, кодирующие домен VH, могут быть получены путем иммунизации антигеном животного, но не человека, предпочтительно трансгенного животного, но не человека, тем самым вырабатывая домены VH, специфичные для этого антигена, и вызывая клональную экспансию В-клеток, вырабатывающих такие домены VH.

Нуклеиновые кислоты, кодирующие домены VH, затем могут быть выделены для синтеза кДНК и использованы в способах, описанных в настоящем изобретении.

В одном способе такую кДНК можно использовать для создания библиотек фагового дисплея для скрининга доменов VH, которые проявляют требуемые

свойства связывания. После отбора желаемые нуклеиновые кислоты VH могут быть трансфицированы в клетку-хозяина для выработки антител.

В другом способе кДНК используют в частотном анализе, при этом кДНК, кодирующие переменные области тяжелой цепи, подвергают высокопроизводительному секвенированию, а кДНК отбирают для трансфекции в клетку-хозяина для выработки антител, в том числе на основе частоты переменной области используемого сегмента гена, полной последовательности переменной области, HCDR3 или других свойств, желательных для специалиста в данной области.

Трансфекция в клетку-хозяин может быть осуществлена известным в данной области образом и как дополнительно описано в настоящем изобретении. Например, для трансфекции можно использовать любой из описанных в настоящем изобретении векторов. Кроме того, предполагают, что клетка-хозяин содержит нуклеиновую кислоту, кодирующую модифицированный домен VH, включающий варианты, кодирующие предпочтительно в первом и втором положениях аминокислоты, интегрированные в ее геном, и которые могут дополнительно содержать дополнительные расположенные ниже по цепи варианты.

Еще один способ включает: получение клетки, содержащей одну или несколько нуклеиновых кислот, кодирующих полипептиды, которые способны собираться в антитело по настоящему изобретению; культивирование указанной клетки в условиях, обеспечивающих экспрессию полипептидов и их сборку в антитело.

Молекулы нуклеиновой кислоты, кодирующие описанный в настоящем изобретении модифицированный домен VH, могут присутствовать в виде внехромосомных копий и/или быть стабильно интегрированными в хромосому клетки-хозяина. Последнее предпочтительнее, и в этом случае локус может быть нацелен на отсутствие сайленсинга генов.

Чтобы вызвать экспрессию нуклеиновых кислот, кодирующих полипептиды, содержащие домен СНЗ, специалистам в данной области хорошо известно, что последовательности, способные управлять такой экспрессией, могут быть функционально связаны с нуклеиновыми кислотами, кодирующими полипептиды, содержащие домен СНЗ. Термин «функционально связанные» означает, что нуклеиновые кислоты, кодирующие полипептиды, содержащие

домен СНЗ, или их предшественников, связаны с последовательностями, способными управлять экспрессией таким образом, что эти последовательности могут управлять экспрессией полипептидов, содержащих домен СНЗ, или их предшественников. Векторы экспрессии, которые могут быть применены и

5 доступные в данной области, например, относятся к серии векторов pcDNA фирмы Invitrogen. Когда последовательность, кодирующая полипептид интереса, надлежащим образом интегрирована относительно последовательности, управляющей транскрипцией и трансляцией кодируемого полипептида, получаемая кассета экспрессии может быть использована для получения

10 полипептида интереса, называемого полипептидом экспрессии. Последовательности, управляющие экспрессией, могут включать промоторы, энхансеры и т.п., а также их комбинации. Они должны функционировать в клетке-хозяине, тем самым стимулируя экспрессию нуклеиновых кислот, которые функционально с ними связаны. Промоторы могут быть

15 конститутивными или индуцируемыми и могут быть получены из различных источников, включая вирусы, прокариотические или эукариотические источники, или могут быть созданы искусственно. Экспрессия представляющих интерес нуклеиновых кислот может происходить с природного промотора, или его производного, или из полностью гетерологичного промотора. Некоторые

20 хорошо известные и часто используемые промоторы для экспрессии в эукариотических клетках включают промоторы, полученные из вирусов, таких как аденовирусы, например промотор E1A, промоторы, полученные из цитомегаловируса (CMV), такие как немедленно-ранний промотор (IE – immediate early) CMV, промоторы, полученные из вируса обезьян 40 (SV40), и

25 т.п. Соответствующие промоторы также могут быть получены из эукариотических клеток, например, промоторы металлотионеина (MT), промотор фактора элонгации 1 α (EF-1 α), промотор актина, промотор иммуноглобулина, промоторы теплового шока и т.п. Любой промотор или энхансер/промотор, способный управлять экспрессией последовательности интереса в клетке-

30 хозяине, подходит для изобретения. В одном из вариантов осуществления настоящего изобретения последовательность, способная управлять экспрессией, содержит область промотора CMV, предпочтительно область, содержащую нуклеотиды от -735 до +95 немедленно-раннего промотор энхансера/промотора гена CMV. Специалисту в данной области должно быть известно, что

последовательности экспрессии, используемые в настоящем изобретении, могут подходящим образом комбинироваться с элементами, которые могут стабилизировать или усиливать экспрессию, такими как инсуляторы, участки прикрепления к матрице, элементами STAR (WO 03/004704) и т.п. Это может

5 повышать стабильность и/или уровни экспрессии.

Выработка белка в рекомбинантных клетках-хозяевах подробно описана, например, в кн.: «Current Protocols in Protein Science», 1995, Coligan JE, Dunn

10 BM, Ploegh HL, Speicher DW, Wingfield PT, ISBN 0-471-11184-8; Bendig, 1988.

Культивирование клеток проводят для того, чтобы произошла активизация метаболизма, и/или рост, и/или деление, и/или выработка представляющих

15 интерес рекомбинантных белков. Это может быть выполнено методами, хорошо известными специалистам в данной области, и включает, но не ограничивается этим, обеспечение клетки питательными веществами. Методы включают рост с прикреплением к поверхностям, рост в суспензии или их комбинации.

Некоторые условия культивирования можно оптимизировать методами, хорошо известными в данной области, для оптимизации выработки белка.

Культивирование можно проводить, например, в чашках, роллерных бутылках или в биореакторах с использованием систем периодического культивирования, систем с подпиткой, непрерывных систем, полых волокон и т.п. Для

20 крупномасштабного (непрерывного) получения рекомбинантных белков с помощью клеточной культуры используют клетки, способные расти в суспензии, которые можно культивировать в отсутствие сыворотки животного или человека или компонентов таких сывороток. Таким образом, очистка упрощается, а безопасность повышается из-за отсутствия дополнительных белков животного

25 или человеческого происхождения, полученных из культуральной среды, в то время как система также очень надежна, поскольку синтетические среды обеспечивают наилучшую воспроизводимость.

Иммуноглобулиноподобные полипептиды экспрессируют в клетках-хозяевах и собирают из клеток или, предпочтительно, из среды для

30 культивирования клеток методами, которые известны специалистам в данной области. После сбора Ig-подобные полипептиды могут быть очищены с использованием методов, известных в данной области. Такие методы могут включать преципитацию, центрифугирование, фильтрацию, эксклюзионную хроматографию, аффинную хроматографию, катионообменную и/или

анионообменную хроматографию, гидрофобную хроматографию и т.п. Для смеси антител, содержащих полипептиды IgG, можно использовать аффинную хроматографию с белком А или белком G (см, например, US 4801687 и US 5151504).

- 5 После захвата с помощью аффинной хроматографии используют стадии ортогональной полировки для удаления любых оставшихся примесей, связанных с процессом и/или продуктом, которые могут включать гомодимеры, варианты заряда, белок клетки-хозяина (HCP – host cell protein) и ДНК клетки-хозяина. Обычно для получения очищенного биспецифического антитела или
- 10 поливалентного мультимера выполняют следующие стадии, включая культивирование клеток-хозяев, осветление с последующим захватом белка, анионообменную хроматографию, в том числе для удаления ДНК клетки-хозяина, затем катионообменную хроматографию (cation exchange chromatography – CIEX) используют для удаления белка клетки-хозяина,
- 15 выщелоченного белка А, потенциальных агрегатов и потенциальных примесей, связанных с продуктом, с последующими дополнительными стадиями, такими как нанофильтрация в качестве заключительной стадии процесса удаления вируса. Специалистам в данной области техники известно, что порядок таких стадий может быть изменен или отдельные стадии могут быть заменены.
- 20 Например, альтернативы второй стадии полировки включают хроматографию гидрофобного взаимодействия и хроматографию смешанного типа.

- Иммуноглобулиноподобные полипептиды и/или их смеси, полученные способами по настоящему изобретению, предпочтительно имеют общую легкую цепь. Таким образом, дополнительно предложен способ согласно настоящему
- 25 изобретению, дополнительно включающий обеспечение указанной клетки-хозяина нуклеиновой кислотой, кодирующей общую легкую цепь. Такая легкая цепь, способная образовывать пары по меньшей мере с двумя разными тяжелыми цепями, образуя тем самым функциональные антигенсвязывающие домены. Функциональный антигенсвязывающий домен способен специфически
- 30 связываться с антигеном. В одном варианте осуществления настоящего изобретения используют общую легкую цепь, которая способна спариваться со всеми тяжелыми цепями, полученными способом согласно настоящему изобретению, тем самым образуя функциональные антигенсвязывающие домены, так что можно избежать неправильного спаривания несовпадающих тяжелых и

легких цепей. Согласно одному объекту настоящего изобретения применяют только общие легкие цепи с одной идентичной аминокислотной последовательностью. В другом варианте специалистам в данной области будет понятно, что термин «общий» также относится к функциональным эквивалентам легкой цепи, аминокислотная последовательность которых не является идентичной. Существует множество вариантов указанной легкой цепи, в которых присутствуют мутации (делеции, замены, инсерции), которые существенно не влияют на образование функциональных участков связывания. Таким образом, такие варианты также способны связывать различные тяжелые цепи и образовывать функциональные антигенсвязывающие домены. Термин «общая легкая цепь», используемый в настоящем изобретении, таким образом, относится к легким цепям, которые могут быть идентичными или иметь некоторые отличия в аминокислотной последовательности, сохраняя при этом специфичность связывания полученного антитела после спаривания с тяжелой цепью. Например, можно получить или найти легкие цепи, которые не являются идентичными, но все же функционально эквивалентны, например, введением и тестированием консервативных аминокислотных замен и/или замен аминокислот в участках, которые не вносят вклад или лишь частично вносят вклад в специфичность связывания в сочетании с тяжелой цепью, и т.п. Комбинация определенной общей легкой цепи и таких функционально эквивалентных вариантов охватывается термином «общая легкая цепь». Ссылка дана на WO 2004/009618 для подробного описания использования обычных легких цепей. Предпочтительно, в настоящем изобретении используют обычную легкую цепь, которая представляет собой легкую цепь, подобную легкой цепи зародышевой линии, более предпочтительно легкую цепь зародышевой линии, предпочтительно перегруппированную легкую цепь каппа зародышевой линии человека, наиболее предпочтительно перегруппированную легкую цепь каппа зародышевой линии человека IgV κ 1-39/ J κ , IGV κ 3-15/J κ или IGV κ 3-20/J κ . Также можно использовать перегруппированную легкую цепь лямбда зародышевой линии человека. Предпочтительная перегруппированная легкая цепь лямбда человека зародышевой линии содержит IGVL3-21/JL.

В качестве альтернативы использованию общей цепи и во избежание ошибочного спаривания несовпадающих тяжелых и легких цепей специалист

может выбрать средства форсированного спаривания тяжелой и легкой цепей с помощью средств, известных специалистам в данной области.

В настоящем описании также представлены клетки-хозяева, которые экспрессируют полипептиды по настоящему изобретению. «Клеткой-хозяином» согласно настоящему изобретению может быть любая клетка-хозяин, способная экспрессировать молекулы рекомбинантной ДНК, включая бактерии, такие как, например, *Escherichia* (например, *E. coli*), *Enterobacter*, *Salmonella*, *Bacillus*, *Pseudomonas*, *Streptomyces*, дрожжи, такие как *S. cerevisiae*, *K. lactis*, *P. pastoris*, *Candida* или *Yarrowia*, мицелиальные грибы, такие как *Neurospora*, *Aspergillus oryzae*, *Aspergillus nidulans* и *Aspergillus niger*, клетки насекомых, такие как клетки *Spodoptera frugiperda* SF-9 или SF-21, и предпочтительно клетки млекопитающих, такие как клетки яичника китайского хомячка (CHO), клетки ВНК, клетки мыши, включая клетки SP2/0 и клетки миеломы NS-0, клетки приматов, такие как клетки COS и Vero, клетки MDCK, клетки BRL 3A, гибридомы, опухолевые клетки, иммортализованные первичные клетки, клетки человека, такие как W138, HepG2, HeLa, HEK293, HT1080, или эмбриональные клетки сетчатки, такие как PER.C6, и т.п. Часто выбранная система экспрессии включает вектор экспрессии клеток млекопитающих и хозяина таким образом, что антитела могут быть соответствующим образом гликозилированы. Линия клеток человека может быть использована для получения антител с полностью человеческим типом гликозилирования. Условия выращивания или размножения клеток (см., например, кн.: «Tissue Culture», под ред. Kruse и Paterson, изд. Academic Press, 1973) и условия экспрессии рекомбинантного продукта могут несколько отличаться, и оптимизацию процесса обычно проводят для увеличения пропорций продукта и/или роста клеток по отношению друг к другу в соответствии со способами, общеизвестными специалисту в данной области. В целом с принципами, протоколами и практическими методами максимизации продуктивности культур клеток млекопитающих можно ознакомиться в кн.: «Mammalian Cell Biotechnology: a Practical Approach» под ред. M. Butler, IRL Press, 1991. Экспрессия антител в рекомбинантных клетках-хозяевах подробно описана в данной области. Нуклеиновые кислоты, кодирующие легкую и тяжелую цепи, могут присутствовать в виде внехромосомных копий и/или стабильно интегрированы в хромосому клетки-хозяина.

Библиотеки

Настоящее изобретение предусматривает библиотеки (т.е. коллекции) различных нуклеиновых кислот, векторов или фагов по настоящему изобретению. Библиотеки могут содержать по меньшей мере около 10^6 различных нуклеиновых кислот, векторов или фагов по настоящему изобретению.

Примером библиотеки согласно настоящему изобретению является библиотека дисплея. Способы получения библиотеки дисплея хорошо известны в данной области техники. Например, способ получения библиотеки дисплея, отображающей различные модифицированные домены VN по настоящему изобретению, может включать интеграцию нуклеиновой кислоты по настоящему изобретению (например, в форме вектора, описанного в другом месте в настоящем описании) в организм, такой как фаг или дрожжи, или другой объект для пептидного дисплея, где указанный организм экспрессирует и демонстрирует указанный модифицированный домен VN на поверхности указанного организма или объекта. Множественные модифицированные домены VN, как правило, многочисленные различные модифицированные домены VN могут быть отображены на поверхности множества организмов, таких как фаги (каждый фаг отображает один модифицированный домен VN), с использованием библиотеки фагового дисплея. Таким образом, в библиотеке фагового дисплея множество модифицированных доменов VN, кодируемых нуклеиновыми кислотами по настоящему изобретению, могут быть спарены с переменной областью общей цепи человека. Библиотека дисплея может представлять собой, например, библиотеку дисплея фага Fab.

Методы библиотеки фагового дисплея

Различные формы методов дисплея, включая фаговый дисплей, дрожжевой дисплей, рибосомальный дисплей, дисплей мРНК, наряду с другими, известны в данной области техники и входят в рамки охвата настоящего изобретения, для использования с модифицированными доменами VN, описанными в настоящем изобретении. Приводимое ниже обсуждение сосредоточено на фаговом дисплее, но такое описание не является ограничивающим и не основывается на описании, предусмотренном в настоящем изобретении, и оно может быть легко применено к другим формам технологии дисплея. Фаговый дисплей – известный метод, используемый, в том числе, для изучения взаимодействий белок-белок, белок-пептид и белок-ДНК, с использованием бактериофагов – вирусов бактерий.

Многие из описанных в настоящем изобретении протоколов являются стандартными протоколами для создания библиотек фагового дисплея и пэннинга фагов для связывания с представляющим интерес антигеном и описаны в кн.: «Antibody Phage Display: Methods and Protocols», под ред. Philippa M.

5 О'Brien, Robert Aitken). Библиотеки можно выращивать и собирать в соответствии с процедурами, известными в данной области, например, описанными Kramer с соавт. (Kramer с соавт., *Nucleic Acids Res*, 2003, 31(11): e59), используя VCSM13 (фирма Stratagene) в качестве штамма хелперного штамма. Фаги могут быть наработаны и обработаны с помощью процедур, 10 известных в данной области, например, описанных Kramer с соавт. (Kramer с соавт., *Nucleic Acids Res*, 2003, 31(11): e59), используя VCSM13 (фирма Stratagene) в качестве штамма хелперного штамма.

В типичном методе нуклеиновая кислота, кодирующая белок интереса, например, нуклеиновая кислота, кодирующая модифицированный домен VH, 15 интегрирована в ген белка оболочки фага, заставляя фаг «отображать» белок снаружи частицы, в то время внутри фаг содержит нуклеиновую кислоту, кодирующую этот белок. Таким образом устанавливают связь между генотипом и фенотипом. Что касается обнаружения антител, при фаговом дисплее большие коллекции (библиотеки) доменов VH и/или VL могут быть экспрессированы на 20 поверхности нитевидных частиц бактериофага таким образом, что они образуют пары для образования связывающих доменов. Из таких библиотек фаги могут быть выбраны посредством связывающего взаимодействия с антигеном и отображаемым связывающим доменом. Таким образом, фаги в методе фагового дисплея можно подвергать скринингу против других белков, пептидов или ДНК- 25 последовательностей или других форм фрагментов-мишеней для обнаружения взаимодействия между отображаемыми VH, VL или связывающим доменом и этими другими фрагментами. Таким образом, большие библиотеки VH, VL или связывающих доменов могут быть подвергнуты скринингу и амплифицированы в процессе, называемом отбором *in vitro*, который аналогичен естественному 30 отбору. Соответственно, модифицированный домен VH по настоящему изобретению может быть отображен на фаге.

Настоящее изобретение обеспечивает эффективный процесс последовательной сборки для получения по существу всех нуклеиновых кислот, кодирующих переменные области тяжелой цепи, от иммунизированного

животного, включая трансгенное животное, и интегрируя нуклеиновые кислоты, кодирующие модифицированные переменные области тяжелой цепи, в технологию дисплея (например, фагового, дрожжевого, рибосомального и т.д.), где каждая из указанных нуклеиновых кислот кодирует неглутаматный и неглутаминовый аминокислотный остаток на N-конце модифицированной переменной области тяжелой цепи, который не является глутаматом или глутамином, тем самым позволяя тестировать практически все переменные области тяжелой цепи от иммунизированного животного, не имеющие такого остатка, независимо от генного сегмента переменной области человека или от которого получена переменная область у указанного животного.

В другом варианте настоящее изобретение предусматривает способ получения определенной популяции связывающих молекул, включающих N-концевые модифицированные переменные области тяжелой цепи, при этом популяция В-клеток, экспрессирующих ограниченный репертуар VL, предпочтительно одну или общую легкую цепь, и экспрессирующих различные переменные областей тяжелой цепи, специфичные в отношении антигена интереса. Указанные В-клетки могут быть получены путем иммунизации трансгенного животного, несущего locus или loci иммуноглобулина человека, антигеном интереса. Нуклеиновые кислоты (РНК или ДНК) из указанных В-клеток секвенируют, кодируя часть и предпочтительно по существу все указанные переменные области тяжелой цепи. Указанные нуклеиновые кислоты, кодирующие переменные области тяжелой цепи иммуноглобулина в указанном образце, предпочтительно амплифицируют и подвергают частотному анализу, причем анализируют использование генного сегмента гена V в указанной популяции, анализируют последовательность VH, HCDR3 и дополнительные свойства репертуара, представляющие интерес для квалифицированного специалиста.

Указанную переменную область (области) тяжелой цепи из такого частотного анализа затем выбирают и вводят в клетку-хозяина способом, описанным в настоящем изобретении, для получения вариаций, включающих первую или первую и вторую кодируемые аминокислоты переменной области по меньшей мере одной последовательности VH, предпочтительно два или более с по меньшей мере одной последовательностью VL указанного ограниченного репертуара VL или общей легкой цепи, дополнительно введенной в указанную

клетку-хозяина. После этого указанную клетку-хозяин культивируют, чтобы обеспечить экспрессию модифицированных полипептидов VH и VL, где одна модифицированная VH снабжена одной VL в указанной клетке-хозяине для

5 модифицированных последовательностей VH снабжены одной VL в указанную клетку-хозяин для получения мультиспецифического антитела.

Способы одновременной амплификации и модификации нуклеиновой кислоты

В настоящем документе также предложен способ одновременной
10 амплификации и модификации нуклеиновой кислоты, кодирующей
вариабельный домен тяжелой цепи иммуноглобулина человека. Этот способ
особенно полезен, поскольку он сочетает амплификацию целевой матрицы с
модификацией последовательности матрицы в одну стадию, так что
амплифицированная нуклеиновая кислота кодирует новый полипептид по
15 настоящему изобретению (т.е. полипептид, содержащий модифицированный
домен VH человека, согласно описанию настоящего изобретения).

Описанный в настоящем изобретении способ включает стадию получения
нуклеиновой кислоты, которая кодирует вариабельный домен тяжелой цепи
иммуноглобулина человека (также называемый в настоящем изобретении
20 доменом VH человека). Нуклеиновая кислота, предназначенная для
использования в способе, также может называться в настоящем изобретении
матричной последовательностью, которая представляет собой
последовательность, подлежащую амплификации и модификации.

Нуклеиновая кислота-матрица может быть получена из любого
25 подходящего источника. Обычно матричная нуклеиновая кислота может быть
кДНК, например кДНК, полученной путем обратной транскрипции образца РНК.
Образец РНК может представлять собой суммарную РНК или мРНК,
полученную из клетки, которая экспрессирует полипептид, содержащий
вариабельный домен тяжелой цепи иммуноглобулина человека.

30 Любая клетка-хозяин, описанная в настоящем изобретении, может быть
использована для получения матричной нуклеиновой кислоты (например,
последовательности кДНК, которая соответствует последовательности РНК,
продуцируемой клеткой, кодирующей вариабельный домен тяжелой цепи
иммуноглобулина человека). В особенно предпочтительном примере матричную

нуклеиновую кислоту получают из клеток человека или клеток трансгенных животных, содержащих генные сегменты вариабельной области иммуноглобулина человека. В другом примере трансгенное животное содержит локус тяжелой цепи иммуноглобулина человека или его часть (например, мини-
 5 локус тяжелой цепи иммуноглобулина человека).

Можно использовать любое подходящее трансгенное животное, например, трансгенную овцу, кролика, крысу, мышь, птицу, включая курицу, и т.д., содержащее генные сегменты вариабельной области человека, которые образуют
 10 человеческие, гуманизированные или химерные антитела или тяжелые цепи, содержащие на N-конце вариабельного домена тяжелой цепи глутамат или глутамин.

Ранее были описаны трансгенные животные, несущие генные сегменты вариабельной области человека. Такие трансгенные животные могут быть использованы в способах, описанных в настоящем изобретении. Трансгенное
 15 животное, применимое для использования в настоящем изобретении и описанное в нем, содержит нуклеиновые кислоты, кодирующие общую цепь иммуноглобулина человека, содержащую реаранжированную легкую или тяжелую вариабельную цепь и кодирующую вариабельную область без перегруппировки родственной цепи (цепей) в зародышевой линии таких
 20 животных согласно описанию в WO 2009/157771. Такие трансгенные животные способны вырабатывать антитела, обладающие разнообразием, происходящим от одной из двух родственных цепей иммуноглобулина, например, не прошедшей реаранжировку тяжелой или легкой цепи, которая подвергается соматической рекомбинации во время развития В-клеток и созревания аффинности после
 25 воздействия антигена. Трансгенные животные, такие как MeMo®, способны вырабатывать разнообразные репертуары антител, направленных против множества антигенов.

Трансгенное животное с ДНК человека может быть иммунизировано представляющим интерес антигеном или эпитопом. Подходящим протоколом
 30 иммунизации обычно является такой, который вызывает селективную экспансию В-клеток, а это означает, что первичная и бустерная иммунизация предназначены для селективной экспансии В-клеток, которые продуцируют антитела, связывающиеся с представляющим интерес антигеном или эпитопом. Протокол иммунизации может, например, использовать различные формы или

фрагменты антигена во время первичной иммунизации и каждой последующей бустерной иммунизации. Например, антиген может быть экспрессирован на мембране клетки, липоидной частице, мицелле, рекомбинантном белке, рекомбинантном белке, слитом с другим белком, домене белка или пептиде как части белка. Протокол иммунизации может включать использование адъюванта во время первичной и/или бустерной иммунизации. Адъювант можно использовать во время первичной иммунизации только для ограничения степени неспецифического размножения «фоновых» В-клеток. «Фоновые» В-клетки отличаются тем, что активируются без стадии связывания антигена с рецептором антитела, экспрессируемым на поверхности В-клетки. В данной области техники известно, что иммунизация, например, Fc-слитыми белками, часто приводит к сильному анти-Fc-ответу, когда примерно до 70% всех В-клеток реагируют на часть Fc слитого белка, а не на антиген интереса. Протокол иммунизации можно использовать без адъюванта для предпочтительного размножения В-клеток, которые были активированы антигеном, используемым для иммунизации.

При получении нуклеиновой кислоты, кодирующей переменный домен тяжелой цепи иммуноглобулина человека от трансгенного животного, обычно выделяют В-клетки. В-клетки могут быть выделены из любого подходящего источника, такого как ткань (например, из лимфатической ткани или из костного мозга (т.е. из ткани, продуцирующей В-клетки)), или из периферической крови (например, от более крупных трансгенных животных, таких как овцы). Например, магнитные микрогранулы, покрытые маркером рап В-клеток человека, CD19, можно использовать для выделения В-клеток из периферической крови (см., например, Bertrand FE, III, с соавт., *Blood*, 1997, 90:736-744).

Нуклеиновые кислоты, кодирующие переменный домен тяжелой цепи иммуноглобулина человека (например, последовательности РНК), обычно выделяют из клеток с использованием стандартных методов. Для получения соответствующей кДНК можно использовать обратную транскрипцию с использованием специфических для гена праймеров или общих праймеров для РНК (например, полиА-праймеров). Как правило, такую кДНК, которая также кодирует переменный домен тяжелой цепи иммуноглобулина человека, используют в качестве матричной нуклеиновой кислоты в описанных в настоящем изобретении способах.

Нуклеиновые кислоты, кодирующие переменный домен тяжелой цепи иммуноглобулина, также могут быть получены путем иммунизации не трансгенного животного. Такие переменные домены тяжелой цепи иммуноглобулина затем соответствующим образом подвергают гуманизации или химеризации методами, известными в данной области.

Способ также включает стадию проведения полимеразной цепной реакции (ПЦР) по меньшей мере с одним 5'-праймером, по меньшей мере с одним 3'-праймером и нуклеиновой кислотой для получения амплифицированной нуклеиновой кислоты.

Обозначения «полимеразная цепная реакция» и «ПЦР» применяют в настоящем описании взаимозаменяемо. Они относятся к методу специфической амплификации области нуклеиновых кислот, например ДНК или РНК. Область может представлять собой отдельный ген, часть гена, кодирующую или некодирующую последовательность или содержать их комбинацию. В большинстве методов ПЦР обычно амплифицируют фрагменты ДНК из сотен пар оснований (п.о.), хотя некоторые методы позволяют амплифицировать фрагменты размером до 40 тысяч пар оснований (т.п.о.). Для базовой установки ПЦР требуется несколько компонентов и реагентов. Эти компоненты включают матрицу нуклеиновой кислоты, содержащую амплифицируемый участок, два праймера, комплементарных 5'- и 3'-концам амплифицируемого участка («5'-праймер» (или прямой праймер) и «3'-праймер (или обратный праймер) соответственно), полимеразу, такую как Taq-полимераза или другая термостабильная полимеразы, дезоксинуклеотидтрифосфаты (dNTP), из которых полимеразы синтезируют новую цепь, буферный раствор, обеспечивающий подходящую химическую среду для оптимальной активности и стабильности полимеразы, двухвалентные катионы, обычно Mg^{2+} , и, наконец, одновалентные катионы, такие как ионы калия.

Точные условия ПЦР, необходимые для амплификации матричной нуклеиновой кислоты, могут быть определены специалистом в данной области с помощью общеизвестных подходов. Примеры условий ПЦР, не ограничивающие рамок охвата настоящего изобретения, которые можно использовать в контексте настоящего изобретения, включают также условия, описанные в настоящем изобретении.

В способе используют по меньшей мере один 5'-праймер, содержащий нуклеиновую кислоту с сайтом модификации, который вносит модификацию в амплифицированную нуклеиновую кислоту таким образом, что амплифицированная нуклеиновая кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевую аминокислоту, которая не является глутамином или глутаматом. Остаток может быть выбран из группы, состоящей из: аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и валина. Это позволяет способу амплифицировать матричную нуклеиновую кислоту, одновременно модифицируя ее таким образом, чтобы она кодировала полипептид по настоящему изобретению (т.е. полипептид, содержащий модифицированный человеческий, гуманизированный или химерный домен VH согласно описанию настоящего изобретения).

По меньшей мере один 5'-праймер может иметь любую последовательность, которая вносит требуемую модификацию в амплифицированную нуклеиновую кислоту. В настоящем описании представлено несколько примеров подходящих 5'-праймеров. Также можно использовать другие последовательности праймеров, которые также способны вносить требуемую модификацию в требуемое положение (положения) в кодируемом домене VH. Как только специалист в данной области техники узнает, что, неожиданно, модификации на N-конце человеческого, гуманизированного или химерного домена VH могут быть применены, становится возможным конструирование соответствующих праймеров.

По меньшей мере один 5'-праймер, используемый в способах по настоящему изобретению, может включать последовательность, которая вносит модификацию в кодируемый домен VH матричной нуклеиновой кислоты таким образом, что в амплифицированной нуклеиновой кислоте N-концевой остаток глутамин или глутамата кодируемого человеческого, гуманизированного или химерного домена VH заменен кислотным или полярным остатком, таким как аспартат или аспарагин. Эти аминокислоты имеют сходные биохимические свойства с глутамином или глутаматом и, таким образом, могут быть успешным выбором.

В конкретном примере по меньшей мере один 5'-праймер, используемый в способах по настоящему изобретению, может включать последовательность, которая вносит модификацию в кодируемый домен VH матричной нуклеиновой кислоты таким образом, что в амплифицированной нуклеиновой кислоте N-концевой остаток глутамина или глутамата кодируемого человеческого, гуманизированного или химерного домена VH заменен аланином. Эта модификация особенно полезна, поскольку устраняет образование пироглутамата, а также поддерживает (например, улучшает) эффективность отщепления сигнального пептида.

Аланин является алифатическим остатком. Соответственно, в другом примере по меньшей мере один 5'-праймер, используемый в способах по настоящему изобретению, может включать последовательность, которая вносит модификацию в кодируемый домен VH матричной нуклеиновой кислоты таким образом, что в амплифицированной нуклеиновой кислоте N-концевой остаток глутамина или глутамата кодируемого человеческого, гуманизированного или химерного домена VH заменен алифатическим остатком, таким как аланин, глицин, аланин, валин, лейцин или изолейцин.

По меньшей мере один 5'-праймер, используемый в способах по настоящему изобретению, может включать последовательность, которая вводит две модификации в кодируемый домен VH матричной нуклеиновой кислоты таким образом, что в амплифицированной нуклеиновой кислоте N-концевая последовательность кодируемого человеческого, гуманизированного или химерного домена VH заменен (первой) N-концевой аминокислотой, выбранной из группы, состоящей из: аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и валина; и (второй) аминокислотой (в расчете на N-конец домена VH), выбранной из группы, состоящей из пролина, валина, аспартата, глутамата, серина, лейцина или треонина. В одном конкретном примере второй аминокислотой на N-конце домена VH выбирают пролин.

В одном конкретном примере по меньшей мере один 5'-праймер, используемый в способах по настоящему изобретению, может включать последовательность, которая вводит две модификации в кодируемый домен VH матричной нуклеиновой кислоты таким образом, что в амплифицированной

нуклеиновой кислоте N-концевая последовательность кодируемого человеческого, гуманизированного или химерного домена VH заменена на аланин-пролин, аланин-аспартат, аланин-глутамат, аланин-треонин, аланин-валин, аланин-серин и аланин-лейцин. Форматирование, применяемое в настоящем изобретении для «аланин-пролин» и других последовательностей, относится к двум соседним аминокислотам на N-конце модифицированного домена VH (т.е. «первая-вторая» аминокислоты на N-конце домена VH) (в направлении от N-конца к C-концу)).

В предпочтительном примере по меньшей мере один 5'-праймер, используемый в способах по настоящему изобретению, включает последовательность, которая вводит две модификации в кодируемый человеческий, гуманизированный или химерный домен VH матричной нуклеиновой кислоты таким образом, что в амплифицированной нуклеиновой кислоте первые две аминокислоты на N-конце кодируемого человеческого, гуманизированного или химерного домена VH представляют собой аланин-пролин. Обнаружено, что эта комбинация особенно выгодна для повышения эффективности отщепления сигнального пептида.

По меньшей мере один 5'-праймер, используемый в способах по настоящему изобретению, может включать последовательность, которая кодирует сигнальный пептид или часть сигнального пептида таким образом, что в амплифицированной (и модифицированной) нуклеиновой кислоте сигнальный пептид кодируется выше по цепи от кодируемого модифицированного человеческого, гуманизированного или химерного домена VH.

Пространственная взаимосвязь между сигнальным пептидом и N-концом модифицированного домена VH в кодируемом полипептиде подробно описана в настоящем изобретении и в равной степени применима в нем.

Таким образом, по меньшей мере один 5'-праймер, используемый в способах по настоящему изобретению, может включать последовательность, кодирующую сигнальный пептид или часть сигнального пептида, содержащую последовательностьAQPAMA (SEQ ID NO: 5), расположенную выше по цепи от сайта модификации (то есть так, что в кодируемом полипептиде последовательность ... AQPAMA (SEQ ID NO: 5) сигнального пептида находится непосредственно рядом с модифицированным N-концом модифицированного домена VH (и выше него по цепи). Например, по меньшей мере один 5'-праймер

может включать последовательность, которая вводит аланин (или комбинацию аланин-пролин) на N-конец модифицированного домена VH, кодируемого амплифицированной нуклеиновой кислотой, и вводит сигнальный пептид выше по цепи кодируемого модифицированного домена VH таким образом, что

5 остатки, фланкирующие сайт расщепления сигнального пептида в модифицированном домене VH, содержат **AQPAMAA** (SEQ ID NO: 6) или **AQPAMAAP** (SEQ ID NO: 7) (с последовательностью на C-конце сигнального пептида, выделенной жирным шрифтом, а N-концевая аминокислота домена VH подчеркнута).

10 В другом примере по меньшей мере один 5'-праймер, используемый в способах по настоящему изобретению, может включать последовательность, кодирующую сигнальный пептид, или часть сигнального пептида, содержащую последовательностьAQPAMA (SEQ ID NO: 5) выше по цепи от сайта модификации (т.е. такой, что в кодируемом полипептиде последовательность ...

15 AQPAMA (SEQ ID NO: 5) сигнального пептида находится непосредственно рядом с (и выше по цепи) модифицированного N-конца модифицированного домена VH). Например, по меньшей мере один 5'-праймер может включать последовательность, которая вводит аланин (или комбинацию аланин-пролин) на N-конце модифицированного домена VH, кодируемого амплифицированной

20 нуклеиновой кислотой, и вводит сигнальный пептид выше по цепи кодируемого модифицированного домена VH таким образом, что остатки, фланкирующие сайт расщепления сигнального пептида в модифицированном человеческом, гуманизированном или химерном домене VH, содержат

25 **MKYLLPTAAAGLLLLAAQPAMAA** (SEQ ID NO: 8) или **MKYLLPTAAAGLLLLAAQPAMAAP** (SEQ ID NO: 9) (с последовательностью на C-конце сигнального пептида, выделенной жирным шрифтом, а N-концевая аминокислота домена VH подчеркнута).

30 В другом примере по меньшей мере один 5'-праймер, используемый в способах по настоящему изобретению, может включать последовательность, кодирующую сигнальный пептид, или часть сигнального пептида, содержащую последовательностьAQPAMA (SEQ ID NO: 5) выше по цепи от сайта модификации (т.е. такой, что в кодируемом полипептиде последовательность ...

AQPAMA (SEQ ID NO: 5) сигнального пептида находится непосредственно рядом с (и выше по цепи) модифицированного N-конца модифицированного

домена VH). Например, по меньшей мере один 5'-праймер может включать последовательность, которая вводит аланин (или комбинацию аланин-пролин) на N-конце модифицированного домена VH, кодируемого амплифицированной нуклеиновой кислотой, и вводит сигнальный пептид выше по цепи кодируемого модифицированного домена VH таким образом, что остатки, фланкирующие сайт расщепления сигнального пептида в модифицированном человеческом, гуманизированном или химерном домене VH, содержат

MGWSCILFLVLLLAQ**PAMA**A (SEQ ID NO: 10) или

MGWSCILFLVLLLAQ**PAMA**AP (SEQ ID NO: 11) (с последовательностью на C-конце сигнального пептида, выделенной жирным шрифтом, а N-концевая аминокислота домена VH подчеркнута).

Примеры подходящих последовательностей 5'-праймеров, которые модифицируют первые два кодона N-конца человеческого, гуманизированного или химерного домена VH (для кодирования аланин-пролин) и вводят сигнальный пептид, непосредственно примыкающий и расположенный выше по цепи от кодируемого модифицированного домена VH, представлены в табл. 1 ниже. Из таблицы следует, что приведенные в качестве примера 5'-праймеры предназначены для амплификации каждого функционального генного сегмента IGHV человека из каждого семейства генов. Это основано на последовательностях после сайта модификации в праймере, которые комплементарны (по крайней мере, частично) немодифицированной нуклеиновой кислоте, обнаруженной в этих положениях в перечисленных сегментах гена VH человека.

Таблица 1. Перечень универсальных 5'-праймеров и их целевых IGHV семейства человека.

Семейство IGHV человека:	5'-универсальные праймеры, описанные в настоящем изобретении, которые применимы для амплификации и модификации AP представителей семейства IGHV:
IGHV1	1308AP, 1308AP2, 2018AP2, 2018AP или 2020AP2
IGHV2	1310AP2, 1310AP3, 1310AP4 или предпочтительно 1310AP5
IGHV3	0508AP, 0508AP2, 2018AP, 2018AP2, 2021AP, 2021AP2, 2021AP3, 2021AP4 или 2021AP5
IGHV4	1312AP2/2019AP2
IGHV5	1313AP или 1313AP2
IGHV6	1310AP2, 1310AP3 и 1310AP4, 1310AP5, 1312AP2/2019AP2,
IGHV7	1314AP или 1314AP2

Таким образом, описанный в настоящем изобретении способ может включать:

(а) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV1, с использованием 5'-праймера, выбранного из 1308AP, 1308AP2, 2018AP, 2018AP2 или 2020AP2; и/или

(б) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV2, с использованием 5'-праймера, выбранного из 1310AP2, 1310AP3, 1310AP4 или 1310AP5; и/или

(в) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV3, с использованием 5'-праймера, выбранного из 0508AP, 0508AP2, 2018AP, 2018AP2, 2021AP, 2021AP2, 2021AP3, 2021AP4 или 2021AP5; и/или

(г) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV4, с использованием 5'-праймера, который представляет собой 1312AP; и/или

(д) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV5, с использованием 5'-праймера, выбранного из 1313AP или 1313AP2;

(е) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV6, с использованием 5'-праймера, выбранного из 1310AP2, 1310AP3, 1310AP4, 1310AP5 или 1312AP2; и/или

(ж) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV7, с использованием 5'-праймера, выбранного из 1314AP или 1314AP2.

В настоящем изобретении отмечают, что праймеры 1312AP2 и 2019AP2 имеют идентичную последовательность, поэтому эти термины могут использоваться взаимозаменяемо. На это также указывает применение термина «1312AP2/2019AP2».

Как показано ниже в примерах, некоторые 5'-праймеры, описанные в настоящем изобретении, могут использоваться предпочтительно.

Например, при амплификации и модификации нуклеиновой кислоты, кодируемой сегментом гена из семейства генов IGHV1, можно использовать 5'-праймер, выбранный из 1308AP, 1308AP2, 2018AP, 2018AP2 или 2020AP2. В качестве альтернативы можно использовать праймер, который аналогичным

образом модифицирует первые две N-концевые аминокислоты вариабельной области человека для кодирования AP. Один конкретный пример включает 5'-праймер, выбранный из 1308AP2, 2018AP2 или 2020AP2.

Кроме того, при амплификации и модификации нуклеиновой кислоты, кодируемой сегментом гена из семейства генов IGHV2, можно использовать 5'-праймер, выбранный из 1310AP2, 1310AP3, 1310AP4 или 1310AP5, или в другом варианте может применяться праймер, который аналогичным образом модифицирует первые две N-концевые аминокислоты вариабельной области человека для кодирования AP. В одном конкретном примере применяют 5'-праймер 1310AP5.

Кроме того, при амплификации и модификации нуклеиновой кислоты, кодируемой сегментом гена из семейства генов IGHV3, можно использовать 5'-праймер, выбранный из 0508AP, 0508AP2, 2018AP, 2018AP2, 2021AP, 2021AP2, 2021AP3, 2021AP4 или 2021AP5, или в другом варианте может применяться праймер, который аналогичным образом модифицирует первые две N-концевые аминокислоты вариабельной области человека для кодирования AP. В одном конкретном примере применяют 5'-праймер, выбранный из 0508AP, 2021AP2 или 2018AP2.

Кроме того, при амплификации и модификации нуклеиновой кислоты, кодируемой сегментом гена из семейства генов IGHV4, можно использовать 5'-праймер, а именно может применяться 1312AP2, или в другом варианте может применяться праймер, который сходным образом модифицирует первые две N-концевые аминокислоты вариабельной области человека для кодирования AP.

Кроме того, при амплификации и модификации нуклеиновой кислоты, кодируемой сегментом гена из семейства генов IGHV5, можно использовать 5'-праймер, выбранный из 1313AP или 1313AP2, или в другом варианте может применяться праймер, который аналогичным образом модифицирует первые две N-концевые аминокислоты вариабельной области человека для кодирования AP. В одном конкретном примере применяют 5'-праймер, а именно 1313AP2.

Кроме того, при амплификации и модификации нуклеиновой кислоты, кодируемой сегментом гена из семейства генов IGHV6, можно использовать 5'-праймер, выбранный из 1310AP2, 1310AP3, 1310AP4, 1310AP5 или 1312AP2, или в другом варианте может применяться праймер, который аналогичным образом модифицирует первые две N-концевые аминокислоты вариабельной области

человека для кодирования AP. В одном конкретном примере применяют 5'-праймер, выбранный из 1312AP2 и 1310AP5.

При амплификации и модификации нуклеиновой кислоты, кодируемой сегментом гена из семейства генов IGHV7, можно использовать 5'-праймер, выбранный из 1314AP или 1314AP2, или в другом варианте может применяться праймер, который аналогичным образом модифицирует первые две N-концевые аминокислоты вариабельной области человека для кодирования AP. В одном конкретном примере применяют 5'-праймер, а именно 1314AP2.

Специалистам в данной области известно, что могут быть обстоятельства, при которых было бы полезно одновременно амплифицировать и модифицировать множество различных матричных нуклеиновых кислот, которые кодируют разные домены VH человека. Например, может быть полезным использование способов, описанные в настоящем изобретении, для амплификации (и модификации) частичного или полного репертуара различных доменов VH человека, закодированных в образце клеток человека и/или трансгенных животных.

Например, способы, описанные в настоящем изобретении, можно использовать для одновременной модификации и амплификации репертуара нуклеиновых кислот, кодирующих домены VH человека у трансгенного животного (например, трансгенного мышинного или птичьего организма, имеющего локус тяжелой цепи иммуноглобулина человека или его часть; например мыши MeMo®). Предложенные в настоящем изобретении способы позволяют одновременно амплифицировать и модифицировать различные генные сегменты IGHV, например, с использованием мультиплексной реакции ПЦР, в которой используют несколько различных 5'-праймеров.

Термины «мультиплексная полимеразная цепная реакция» или «мультиплексная ПЦР» используются в настоящем изобретении взаимозаменяемо для обозначения полимеразной цепной реакции с использованием нескольких уникальных праймеров в одной реакции/смеси ПЦР для получения амплифицированных нуклеиновых кислот с различными последовательностями. Воздействуя одновременно на несколько генов, можно получить дополнительную информацию из одного тестирования, которое в противном случае потребовало бы в несколько раз больше реагентов и больше

времени для выполнения. Температуры отжига для каждого набора праймеров должны быть оптимизированы для правильной работы в рамках одной реакции.

Таким образом, способ может включать стадию получения множества различных нуклеиновых кислот, кодируемых по меньшей мере одним генным сегментом человека, выбранным из каждого из следующих семейств генов человека: IGHV1, IGHV2, IGHV3, IGHV4, IGHV5, IGHV6 и IGHV7.

Преимущественно множество различных нуклеиновых кислот (в качестве матричной нуклеиновой кислоты) могут быть амплифицированы и модифицированы одновременно в одной реакции ПЦР.

Например, по меньшей мере один праймер из каждой строки в табл. 1 может быть выбран и использован в способах по настоящему изобретению для амплификации и модификации матричной нуклеиновой кислоты, которая кодирует функциональные генные сегменты в следующих семействах генов человека IGHV1, IGHV2, IGHV3, IGHV4, IGHV5, IGHV6 и IGHV7 одновременно.

В примерах ниже показано, что некоторые 5'-праймеры, описанные в настоящем документе, могут использоваться предпочтительно.

Например, для использования в способах настоящего изобретения может быть выбрана комбинация по меньшей мере одного праймера из каждой из следующих категорий (в результате в одной реакции используют смесь по меньшей мере шести различных 5'-праймеров, в зависимости от количество семейств генов V, дающих начало репертуару VH):

а) 5' праймер, выбранный из 1308AP, 1308AP2, 2020AP2, 2018AP или 2018AP2;

б) 5' праймер, выбранный из 1310AP2, 1310AP3, 1310AP4 или 1310AP5;

в) 5' праймер, выбранный из 0508AP, 0508AP2, 2018AP, 2018AP2, 2021AP, 2021AP2, 2021AP3, 2021AP4 или 2021AP5;

г) 5' праймер, выбранный из 1312AP2;

д) 5' праймер, выбранный из 1313AP или 1313AP2;

е) 5' праймер, выбранный из 1310AP2, 1310AP3, 1310AP4, 1310AP5 или 1312AP2;

ж) 5' праймер, выбранный из 1314AP или 1314AP2.

Специалистам в данной области очевидно, что комбинация по меньшей мере одного праймера из каждого из приведенных выше пунктов от а) до ж)

была бы особенно выгодной, поскольку она обеспечивала бы универсальную смесь 5'-праймеров, которая одновременно амплифицирует и модифицирует матричную нуклеиновую кислоту, которая кодирует сегменты генов IGHV1, IGHV2, IGHV3, IGHV4, IGHV5, IGHV6 и IGHV7 человека в одной реакции.

5 Также могут быть обстоятельства, при которых может быть предпочтительным подмножество праймеров, перечисленных в табл. 1 (или перечисленных пунктах от а) до ж) выше). Например, если сосредоточиться на матричной нуклеиновой кислоте, которая кодирует только IGHV1, IGHV2, IGHV3 человека, может оказаться полезным выбрать комбинацию праймеров
10 только в соответствующих строках табл. 1 (или только в пунктах (а)-(в) выше). Предпочтительность использования праймеров коррелирует, например, с семействами генов в платформе, используемой для получения антител или тяжелых цепей. Например, когда применяют трансгенного хозяина, несущего человеческий мини-локус, который содержит сегменты гена VH из семейств
15 генов IGHV1, IGHV5 и IGHV7, праймеры, перечисленные в табл. 1, соответствующие указанным семействам генов, предпочтительно используют для создания панели тяжелых цепей, включающей варианты на N-конце. В дополнение к праймерам, предоставленным выше, на основе положений, описанных в настоящем изобретении, специалист в данной области техники
20 может дополнительно разработать праймеры, применимые в каждом семействе генов VH для амплификации любого домена VH человека, который также включен в настоящее изобретение.

Описанные в настоящем изобретении способы включают проведение полимеразной цепной реакции (ПЦР) по меньшей мере с одним 5'-праймером,
25 по меньшей мере с одним 3'-праймером и нуклеиновой кислотой для получения амплифицированной нуклеиновой кислоты. По крайней мере, один 5'-праймер и нуклеиновая кислота (матрица) подробно описаны выше.

Можно использовать любой подходящий 3'-праймер или смесь 3'-праймеров. Специалисту в данной области очевидно, это относится к 3'-
30 праймерам, комплементарным нуклеиновой кислоте, кодирующей область FR4 кодируемого человеческого, гуманизированного или химерного домена VH, или 3'-праймеры, комплементарные нуклеиновой кислоте, кодирующей константный домен тяжелой цепи человека. Для доменов VH человека область FR4 кодируется перегруппированными сегментами гена J человека (или сегментами

гена J в контексте общей тяжелой цепи). Соответственно, дизайн подходящего 3'-праймера хорошо известен специалистам в данной области.

Ниже приведены типичные праймеры, которые включают область, комплементарную концу FR4, и включают сайты рестрикции BstEII и XhoI.

5 HuJH1/2xho = TATTGTTACCTCGAGACGGTGACCAGGGTGCC (SEQ ID NO: 12)

 HuJH3xho = TATTGTTACCTCGAGACGGTGACCATTGTCCC (SEQ ID NO: 13)

10 HuJH4/5xho = TATTGTTACCTCGAGACGGTGACCAGGGTTCC (SEQ ID NO: 14)

 HuJH6xho = TATTGTTACCTCGAGACGGTGACCGTGGTCCC (SEQ ID NO: 15)

15 Описанные в настоящем изобретении способы могут дополнительно включать стадию введения каждой амплифицированной и модифицированной нуклеиновой кислоты в вектор. Способы введения нуклеиновых кислот в векторы хорошо известны и включают расщепление рестрикционными ферментами и лигирование. Подходящие векторы описаны в настоящем описании и включают фагмиды или плазмиды.

20 Описанные в настоящем изобретении способы также могут дополнительно включать трансформацию или трансфекцию каждого вектора в клетку для создания библиотеки. Способы введения векторов в клетки хорошо известны. Подходящие клетки-хозяева описаны в настоящем описании и включают компетентные в отношении фага клетки, такие как компетентные в отношении фага клетки *E.coli* или компетентные в отношении фагов дрожжи.

25 Соответствующие библиотеки также описаны в настоящем описании.

Наборы

В настоящем изобретении также предусмотрены наборы. Наборы содержат множество описанных в настоящем изобретении 5'-праймеров. Набор может содержать по меньшей мере два, по меньшей мере три, по меньшей мере четыре, по меньшей мере пять, по меньшей мере шесть, по меньшей мере семь, по меньшей мере восемь, по меньшей мере девять и т. д. различных 5'-праймеров, описанных в настоящем изобретении. Необязательно наборы также содержат по меньшей мере один 3'-праймер, описанный в настоящем изобретении, в зависимости от природы репертуара человеческих,

30

гуманизированных или химерных доменов VN. Подробная информация о подходящих праймерах, приведенная выше, в равной степени применима здесь.

Компоненты набора могут быть размещены в контейнере, удобном для транспортировки.

- 5 Кроме того, наборы могут включать учебные материалы, содержащие инструкции (т. е. протоколы) по использованию материалов, входящих в набор. Хотя учебные материалы обычно содержат письменные или печатные материалы, они могут быть представлены на любом носителе, способном хранить такие инструкции и передавать их конечному пользователю.
- 10 Подходящие носители включают, но не ограничиваются ими, электронные носители данных (например, магнитные диски, ленты, картриджи, чипы) и оптические носители (например, CD-ROM). Средства массовой информации могут включать адреса интернет-сайтов, на которых размещены учебные материалы. Такие инструкции могут соответствовать любому из способов или
- 15 применений, подробно описанных в настоящем изобретении.

Фармацевтические композиции и способы их применения

- Настоящее изобретение также предусматривает фармацевтическую композицию, которая содержит антитело, фрагмент антитела или вариант
- 20 антитела и фармацевтически приемлемый носитель и/или разбавитель. Соответственно, настоящее изобретение относится к антителу, фрагменту антитела или варианту антитела, описанным в настоящем изобретении, для применения в лечении человека или животного с помощью терапии. Кроме того, настоящее изобретение предусматривает способ лечения больного человека или
- 25 животного, который включает введение человеку или животному терапевтически эффективного количества антитела, фрагмента антитела или варианта антитела, согласно описанию в настоящем изобретении. Количество антитела, фрагмента антитела или варианта антитела по настоящему изобретению, которое должно быть введено пациенту, обычно находится в терапевтическом окне, что означает,
- 30 что используется достаточное количество для получения терапевтического эффекта, при этом количество не превышает порогового значения, приводящего к неприемлемой степени побочных эффектов. Чем меньше количество антитела, фрагмента антитела или варианта антитела, необходимого для получения желаемого терапевтического эффекта, тем больше обычно будет

терапевтическое окно. Антитело, фрагмент антитела или вариант антитела по настоящему изобретению, оказывающие достаточное терапевтическое действие при низких дозах, соответственно, являются предпочтительными.

В настоящем описании ссылка на патентный документ или другой материал, который приводится в качестве предшествующего уровня техники, не должна восприниматься как признание того, что этот документ или материал был известен или что содержащаяся в нем информация была частью общеизвестных сведений на дату приоритета, касающихся любого пункта формулы настоящего изобретения.

Если в настоящем описании не указано иное, все используемые здесь технические и научные термины имеют то же значение, которое обычно понимают специалисты в области, к которой относится данное изобретение. Хотя любые способы и материалы, подобные или эквивалентные описанным в настоящем изобретении, находят применение в практике настоящего изобретения, здесь описаны предпочтительные способы и материалы. Соответственно, термины, определенные непосредственно ниже, более полно описанные со ссылкой на спецификацию в целом. Кроме того, в настоящем изобретении упоминание в единственном числе может также подразумевать множественное число, если из контекста явно не следует иное. Если иное не указано, нуклеиновые кислоты записывают слева направо в ориентации от 5' к 3'; аминокислотные последовательности записывают слева направо в ориентации от amino- к карбокси-концу, соответственно. Следует учитывать, что настоящее изобретение не ограничивается описанными конкретными методами, протоколами и реагентами, поскольку они могут варьировать в зависимости от контекста, в котором их используют специалисты в данной области.

Объекты настоящего изобретения поясняются приводимыми ниже примерами, которые однако не ограничивают рамок охвата настоящего изобретения.

ПРИМЕРЫ

Были идентифицированы новые праймеры, которые могут амплифицировать все переменные области, полученные из каждого функционального сегмента гена VH в каждом семействе генов VH человека. Новые праймеры модифицируют любой N-конец любого домена VH человека, полученного в результате рекомбинации любого функционального сегмента гена

VH человека, что приводит к предотвращению образования N-конца пироглутаминовой кислоты и/или увеличения экспрессии.

Приведенные ниже примеры демонстрируют изобретение с использованием ДНК, кодирующей переменные области, продуцируемые мышами MeMus MeMo®, и интеграцию такой ДНК в векторы при варьировании первой (или 5 первой и второй) N-концевой кодируемой аминокислоты переменного домена тяжелой цепи иммуноглобулина человека. Панели ДНК, кодирующие переменные области, выработанные двумя разными мышами MeMo®, успешно интегрированы в векторы, имеющие различающиеся кодируемые первую (или 10 первую и вторую) N-концевую аминокислоту. Такие мыши MeMo® имеют синтетические мини-локусы тяжелой цепи, которые содержат репрезентативный сегмент гена VH из семейств генов VH человека. Показано, что праймеры работают во всех подсемействах генов VH. Праймеры были оптимизированы для эффективности амплификации VH и разнообразия VH. Их используют для 15 успешного создания библиотек фагового дисплея и для последующей экспрессии Fab.

Обоснование исследования

Описанные ранее линии мышей MeMus (линия MeMo®) экспрессируют антитела с областями VH человека. После иммунизации этих мышей можно 20 выделить РНК с последующим синтезом кДНК и ПЦР-амплификацией областей VH. Следует отметить, что последовательности VH начинаются с E или Q.

В настоящее время разработаны праймеры для ПЦР, которые заменяют E или Q на N-конце всех последовательностей VH во время амплификации. Праймеры специально разработаны для амплификации незапланированных 25 репертуаров VH для любой переменной области тяжелой цепи, состоящей из любого рекомбинированного сегмента функционального гена V человека и семейства функциональных генов в репертуаре человека.

При изменении последовательности N-конца VH также следует соблюдать осторожность, чтобы не повлиять на такие свойства антител, как структура, 30 связывание антигена и стабильность, а также отщепление сигнального пептида (SP). Частоты аминокислот в отдельных положениях SP и ассоциированных зрелых белков были проанализированы для 2352 секретируемых белков эукариот, грамположительных бактерий и грамотрицательных бактерий [Choo, Ranganathan, 2008]. Этот анализ показал сходство, а также различия между

группами видов. В целом, аминокислотные предпочтения в основном наблюдаются внутри SP, однако определенные предпочтения также наблюдаются для первых нескольких остатков зрелых пептидов:

Из анализа этих данных следует, что A и Q являются предпочтительными в первом положении как у эукариотических (25% A или Q), так и граммотрицательных (54% A или Q) зрелых белков. У эукариот P относительно часто встречается во втором (16%) и четвертом (11%) положении. У граммотрицательных бактерий D, E, P и T часто встречаются во втором положении (56% всех проанализированных белков имеют один из этих четырех остатков). Для третьей и четвертой позиции T преобладает на третьей (11%) и четвертой (13%) позиции.

С другой стороны, некоторые аминокислоты явно недостаточно представлены в определенных положениях. Например, W встречается с частотой всего около 1% в первой позиции у эукариот и граммотрицательных бактерий. Отсюда делают вывод, что, поскольку определенные аминокислоты являются или предпочтительными, или нежелательными, оптимизация отщепления SP может быть возможным путем адаптации последовательности первых нескольких остатков зрелых пептидов.

При изменении последовательности N-конца у VH модифицированные последовательности должны хорошо комбинироваться как с прокариотическими (бактериальными), так и с эукариотическими SP.

К бактериальным SP, например, относится MKYLLPTAAAGLLLLAAQPAMA (SEQ ID NO: 1). К эукариотическим SP, например, относится MGWSCIILFLVLLLAQPAMA (SEQ ID NO: 4). Эти сигнальные пептиды приведены в качестве примеров, но они не ограничивают рамок охвата настоящего изобретения, варианты осуществления которого приведены ниже.

Чтобы иметь возможность проверить *in silico*, что отщепление SP от модифицированной VH, по крайней мере, так же хорошо, как и от соответствующей VH дикого типа (WT), был разработан набор из восемнадцати репрезентативных последовательностей, каждая из которых содержит SP и первые 20 N-концевых аминокислот области VH (9 областей VH в сочетании с 2 разными последовательностями SP = 18 репрезентативных последовательностей). Выбранные последовательности области VH специально

выбраны в качестве репрезентативных последовательностей для всех подсемейств генов VN. Полагают, что остатки VN за пределами положения 20 в области VN существенно не влияют на отщепление SP [Choo, Ranganathan, 2008].

Первый остаток VN в каждой из 18 последовательностей («положение 1») меняли, чтобы включить все 20 возможных аминокислот, в результате чего получилось $18 \times 20 = 360$ последовательностей. Всем последовательностям был присвоен код формата P#X или E#X, где:

P = прокариотическая SP; E = эукариотическая SP

= внутренний идентификационный номер

X = аминокислота в положении 1 в VN

Например, P1A содержит прокариотическую SP, и в последовательности VN из 20 остатков первая аминокислота E, кодируемая VN, заменена на A.

Чтобы изучить влияние первого остатка VN на расщепление SP, все 360 последовательностей проанализированы *in silico* с применением средства прогнозирования SignalP 4.1 на сайте www.cbs.dtu.dk/services/SignalP [Petersen с соавт., 2011] с использованием следующих параметров:

группа организмов: «граммотрицательные бактерии» для последовательностей P#X и «эукариоты» для последовательностей E#X

выходной формат: «Короткий (без графики)»

все остальные параметры: стандарт/по умолчанию

Для всех 360 последовательностей прогноз положения сайта отщепления SP получен правильно. Так называемые оценки D (дискриминация) прогноза сравнивают для всех 360 последовательностей (см. фиг. 2). Высокий показатель D указывает на высокую вероятность того, что последовательность, предшествующая области VN, на самом деле является сигнальным пептидом. Предполагают, что более высокий показатель D соответствует более высокой вероятности эффективного расщепления SP. На фиг. 2 также приведены немодифицированные частоты аминокислот в положении 1 для панели из 307 белков грамотрицательных бактерий и 1877 эукариотических белков, содержащих N-концевые SP [Choo, Ranganathan, 2008; дополнительный файл 2]. В настоящем изобретении предполагают, что более высокая частота соответствует более высокой вероятности эффективного отщепления SP. В последовательностях P#X самые высокие баллы D наблюдают для X=A (в

среднем 0,911), а А является наиболее частой аминокислотой в секреторных белках прокариот (41,7%). В качестве другого примера, в последовательностях E#X самые низкие оценки D наблюдаются для X=P (в среднем 0,863), а Р является наименее часто встречающейся аминокислотой в эукариотических секреторных белках (0,3%).

На основании оценки D следующие три остатка, имеющие наивысшую оценку, выбраны в качестве потенциальных альтернативных остатков в первом положении VH:

А: большинство оценок/частот (намного) выше по сравнению с Е и Q; это, безусловно, наиболее часто встречающаяся аминокислота в немодифицированных SP

D: боковая цепь химически подобна Е; баллы/частоты сопоставимы с Е и Q

S: часть баллов/частот немного выше по сравнению с Е и/или Q

Вместе с последовательностями, содержащими Е или Q в положении 1, остается $5 \times 9 \times 2 = 90$ последовательностей для дальнейшего анализа (5 аминокислот в положении 1×9 последовательностей области $VH \times 2$ SP).

Для каждой из 90 последовательностей (т. е. с А, D, Е, Q или S в положении 1) варьируют второй остаток VH («положение 2»). Следующие 7 остатков опущены, так как они редко или вообще не встречаются в немодифицированных SP [Choo, Ranganathan, 2008]; частота в положении 2 граммотрицательных/эукариотических SP даны в скобках:

C (0,0% / 1,9%)

F (1,0% / 2,2%)

H (0,3% / 2,5%)

M (0,0% / 0,7%)

R (0,3% / 4,7%)

W (1,6% / 0,9%)

Y (0,3% / 2,5%)

Это оставляет следующие 13 остатков для изменения в положении 2:

A (5,9% / 3,7%)

D (17,3% / 8,0%)

E (16,9% / 8,8%)

G (6,5% / 5,0%)

I (2,9% / 3,7%)

K (1,6% / 5,0%)

L (1,6% / 4,6%)

N (5,9% / 4,3%)

P (10,8% / 15,8%)

5 Q (4,9% / 4,9%)

S (5,5% / 9,0%)

T (10,8% / 5,6%)

V (5,9% / 6,3%)

Вышеуказанные результаты составляют $90 \times 13 = 1170$

10 последовательностей для анализа.

Получают 1170 последовательностей. Всем последовательностям присвоен код формата P#XZ или E#XZ, где:

P = прокариотическая SP; E = эукариотическая SP

= внутренний идентификационный номер

15 X = аминокислота в VH в положении 1

Z = аминокислота в VH в положении 2

Например, P1AD содержит прокариотическую SP и последовательность VH из 20 остатков, где первая кодируемая VH аминокислота E заменена на A, а вторая кодируемая VH аминокислота V заменена на D.

20 Чтобы изучить влияние первых двух различных остатков VH на отщепление SP, все 1170 последовательностей проанализированы *in silico* с использованием инструмента прогнозирования SignalP 4.1 с параметрами, дополнительно предусмотренными в настоящем изобретении.

Для всех 1170 последовательностей прогноз положение сайта отщепления
25 SP был правильным. Были сопоставлены оценки D для 1170 последовательностей (фиг. 3). Установлено, что в целом влияние идентичности остатка в положении 2 на расщепление SP не зависит существенно от идентичности остатка в положении 1. Например, относительно низкие показатели D наблюдают с K в положении 2, независимо от остатка в положении
30 1, которым может быть A, D, E, Q или S. Кроме того, самые высокие баллы обычно получают для последовательностей с A в положении 1.

Результаты на фиг. 3 используют следующим образом для определения оптимальной комбинации остатков в положениях 1 и 2 для каждого из девяти праймеров, что показано на фиг 4.

Для каждой из 18 комбинаций SP и генных сегментов VH идентифицированы варианты с наивысшим показателем D. Например, для 65 последовательностей кода P1XZ вариант с AV в положении 1+2 имеет самый высокий показатель D (0,907). Интересно, что все лучшие варианты последовательностей с бактериальным SP имеют AV (или иногда также AT) в положении 1+2. Сходным образом лучшие варианты последовательностей с эукариотической SP имеют AP (или иногда также AV) в положении 1+2.

Очевидно, что на основании представленных в настоящем изобретении положений специалист в данной области может также идентифицировать отдельные праймеры для использования с прокариотическим или эукариотическим сигнальным пептидом.

Предпочтительно не требуются отдельные праймеры для использования VH с прокариотической или эукариотической SP, консенсусная последовательность в положении 1+2 определена для каждого из 9 праймеров, что в сочетании с обеими SP дает более высокую оценку/частоту, чем соответствующая последовательность дикого типа (WT). На основании этих данных консенсусная последовательность была определена как AP:

В сочетании с бактериальной SP последовательность AP в положении 1+2 дает оценку D, которая на 0,030–0,065 (в среднем 0,040) выше, чем у последовательности WT.

В сочетании с эукариотической SP последовательность AP в положении 1+2 дает оценку D, которая на 0,002–0,005 (в среднем 0,003) выше, чем у последовательности WT.

Аминокислота А является наиболее часто встречающейся аминокислотой в положении 1 в сочетании с немодифицированными бактериальными (41,7%) и эукариотическими (13,5%) SP (см. фиг. 3).

Аминокислота Р является третьей, наиболее часто встречающейся аминокислотой в положении 2 в сочетании с немодифицированными бактериальными SP (10,8%) и наиболее часто встречающейся аминокислотой в положении 2 в сочетании с немодифицированными эукариотическими SP (15,8%; см. фиг. 3).

Были разработаны новые праймеры FW (5'), которые аналогичны праймерам, перечисленным на фиг. 5, за исключением того, что во всех новых праймерах FW первые два кодона VH изменены таким образом, что они

кодируют AP вместо EV, EQ, QI или QL (названные 0508AP, 1308AP, 1310AP и т. д.). Из-за вырожденности генетического кода для A (GCN) и P (CCN) существует четыре разных кодона. Для праймеров были выбраны кодоны, которые наиболее гомологичны кодонам в текущих праймерах (это зависит от праймера). Были приняты меры, чтобы не ввести новые сайты клонирования SfiI, BstEII и XhoI. Последовательности полученных 9 новых праймеров представлены на фиг. 5; белковые трансляции этих праймеров приведены на фиг. 6.

Сконструированные праймеры тестируют параллельно с текущими праймерами, которые не модифицируют первую и вторую аминокислоты, кодируемые областью VH. Резюме результатов выглядит следующим образом:

Результаты экспериментов

Дизайн праймеров и анализ эффективности амплификации

Используют две разные линии мышей MeMo®. Используют кДНК на основе нуклеиновой кислоты этих мышей для амплификации VH с новыми праймерами и текущими праймерами. Эффективность амплификации анализируют для каждого праймера путем сравнения выхода продуктов ПЦР на агарозном геле. См. фиг. 7.

При применении следующих праймеров получают удовлетворительный выход ПЦР продукта:

Для семейства IGHV1: 1308AP; 2018AP

Для семейства IGHV3: 0508AP; 2018AP

Для семейства IGVH4: 1312AP; 2019AP

Для семейства IGHV5: 1313AP

Для семейства IGHV7: 1314AP

Для следующих праймеров получают низкий выход или нулевой:

Для семейства IGHV1: 2020AP (нулевой выход)

Для семейства IGHV2: 1310AP (низкий выход)

Три новых (AP) праймера не показали хороших результатов в ПЦР для амплификации сегментов гена VH из кДНК, поэтому дизайн праймера был пересмотрен. Плохие результаты могут быть вызваны мотивами последовательности, такими как длинные участки G/C, которые могут вызывать нежелательные вторичные структуры, такие как димеры и шпильки. Проверка используемых последовательностей праймеров показала, что введение мутаций

приводит к относительно длинным участкам G/C длиной 10-11 п.н., частично из-за использования кодона GCG для Ala и CCC/CCG для Pro. Хотя такие отрезки также присутствуют в праймерах, проявивших хорошие результаты, любой отрицательный эффект может зависеть от свойств последовательности в других местах праймеров, которые варьируют в зависимости от праймера.

Вырожденность генетического кода позволяет уменьшить длину отрезков G/C за счет выбора других кодонов. Поэтому были разработаны новые праймеры (версия 2; AP2; обозначенные 0508AP2, 1310AP2 и т. д.), в которых кодоны GCA и CCT используют для Ala и Pro, соответственно. Это было сделано не только для трех неэффективных праймеров, которые не проявили хорошей активности, но и для остальных 6 праймеров.

Чтобы проверить *in silico*, ожидается ли, что праймеры версии 2/AP2 будут работать лучше, было проанализировано подмножество с использованием возможностей Oligo Analyzes программного обеспечения Vector NTI. Например, для праймера DO_2020 и его обновленных версий по прогнозу отмечены следующие возможные нежелательные димеры и шпильки:

2020AP: 65 димеров и 26 шпилек (всего 91 объект). Это относительно много и может служить объяснением (частично), почему этот праймер не проявил себя эффективно.

2020AP2: 49 димеров и 19 шпилек (всего 68 объектов). Это намного меньше чем у 2020AP. Это означает, что 2020AP2 может действовать лучше, чем 2020AP.

Также было проверено, может ли дальнейшее снижение содержания G/C привести к еще меньшему количеству вторичных структур согласно прогнозу:

Когда кодон Ala GCC, который предшествует VH, был заменен на GCT или GCA в праймере 2020AP2, по прогнозу было обозначено 55 или 54 возможных димера и 22 возможные шпильки (всего 77 или 76 объектов). Это больше, чем для 2020AP2, поэтому кодон Ala был сохранен как GCC.

Аналогичные результаты прогноза получены для других праймеров (не показаны).

Выравнивание нового (AP) и оптимизированного (AP2) праймеров показано на фиг. 8.

Было показано, что большинство праймеров AP2 дают достаточное количество продукта ПЦР. См. фиг. 9. Однако праймеры 0508AP2 и 1310AP2;

(дорожки № 2 и 4 для мыши 1), похоже, не показывают (значительного) улучшения по сравнению с предыдущим дизайном. Праймер 2020AP2 (дорожка № 5 для мыши 2) теперь работает, тогда как предыдущий дизайн (2020AP) не работал. Отрицательные контроли (-) дали отрицательный результат.

5 Все новые версии праймеров (AP и AP2) тестируют параллельно для анализа эффективности амплификации. См. фиг. 10. Ранее наблюдали, что праймер 2020AP2 (№ 8 мышь 2) работает, тогда как 2020AP (№ 7 мышь 2) не работает. Для обеих мышей праймер 0508AP (№ 1), по-видимому, работает немного лучше, чем 0508AP2 (№ 2). Во всех других реакциях выходы AP2
10 аналогичны или немного выше по сравнению с AP.

Была начата дальнейшая оптимизация праймера 1310AP. Кроме того, был протестирован дополнительный праймер к семейству IGHV3 (2021AP): протестировано 5 различных вариантов праймера 1310AP вместе с пятью различными вариантами праймера 2021AP. См. фиг. 11. Новый праймер 1310AP5
15 (№ 6 мышь 1) дает явно лучшие результаты, чем праймер 1310AP2 (№ 3 мышь 1), который ранее работал лучше всего. Все новые праймеры от 2021AP до 2021AP5 (№№ 8-12) работают одинаково хорошо.

В целом, следующие праймеры способны индуцировать достаточное количество продукта ПЦР:

20 Для семейства IGHV1: 1308AP, 1308AP2, 2020AP2, 2018AP, 2018AP2
Для семейства IGHV2: 1310AP2, 1310AP3, 1310AP4, 1310AP5
Для семейства IGHV3: 0508AP, 0508AP2, 2018AP, 2018AP2, 2021AP, 2021AP2, 2021AP3, 2021AP4, 2021AP5
Для семейства IGVN4: 1312AP2 / 2019AP2
25 Для семейства IGHV5: 1313AP, 1313AP2
Для семейства IGHV6: 1310AP2, 1310AP3, 1310AP4, 1310AP5, 1312AP2/2019AP2
Для семейства IGHV7: 1314AP, 1314AP2.

30 Продукты ПЦР различных вышеуказанных праймеров очищают и лигируют в вектор для трансформации в фаг-компетентные бактериальные клетки. Созданы библиотеки фагового дисплея. Для определения частоты вставок и разнообразия последовательностей проводили ПЦР колоний и секвенирование.

Анализ разнообразия VH

Библиотеки фагового дисплея (размером приблизительно 10^6 - 10^7) строят путем клонирования амплифицированных сегментов гена VH в вектор Fab-фаг. Отдельные клоны из этих библиотек секвенируют для определения репрезентативности различных семейств генов VH. Полученные последовательности анализируют для определения репрезентативности различных семейств генов VH. Результат (т.е. процент каждой VH в общем числе амплифицированных сегментов гена VH) оказался сходным для праймеров, индуцирующих новый вариант, и праймеров, которые амплифицируют неизменные последовательности VH, демонстрируя праймеры, которые генерируют варианты, содержащие первые два положения последовательностей VH, которые не влияют на представление соответствующих сегментов гена V и семейств VH в полученной фаговой библиотеке.

Анализ экспрессии Fab

Отдельные клоны из библиотек используют для получения неочищенных периплазматических экстрактов, содержащих растворимые Fab. Концентрации Fab определяют с использованием количественного анализа Octet. Было обнаружено, что выход Fab в большинстве продуктов находится в одном и том же диапазоне (приблизительно 10-15 мкг/мл) (данные не показаны). Fab с мутацией AP вырабатываются хорошо и в целом приводят к повышенным средним выходам, чем Fab WT (11,4 против 10,0 мкг/мл), демонстрируя полезность праймеров, генерирующих варианты, и вариабельность N-концевых вариабельных областей.

Анализ целостности Fab

Подмножество полученных растворимых Fab подвергали SDS-PAGE и Вестерн-блоттингу. На полученных пятнах были видны полосы ожидаемых размеров (данные не показаны).

Краткое описание результатов

Эксперименты показали, что описанные здесь праймеры изобретения можно использовать для создания библиотек фагового дисплея с разнообразным репертуаром VH, и что Fab-фрагменты могут экспрессироваться представителями этих библиотек. Новые праймеры можно использовать для

амплификации кДНК, кодирующей генные сегменты VH, во всем репертуаре генных сегментов вариабельной области человека, при этом модифицируя N-конец кодируемого вариабельного домена для предотвращения образования N-концевой пироглутаминовой кислоты и/или повышения экспрессии Fab.

5 Внимание читателей обращено на все статьи и документы, которые были поданы одновременно с этой спецификацией или до нее в связи с этой заявкой и которые открыты для публичного ознакомления с этой спецификацией, и содержание всех бумаг и документов включено в настоящий документ посредством ссылок.

10 Все признаки, описанные в настоящем изобретении (включая любые прилагаемые пункты формулы изобретения, реферат и фигуры), и/или все стадии любого способа или процесса, раскрытые таким образом, могут быть объединены в любую комбинацию, за исключением комбинаций, в которых по крайней мере некоторые из таких признаков и/или стадий являются
15 взаимоисключающими.

 Каждая функция, раскрытая в настоящем описании (включая любые прилагаемые пункты формулы изобретения, реферат и фигуры), может быть заменена альтернативными функциями, служащими той же, эквивалентной или аналогичной цели, если прямо не указано иное. Таким образом, если прямо не
20 указано иное, каждый раскрытый признак является лишь одним примером общего ряда эквивалентных или подобных признаков.

 Изобретение не ограничивается деталями любых предшествующих вариантов его осуществления. Изобретение распространяется на любой новый элемент или любую новую комбинацию признаков, раскрытых в данном
25 описании (включая любые прилагаемые пункты формулы изобретения, реферат и фигуры), или на любой новый элемент или любую новую комбинацию стадий любого описанного способа или процесса.

Литературные источники

30 Choo K. H., Ranganathan S. (2008). Flanking signal and mature peptide residues influence signal peptide cleavage. *BMC Bioinformatics* 9, приложение 12, S15.

 Fowler E., Moyer M., Krishna R. G., Chin C. C. Q., Wold F. (1996). Removal of N-terminal blocking groups from proteins, *Current Protocols in Protein Science*.

 Jefferis R. (2016). Review article. Posttranslational modifications and the immunogenicity of biotherapeutics. *Journal of Immunology Research* 2016.

Liu YD, Goetze AM, Bass RB, Flynn GC (2011). N-terminal glutamate to pyroglutamate conversion in vivo for human IgG2 antibodies. *J Biol Chem.* 2011, 286(13):11211-11217.

5 Petersen T. N., Brunak S., von Heijne G., Nielsen H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* 8, 785-786.

Yu L., Vizek A., Huff M.B., Young M., Remmele R.L. Jr, He B. (2006). Investigation of N-terminal glutamate cyclization of recombinant monoclonal antibody in formulation development. *J. Pharm. Biomed. Anal.* 42(4): 455-63.

10 Ambrogelly A., Gozo S., Katiyar A., Dellatore S., Kune Y., Bhat R., Sun J., Li N., Wang D., Nowak C., Neill A., Ponniah G., King C., Mason B., Beck A, Liu H. (2018). Analytical comparability of recombinant monoclonal antibody therapeutics. *MAbs* 10(4): 513-538

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Полипептид, содержащий вариабельный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, причем
5 указанный вариабельный домен содержит:

- N-концевую аминокислоту, выбранную из группы, состоящей из:
аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина,
лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана и
тирозина; или

10 - N-концевую последовательность, выбранную из группы, состоящей из
последовательностей: аланин-пролин, аланин-аспартат, аланин-глутамат,
аланин-треонин, аланин-валин, аланин-серин и аланин-лейцин.

15 2. Полипептид по п. 1, причем вариабельный домен тяжелой цепи
человеческого, гуманизированного или химерного иммуноглобулина содержит
N-концевую последовательность аланин-пролин.

20 3. Полипептид по п.п. 1 или 2, причем полипептид содержит сигнальный
пептид выше по цепи N-концевой аминокислоты вариабельного домена тяжелой
цепи человеческого, гуманизированного или химерного иммуноглобулина.

4. Антитело, вариант антитела или фрагмент антитела, содержащие
полипептид по п.п. 1 или 2.

25 5. Нуклеиновая кислота, кодирующая полипептид, антитело, вариант
антитела или фрагмент антитела по п.п. 1-4, где необязательно нуклеиновая
кислота находится в векторе или фаге.

30 6. Библиотека, содержащая не менее примерно 10^6 различных нуклеиновых
кислот, векторов или фагов по п. 5.

7. Способ одновременной амплификации и модификации нуклеиновой
кислоты, кодирующей вариабельный домен тяжелой цепи человеческого,
гуманизированного или химерного иммуноглобулина, включающий:

(а) получение нуклеиновой кислоты, кодирующей переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина; и

(б) осуществление полимеразной цепной реакции по меньшей мере с одним 5'-праймером, по меньшей мере с одним 3'-праймером и нуклеиновой кислотой для получения амплифицированной нуклеиновой кислоты,

причем по меньшей мере один 5'-праймер включает нуклеиновую кислоту с сайтом модификации, который внедряет модификацию в амплифицированную нуклеиновую кислоту таким образом, что амплифицированная нуклеиновая кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, включающий N-концевую аминокислоту, выбранную из группы, состоящей из: аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и валина.

8. Способ по п. 7, где каждая амплифицированная нуклеиновая кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевой аланин, при этом каждая амплифицированная нуклеиновая кислота необязательно кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевую последовательность аланин-пролин.

9. Способ по п. 7 или 8, где по меньшей мере один 5'-праймер кодирует сигнальный пептид или часть сигнального пептида по ходу транскрипции от сайта модификации.

10. Способ по любому из п.п. 7-9, где нуклеиновая кислота (кислоты) на стадии (а) являются кДНК, причем необязательно способ включает предварительную стадию выделения нуклеиновых кислот из В-клеток животного и получения кДНК из нуклеиновых кислот для получения нуклеиновой кислоты (кислот), полученных на стадии (а); при этом также необязательно

i) животное иммунизируют интересующим антигеном, а нуклеиновые кислоты из В-клеток кодируют тяжелые цепи, обладающие специфичностью и аффинностью к представляющему интерес антигену; и/или

5 ii) животное представляет собой трансгенную мышь, содержащую локус тяжелой цепи иммуноглобулина человека; и/или

iii) животное является трансгенной мышью, содержащую общую легкую цепь.

10 11. Способ по любому из п.п. 7-10, **отличающийся тем**, что стадия (а) включает получение множества различных нуклеиновых кислот, кодируемых по меньшей мере одним сегментом рекомбинированного гена человека, или основывающихся на них, выбранных из каждого из следующих семейств генов человека: IGHV1, IGHV2, IGHV3, IGHV4, IGHV5, IGHV6 и IGHV7.

15 12. Способ по любому из п.п. 7-11, **отличающийся тем**, что указанный способ включает:

(а) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV1, с использованием 5'-праймера, выбранного из 1308AP, 1308AP2, 2020AP2, 2018AP или 2018AP2; и/или

20 (б) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV2, с использованием 5'-праймера, выбранного из 1310AP2, 1310AP3, 1310AP4 или 1310AP5; и/или

(в) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV3, с использованием 5'-праймера, выбранного из 0508AP, 0508AP2, 2018AP, 2018AP2, 2021AP, 2021AP2, 2021AP3, 2021AP4 или 2021AP5; и/или

(г) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV4, с использованием 5'-праймера, который представляет собой 1312AP; и/или

30 (д) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV5, с использованием 5'-праймера, выбранного из 1313AP или 1313AP2;

(е) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV6, с использованием 5'-праймера, выбранного из 1310AP2, 1310AP3, 1310AP4, 1310AP5 или 1312AP2; и/или

5 (ж) амплификацию и модификацию нуклеиновой кислоты, кодируемой геном семейства IGHV7, с использованием 5'-праймера, выбранного из 1314AP или 1314AP2.

13. 5'-праймер для амплификации и модификации любой нуклеиновой кислоты, которая кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, выбранный из одного или 10 нескольких из следующих семейств генов VH человека или основанный на них: IGHV1, IGHV2, IGHV3, IGHV4, IGHV5, IGHV6 и IGHV7, где праймер содержит сайт модификации, который внедряет модификацию в амплифицированную нуклеиновую кислоту таким образом, что амплифицированная нуклеиновая 15 кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевую аминокислоту, выбранную из группы, состоящей из: аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и 20 валина.

14. Праймер по п. 13, где сайт модификации таков, что амплифицированная нуклеиновая кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, 25 содержащий N-концевой аланин, причем необязательно сайт модификации таков, что амплифицированная нуклеиновая кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий на N-конце аланин-пролин.

30 15. Набор, включающий по меньшей мере один 5'-праймер, выбранный из каждой из следующих групп:

(а) 1308AP, 1308AP2, 2020AP2, 2018AP или 2018AP2;

(б) 1310AP2, 1310AP3, 1310AP4 или 1310AP5;

(в) 0508AP, 0508AP2, 2018AP, 2018AP2, 2021AP, 2021AP2, 2021AP3, 2021AP4 или 2021AP5;

(г) 1312AP2;

(д) 1313AP или 1313AP2;

5 (е) 1310AP2, 1310AP3, 1310AP4, 1310AP5 или 1312AP2;

(ж) 1314AP2 или 1314AP.

16. Способ получения антитела, варианта антитела или фрагмента антитела, включающий:

10 - модификацию нуклеиновой кислоты, которая кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина таким образом, что модифицированная нуклеиновая кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевую аминокислоту,
15 выбранную из группы, состоящей из: аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и валина;

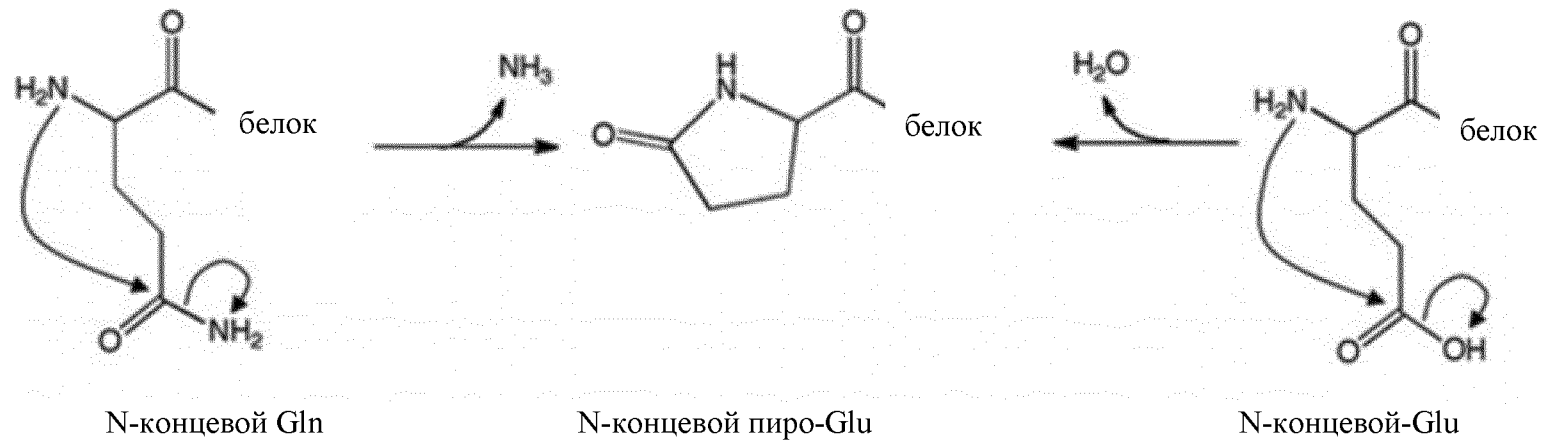
- последующее использование технологии скрининга антител для идентификации переменного домена тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащего N-концевую аминокислоту, выбранную из аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и валина для связывания с антигеном-мишенью;

25 - отбор переменного домена тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, который связывает антиген-мишень; и

- применение указанного переменного домена тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина для
30 разработки терапевтического антитела, варианта антитела или фрагмента антитела без дополнительной модификации N-концевой аминокислоты.

17. Способ снижения образования пироглутамата в переменном домене тяжелой цепи человеческого, гуманизированного или химерного

иммуноглобулина, включающий: модификацию нуклеиновой кислоты, кодирующей переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, таким образом, что модифицированная нуклеиновая кислота кодирует переменный домен тяжелой цепи человеческого, гуманизированного или химерного иммуноглобулина, содержащий N-концевую аминокислоту, выбранную из группы, состоящей из: аланина, аргинина, аспарагина, аспартата, глицина, гистидина, изолейцина, лейцина, лизина, метионина, фенилаланина, пролина, серина, треонина, триптофана, тирозина и валина.



Фигура 1.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	avg.	Естеств. частота (%)			E1	E2	E3	E4	E5	E6	E7	E8	E9	avg.	Естеств. частота (%)
A	0.907	0.919	0.917	0.904	0.912	0.915	0.907	0.899	0.919	0.911	41.7		A	0.949	0.948	0.950	0.945	0.948	0.947	0.950	0.942	0.948	0.947	13.5
C	0.873	0.890	0.887	0.871	0.881	0.884	0.873	0.861	0.890	0.879	0.0		C	0.944	0.943	0.947	0.940	0.943	0.941	0.945	0.936	0.943	0.942	3.4
D	0.844	0.866	0.866	0.846	0.855	0.860	0.845	0.838	0.866	0.854	7.2		D	0.944	0.944	0.945	0.942	0.944	0.941	0.946	0.940	0.944	0.943	5.3
E	0.834	0.857	0.858	0.836	0.845	0.850	0.835	0.826	0.857	0.844	6.2		E	0.946	0.945	0.946	0.943	0.945	0.943	0.948	0.942	0.945	0.945	7.8
F	0.878	0.892	0.616	0.874	0.884	0.887	0.878	0.863	0.892	0.852	1.0		F	0.951	0.951	0.951	0.947	0.950	0.948	0.952	0.945	0.951	0.950	3.6
G	0.868	0.886	0.883	0.866	0.877	0.880	0.868	0.858	0.886	0.875	6.2		G	0.938	0.938	0.942	0.935	0.938	0.936	0.939	0.930	0.938	0.937	7.1
H	0.848	0.869	0.867	0.847	0.858	0.862	0.848	0.835	0.869	0.856	1.0		H	0.943	0.943	0.944	0.940	0.943	0.940	0.944	0.937	0.943	0.942	1.8
I	0.857	0.875	0.570	0.855	0.866	0.868	0.857	0.840	0.875	0.829	0.3		I	0.947	0.947	0.948	0.943	0.946	0.944	0.948	0.940	0.947	0.946	3.3
K	0.829	0.854	0.853	0.829	0.841	0.845	0.829	0.819	0.854	0.839	2.9		K	0.941	0.940	0.942	0.938	0.940	0.938	0.943	0.935	0.940	0.940	4.6
L	0.865	0.880	0.563	0.861	0.872	0.874	0.865	0.846	0.880	0.834	3.3		L	0.950	0.950	0.951	0.946	0.949	0.948	0.951	0.943	0.950	0.949	8.5
M	0.880	0.894	0.892	0.877	0.886	0.889	0.880	0.865	0.894	0.884	0.0		M	0.945	0.944	0.947	0.941	0.944	0.942	0.946	0.937	0.944	0.943	1.1
N	0.862	0.879	0.879	0.861	0.870	0.874	0.862	0.851	0.879	0.869	3.6		N	0.939	0.938	0.941	0.936	0.938	0.936	0.940	0.932	0.938	0.938	2.5
P	0.836	0.854	0.854	0.834	0.845	0.848	0.836	0.816	0.854	0.842	0.0		P	0.864	0.862	0.888	0.861	0.863	0.860	0.866	0.842	0.862	0.863	0.3
Q	0.864	0.883	0.883	0.864	0.873	0.877	0.864	0.857	0.883	0.872	12.1		Q	0.949	0.948	0.948	0.946	0.948	0.946	0.950	0.944	0.948	0.947	11.0
R	0.813	0.840	0.839	0.815	0.826	0.830	0.814	0.801	0.840	0.824	0.7		R	0.941	0.941	0.943	0.939	0.941	0.939	0.943	0.935	0.941	0.940	5.0
S	0.871	0.887	0.885	0.869	0.879	0.882	0.871	0.858	0.887	0.877	4.2		S	0.943	0.942	0.945	0.939	0.942	0.940	0.944	0.936	0.942	0.941	8.2
T	0.867	0.882	0.881	0.865	0.874	0.877	0.866	0.852	0.882	0.872	3.9		T	0.944	0.943	0.946	0.940	0.943	0.941	0.945	0.937	0.943	0.942	4.3
V	0.870	0.886	0.883	0.867	0.878	0.880	0.870	0.855	0.886	0.875	3.3		V	0.947	0.947	0.949	0.943	0.946	0.945	0.948	0.940	0.947	0.946	4.5
W	0.867	0.883	0.579	0.864	0.874	0.877	0.867	0.853	0.883	0.839	0.7		W	0.947	0.947	0.950	0.944	0.947	0.946	0.948	0.940	0.947	0.946	1.0
Y	0.863	0.883	0.881	0.862	0.873	0.876	0.863	0.853	0.883	0.871	2.0		Y	0.944	0.944	0.946	0.941	0.944	0.942	0.946	0.939	0.944	0.943	3.1

Фигура 2.

Естеств.											Естеств.												
	P1A	P2A	P3A	P4A	P5A	P6A	P7A	P8A	P9A	avg.	частота (%)		E1A	E2A	E3A	E4A	E5A	E6A	E7A	E8A	E9A	avg.	частота (%)
A	0.900	0.910	0.915	0.897	0.903	0.907	0.900	0.897	0.910	0.904	5.9	A	0.932	0.931	0.943	0.928	0.931	0.930	0.933	0.928	0.931	0.932	3.7
D	0.888	0.905	0.911	0.888	0.895	0.901	0.888	0.888	0.905	0.897	17.3	D	0.946	0.946	0.950	0.944	0.946	0.944	0.948	0.944	0.946	0.946	8.0
E	0.882	0.900	0.906	0.882	0.890	0.895	0.882	0.882	0.900	0.891	16.9	E	0.947	0.947	0.951	0.945	0.947	0.945	0.949	0.945	0.947	0.947	8.8
G	0.897	0.911	0.915	0.895	0.902	0.906	0.897	0.895	0.911	0.903	6.5	G	0.939	0.939	0.947	0.936	0.939	0.937	0.941	0.936	0.939	0.939	5.0
I	0.899	0.912	0.917	0.897	0.904	0.908	0.899	0.897	0.912	0.905	2.9	I	0.945	0.945	0.950	0.941	0.945	0.943	0.946	0.941	0.945	0.945	3.7
K	0.856	0.875	0.884	0.855	0.864	0.869	0.856	0.855	0.875	0.865	1.6	K	0.942	0.942	0.947	0.940	0.942	0.940	0.944	0.940	0.942	0.942	5.0
L	0.902	0.915	0.918	0.899	0.907	0.910	0.902	0.899	0.915	0.907	1.6	L	0.946	0.946	0.952	0.942	0.946	0.944	0.947	0.942	0.946	0.946	4.6
N	0.901	0.914	0.919	0.899	0.906	0.910	0.901	0.899	0.914	0.907	5.9	N	0.946	0.946	0.950	0.944	0.946	0.944	0.948	0.944	0.946	0.946	4.3
P	0.899	0.913	0.918	0.897	0.905	0.909	0.898	0.897	0.913	0.905	10.8	P	0.950	0.950	0.953	0.948	0.950	0.949	0.952	0.948	0.950	0.950	15.8
Q	0.890	0.905	0.911	0.889	0.896	0.900	0.890	0.889	0.905	0.897	4.9	Q	0.946	0.945	0.950	0.943	0.945	0.944	0.947	0.943	0.945	0.945	4.9
S	0.900	0.912	0.916	0.897	0.904	0.908	0.899	0.897	0.912	0.905	5.5	S	0.945	0.945	0.950	0.942	0.945	0.943	0.946	0.942	0.945	0.945	9.0
T	0.906	0.918	0.922	0.904	0.910	0.914	0.905	0.904	0.918	0.911	10.8	T	0.947	0.947	0.951	0.944	0.947	0.945	0.948	0.944	0.947	0.947	5.6
V	0.907	0.919	0.922	0.904	0.912	0.915	0.907	0.904	0.919	0.912	5.9	V	0.949	0.948	0.953	0.945	0.948	0.947	0.950	0.945	0.948	0.948	6.3

Естеств.											Естеств.												
	P1D	P2D	P3D	P4D	P5D	P6D	P7D	P8D	P9D	avg.	частота (%)		E1D	E2D	E3D	E4D	E5D	E6D	E7D	E8D	E9D	avg.	частота (%)
A	0.843	0.861	0.872	0.843	0.851	0.856	0.843	0.843	0.861	0.853	5.9	A	0.932	0.931	0.941	0.930	0.931	0.930	0.934	0.930	0.931	0.932	3.7
D	0.823	0.847	0.862	0.828	0.834	0.842	0.824	0.828	0.847	0.837	17.3	D	0.940	0.940	0.944	0.939	0.940	0.937	0.943	0.939	0.940	0.940	8.0
E	0.813	0.839	0.854	0.819	0.825	0.833	0.814	0.819	0.839	0.828	16.9	E	0.941	0.941	0.945	0.940	0.941	0.938	0.943	0.940	0.941	0.941	8.8
G	0.831	0.854	0.865	0.834	0.843	0.848	0.832	0.834	0.854	0.844	6.5	G	0.937	0.936	0.943	0.935	0.936	0.934	0.939	0.935	0.936	0.937	5.0
I	0.832	0.856	0.866	0.835	0.844	0.849	0.833	0.835	0.856	0.845	2.9	I	0.940	0.939	0.945	0.937	0.939	0.937	0.942	0.937	0.939	0.939	3.7
K	0.770	0.797	0.818	0.777	0.782	0.790	0.772	0.777	0.797	0.787	1.6	K	0.936	0.935	0.941	0.935	0.935	0.933	0.939	0.935	0.935	0.936	5.0
L	0.837	0.860	0.868	0.838	0.848	0.853	0.837	0.838	0.860	0.849	1.6	L	0.943	0.942	0.947	0.940	0.942	0.940	0.944	0.940	0.942	0.942	4.6
N	0.838	0.859	0.872	0.841	0.848	0.854	0.839	0.841	0.859	0.850	5.9	N	0.940	0.940	0.944	0.939	0.940	0.937	0.942	0.939	0.940	0.940	4.3
P	0.834	0.856	0.868	0.837	0.844	0.850	0.834	0.837	0.856	0.846	10.8	P	0.944	0.944	0.947	0.943	0.944	0.941	0.946	0.943	0.944	0.944	15.8
Q	0.824	0.847	0.862	0.828	0.835	0.841	0.825	0.828	0.847	0.837	4.9	Q	0.941	0.940	0.945	0.939	0.940	0.938	0.943	0.939	0.940	0.941	4.9
S	0.836	0.857	0.869	0.838	0.846	0.852	0.837	0.838	0.857	0.848	5.5	S	0.941	0.940	0.945	0.938	0.940	0.938	0.942	0.938	0.940	0.940	9.0
T	0.846	0.866	0.877	0.848	0.855	0.861	0.846	0.848	0.866	0.857	10.8	T	0.942	0.942	0.946	0.940	0.942	0.939	0.944	0.940	0.942	0.942	5.6
V	0.844	0.866	0.875	0.846	0.855	0.860	0.845	0.846	0.866	0.856	5.9	V	0.944	0.944	0.948	0.942	0.944	0.941	0.946	0.942	0.944	0.944	6.3

Естеств.											Естеств.												
	P1E	P2E	P3E	P4E	P5E	P6E	P7E	P8E	P9E	avg.	частота (%)		E1E	E2E	E3E	E4E	E5E	E6E	E7E	E8E	E9E	avg.	частота (%)
A	0.831	0.850	0.863	0.830	0.839	0.844	0.831	0.830	0.850	0.841	5.9	A	0.935	0.935	0.943	0.933	0.935	0.933	0.937	0.933	0.935	0.935	3.7
D	0.813	0.838	0.855	0.819	0.825	0.833	0.814	0.819	0.838	0.828	17.3	D	0.942	0.941	0.945	0.940	0.941	0.938	0.944	0.940	0.941	0.941	8.0
E	0.801	0.829	0.846	0.808	0.814	0.823	0.803	0.808	0.829	0.818	16.9	E	0.942	0.942	0.945	0.941	0.942	0.939	0.944	0.941	0.942	0.942	8.8
G	0.819	0.844	0.857	0.823	0.832	0.838	0.821	0.823	0.844	0.833	6.5	G	0.939	0.938	0.944	0.937	0.938	0.936	0.941	0.937	0.938	0.939	5.0
I	0.820	0.845	0.858	0.823	0.833	0.838	0.821	0.823	0.845	0.834	2.9	I	0.942	0.941	0.946	0.939	0.941	0.939	0.944	0.939	0.941	0.941	3.7
K	0.752	0.780	0.804	0.760	0.764	0.773	0.754	0.760	0.780	0.770	1.6	K	0.938	0.937	0.942	0.936	0.937	0.934	0.940	0.936	0.937	0.937	5.0
L	0.825	0.849	0.859	0.826	0.837	0.841	0.825	0.826	0.849	0.837	1.6	L	0.945	0.945	0.949	0.942	0.944	0.942	0.946	0.942	0.945	0.944	4.6
N	0.827	0.850	0.865	0.832	0.838	0.844	0.828	0.832	0.850	0.841	5.9	N	0.941	0.941	0.945	0.940	0.941	0.939	0.944	0.940	0.941	0.941	4.3
P	0.823	0.847	0.861	0.827	0.835	0.841	0.824	0.827	0.847	0.837	10.8	P	0.945	0.945	0.948	0.943	0.945	0.942	0.947	0.943	0.945	0.945	15.8
Q	0.812	0.836	0.853	0.817	0.823	0.830	0.813	0.817	0.836	0.826	4.9	Q	0.942	0.942	0.946	0.941	0.942	0.940	0.944	0.941	0.942	0.942	4.9
S	0.825	0.847	0.861	0.827	0.836	0.841	0.825	0.827	0.847	0.837	5.5	S	0.942	0.942	0.946	0.940	0.941	0.939	0.944	0.940	0.942	0.942	9.0
T	0.835	0.856	0.870	0.838	0.845	0.851	0.835	0.838	0.856	0.847	10.8	T	0.944	0.943	0.947	0.942	0.943	0.941	0.946	0.942	0.943	0.943	5.6
V	0.834	0.857	0.868	0.836	0.845	0.850	0.835	0.836	0.857	0.846	5.9	V	0.946	0.945	0.949	0.943	0.945	0.943	0.948	0.943	0.945	0.945	6.3

Фигура 3.

										Естеств. частота(%)	
	P1Q	P2Q	P3Q	P4Q	P5Q	P6Q	P7Q	P8Q	P9Q	avg.	
A	0.862	0.877	0.887	0.861	0.868	0.873	0.862	0.861	0.877	0.870	5.9
D	0.843	0.866	0.878	0.847	0.853	0.860	0.844	0.847	0.866	0.856	17.3
E	0.833	0.858	0.871	0.838	0.845	0.852	0.834	0.838	0.858	0.847	16.9
G	0.851	0.872	0.882	0.853	0.861	0.866	0.851	0.853	0.872	0.862	6.5
I	0.852	0.873	0.883	0.853	0.862	0.866	0.853	0.853	0.873	0.863	2.9
K	0.792	0.817	0.837	0.797	0.803	0.810	0.793	0.797	0.817	0.807	1.6
L	0.857	0.878	0.885	0.857	0.867	0.871	0.858	0.857	0.878	0.868	1.6
N	0.857	0.876	0.888	0.859	0.866	0.871	0.857	0.859	0.876	0.868	5.9
P	0.853	0.874	0.885	0.855	0.863	0.868	0.853	0.855	0.874	0.864	10.8
Q	0.843	0.864	0.877	0.846	0.853	0.858	0.844	0.846	0.864	0.855	4.9
S	0.856	0.875	0.885	0.856	0.864	0.869	0.856	0.856	0.875	0.866	5.5
T	0.864	0.882	0.892	0.866	0.873	0.877	0.864	0.866	0.882	0.874	10.8
V	0.864	0.883	0.891	0.864	0.873	0.877	0.864	0.864	0.883	0.874	5.9

										Естеств. частота(%)	
	E1Q	E2Q	E3Q	E4Q	E5Q	E6Q	E7Q	E8Q	E9Q	avg.	
A	0.938	0.938	0.945	0.935	0.938	0.936	0.940	0.935	0.938	0.938	3.7
D	0.944	0.944	0.947	0.942	0.944	0.942	0.946	0.942	0.944	0.944	8.0
E	0.945	0.945	0.948	0.943	0.944	0.942	0.947	0.943	0.945	0.945	8.8
G	0.942	0.941	0.946	0.939	0.941	0.939	0.943	0.939	0.941	0.941	5.0
I	0.945	0.945	0.948	0.942	0.945	0.942	0.947	0.942	0.945	0.945	3.7
K	0.941	0.940	0.944	0.939	0.940	0.938	0.943	0.939	0.940	0.940	5.0
L	0.948	0.947	0.951	0.944	0.947	0.945	0.949	0.944	0.947	0.947	4.6
N	0.944	0.944	0.947	0.942	0.944	0.942	0.946	0.942	0.944	0.944	4.3
P	0.948	0.947	0.950	0.946	0.947	0.945	0.949	0.946	0.947	0.947	15.8
Q	0.945	0.945	0.948	0.943	0.944	0.943	0.947	0.943	0.945	0.945	4.9
S	0.944	0.944	0.948	0.942	0.944	0.942	0.946	0.942	0.944	0.944	9.0
T	0.946	0.946	0.949	0.944	0.946	0.944	0.948	0.944	0.946	0.946	5.6
V	0.949	0.948	0.951	0.946	0.948	0.946	0.950	0.946	0.948	0.948	6.3

										Естеств. частота(%)	
	P1S	P2S	P3S	P4S	P5S	P6S	P7S	P8S	P9S	avg.	
A	0.853	0.865	0.876	0.849	0.857	0.861	0.852	0.849	0.865	0.859	5.9
D	0.853	0.873	0.883	0.854	0.862	0.867	0.853	0.854	0.873	0.864	17.3
E	0.843	0.865	0.876	0.845	0.854	0.859	0.844	0.845	0.865	0.855	16.9
G	0.857	0.876	0.884	0.856	0.866	0.869	0.857	0.856	0.876	0.866	6.5
I	0.859	0.877	0.885	0.858	0.868	0.871	0.859	0.858	0.877	0.868	2.9
K	0.789	0.812	0.829	0.791	0.800	0.805	0.790	0.791	0.812	0.802	1.6
L	0.861	0.877	0.885	0.858	0.868	0.871	0.860	0.858	0.877	0.868	1.6
N	0.864	0.881	0.890	0.863	0.871	0.876	0.863	0.863	0.881	0.872	5.9
P	0.862	0.880	0.889	0.861	0.871	0.875	0.862	0.861	0.880	0.871	10.8
Q	0.847	0.866	0.877	0.848	0.856	0.860	0.848	0.848	0.866	0.857	4.9
S	0.859	0.875	0.884	0.857	0.866	0.870	0.858	0.857	0.875	0.867	5.5
T	0.869	0.884	0.893	0.868	0.876	0.879	0.868	0.868	0.884	0.877	10.8
V	0.871	0.887	0.894	0.869	0.879	0.882	0.871	0.869	0.887	0.879	5.9

										Естеств. частота(%)	
	E1S	E2S	E3S	E4S	E5S	E6S	E7S	E8S	E9S	avg.	
A	0.907	0.904	0.937	0.920	0.905	0.902	0.910	0.920	0.904	0.912	3.7
D	0.941	0.940	0.945	0.938	0.940	0.938	0.942	0.938	0.940	0.940	8.0
E	0.942	0.941	0.946	0.939	0.941	0.939	0.943	0.939	0.941	0.941	8.8
G	0.933	0.932	0.942	0.930	0.932	0.930	0.934	0.930	0.932	0.933	5.0
I	0.938	0.938	0.945	0.934	0.938	0.936	0.940	0.934	0.938	0.938	3.7
K	0.936	0.935	0.942	0.934	0.935	0.933	0.938	0.934	0.935	0.936	5.0
L	0.940	0.939	0.947	0.936	0.939	0.937	0.941	0.936	0.939	0.939	4.6
N	0.941	0.940	0.945	0.938	0.940	0.938	0.942	0.938	0.940	0.940	4.3
P	0.945	0.945	0.949	0.943	0.945	0.943	0.947	0.943	0.945	0.945	15.8
Q	0.940	0.939	0.946	0.937	0.939	0.938	0.941	0.937	0.939	0.940	4.9
S	0.939	0.938	0.945	0.936	0.939	0.937	0.940	0.936	0.938	0.939	9.0
T	0.941	0.941	0.947	0.938	0.941	0.939	0.942	0.938	0.941	0.941	5.6
V	0.943	0.942	0.948	0.939	0.942	0.940	0.944	0.939	0.942	0.942	6.3

Фигура 3 (продолжение).

	P1	P2	P3	P4	P5	P6	P7	P8	P9		Средн.		E1	E2	E3	E4	E5	E6	E7	E8	E9		Средн.
Последовательность WT	EV	QV	QI	QV	EV	QV	QV	QL	QV		EV/QV/QI/QL		EV	QV	QI	QV	EV	QV	QV	QL	QV		EV/QV/QI/QL
Оценка WT	0.834	0.883	0.883	0.864	0.845	0.877	0.864	0.857	0.883		0.866		0.946	0.948	0.948	0.946	0.945	0.946	0.950	0.944	0.948		0.947
Последовательность наилучшего варианта	AV	AV	AT/AV	AT/AV	AV	AV	AV	AT/AV	AV		AT/AV		AP	AP	AP/AV	AP	AP	AP	AP	AP	AP		AP/AV
Оценка консенсусного наилучшего варианта	0.907	0.919	0.922	0.904	0.912	0.915	0.907	0.904	0.919		0.912		0.950	0.950	0.953	0.948	0.950	0.949	0.952	0.948	0.950		0.950
Последовательность консенсусного наилучшего варианта	AP	AP	AP	AP	AP	AP	AP	AP	AP		AP		AP	AP	AP	AP	AP	AP	AP	AP	AP		AP
Оценка наилучшего варианта	0.899	0.913	0.918	0.897	0.905	0.909	0.898	0.897	0.913		0.905		0.950	0.950	0.953	0.948	0.950	0.949	0.952	0.948	0.950		0.950
Оценка отличия от WT	0.065	0.030	0.035	0.033	0.060	0.032	0.034	0.040	0.030		0.040		0.004	0.002	0.005	0.002	0.005	0.003	0.002	0.004	0.002		0.003

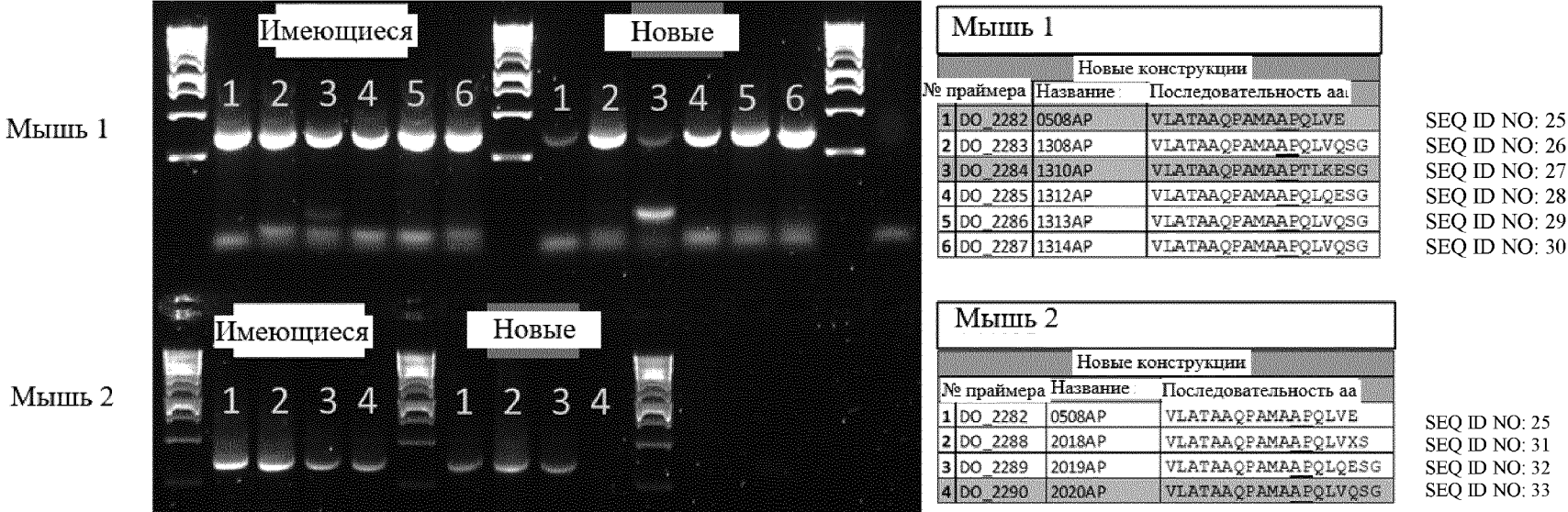
Фигура 4.

0508AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGGTGGAG-----	SEQ ID NO: 16
1308AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGGTGCAGTCTGGGGC	SEQ ID NO: 17
1310AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCC</u> ACCTTGAAGGAGTCTGGTC-	SEQ ID NO: 18
1312AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGCAGGAGTCGGGCC-	SEQ ID NO: 19
1313AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGGTGCAGTCTGGAG-	SEQ ID NO: 20
1314AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGGTGCAATCTGGGTC	SEQ ID NO: 21
2018AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGGTGSAGTCTGG---	SEQ ID NO: 22
2019AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGCAGGAGTCGGGCC-	SEQ ID NO: 23
2020AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCC</u> CAGCTGGTACAGTCTGGGGC	SEQ ID NO: 24

Фигура 5.

0508AP	VLATAAQPAMAA <u>AP</u> QLVE--	SEQ ID NO: 25
1308AP	VLATAAQPAMAA <u>AP</u> QLVQSG	SEQ ID NO: 26
1310AP	VLATAAQPAMAA <u>APT</u> LKESG	SEQ ID NO: 27
1312AP	VLATAAQPAMAA <u>AP</u> QLQESG	SEQ ID NO: 28
1313AP	VLATAAQPAMAA <u>AP</u> QLVQSG	SEQ ID NO: 29
1314AP	VLATAAQPAMAA <u>AP</u> QLVQSG	SEQ ID NO: 30
2018AP	VLATAAQPAMAA <u>AP</u> QLVXS-	SEQ ID NO: 31
2019AP	VLATAAQPAMAA <u>AP</u> QLQESG	SEQ ID NO: 32
2020AP	VLATAAQPAMAA <u>AP</u> QLVQSG	SEQ ID NO: 33

Фигура 6.



Выравнивание новых праймеров (первые два кодона VN выделены жирным шрифтом и подчеркнуты)

SEQ ID NO:

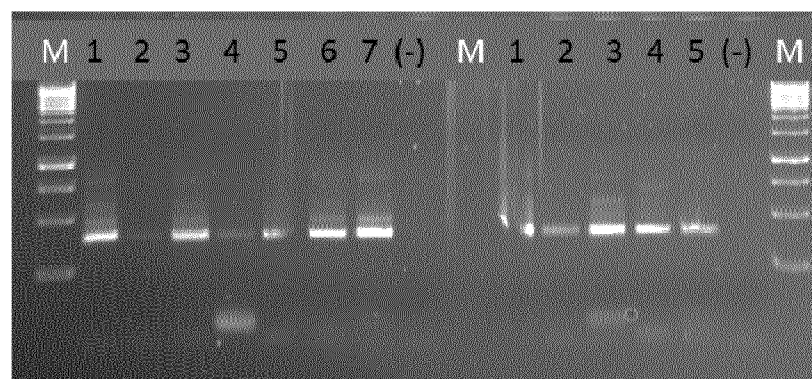
0508AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGGTGGAG-----	16
1308AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGGTGCAGTCTGGGGC	17
1310AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCC</u> ACCTTGAAGGAGTCTGGTC-	18
1312AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGCAGGAGTCGGGCC-	19
1313AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGGTGCAGTCTGGAG-	20
1314AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGGTGCATCTGGGTC	21
2018AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGGTGSAGTCTGG---	22
2019AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCG</u> CAGCTGCAGGAGTCGGGCC-	23
2020AP	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCGCCC</u> CAGCTGGTACAGTCTGGGGC	24

Выравнивание новых праймеров (первые два кодона VN выделены жирным шрифтом и подчеркнуты)

SEQ ID NO:

0508AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCACCT</u> CAGCTGGTGGAG-----	34
1308AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCACCT</u> CAGCTGGTGCAGTCTGGGGC	35
1310AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCACCT</u> ACCTTGAAGGAGTCTGGTC-	36
1312AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCACCT</u> CAGCTGCAGGAGTCGGGCC-	37
1313AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCACCT</u> CAGCTGGTGCAGTCTGGAG-	38
1314AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCACCT</u> CAGCTGGTGCATCTGGGTC	39
2018AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCACCT</u> CAGCTGGTGSAGTCTGG---	40
2019AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCACCT</u> CAGCTGCAGGAGTCGGGCC-	41
2020AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC <u>GCACCT</u> CAGCTGGTACAGTCTGGGGC	42

Фигура 8.

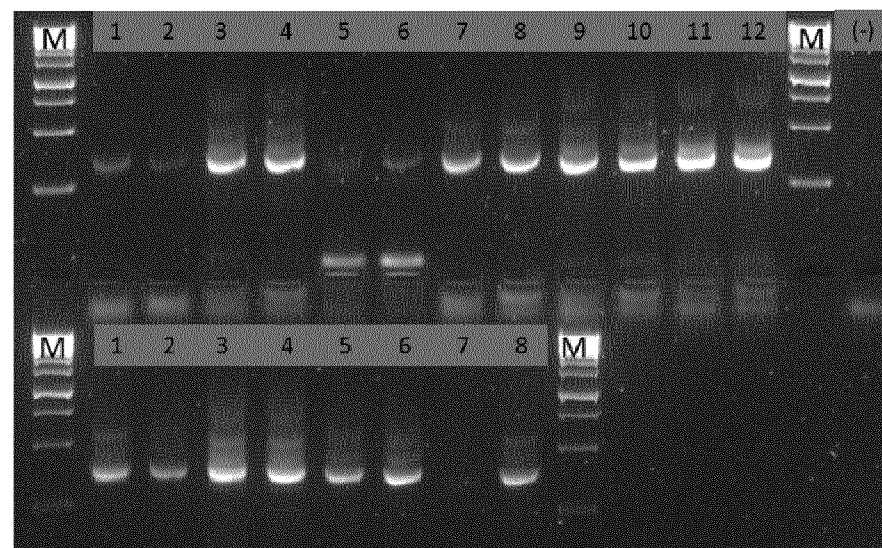


Мышь 1

Мышь 2

Мышь 1			
№№	№ праймера	Название праймера	Примечание
1	DO_0508	HuVN3A	Контроль, старый праймер
2	DO_2294	0508AP2	Субоптимален в предшествующей разработке
3	DO_2295	1308AP2	
4	DO_2296	1310AP2	Субоптимален в предшествующей разработке
5	DO_2297	1312AP2	
6	DO_2298	1313AP2	
7	DO_2299	1314AP2	
Мышь 2			
№№	№ праймера	Название праймера	Примечание
1	DO_0508	HuVN3A	Контроль, старый праймер
2	DO_2294	0508AP2	
3	DO_2300	2018AP2	
4	DO_2301	2019AP2	Субоптимален в предшествующей разработке
5	DO_2302	2020AP2	Отрицателен в предшествующей разработке

Фигура 9.



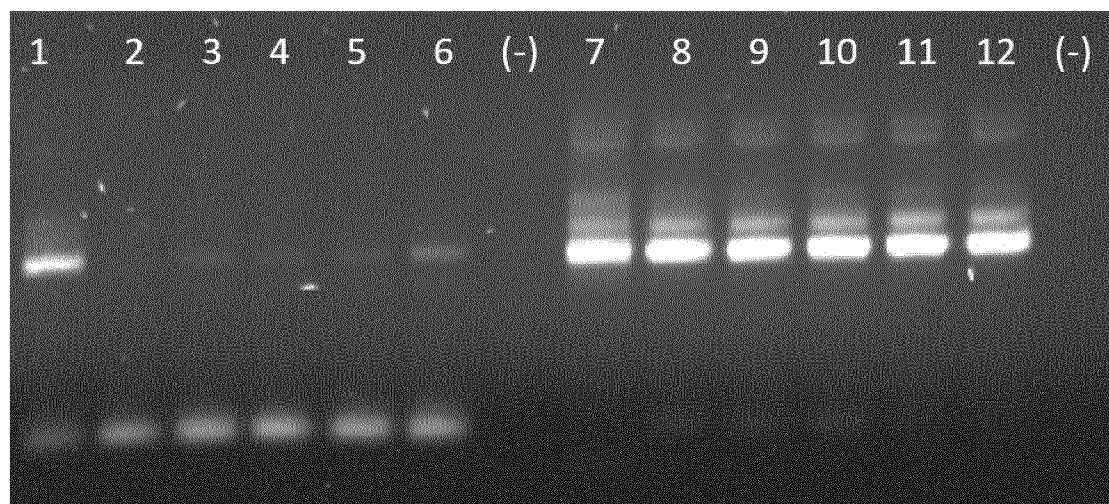
Мышь 1

Мышь 2

Мышь 1					
№	№ праймера	Название			
1	DO_2282	0508AP			
2	DO_2294	0508AP2			
3	DO_2283	1308AP			
4	DO_2295	1308AP2			
5	DO_2284	1310AP			
6	DO_2296	1310AP2			
7	DO_2285	1312AP			
8	DO_2297	1312AP2			
9	DO_2286	1313AP			
10	DO_2298	11313AP2			
11	DO_2287	1314AP			
12	DO_2299	1314AP2			

Мышь 2		
№	№ праймера	Название
1	DO_2282	0508AP
2	DO_2294	0508AP2
3	DO_2288	2018AP
4	DO_2300	2018AP2
5	DO_2289	2019AP
6	DO_2301	2019AP2
7	DO_2290	2020AP
8	DO_2302	2020AP2

Фигура 10.



Мышь 1

Мышь 2

Мышь 1				Мышь 2			
Реакция №	Праймер	DO	Кодоны 1+2	Реакция №	Праймер	DO	Кодоны 1+2
1	DO 1310	DO 1310	cag atc	7	DO 2021	DO	gag gtg
2	1310AP	DO 2284	gcg ccc	8	2021AP	DO	gcg ccg
3	1310AP2	DO 2296	gca cct	9	2021AP2	DO	gca cct
4	1310AP3	DO 2312	gca cca	10	2021AP3	DO	gca cca
5	1310AP4	DO 2313	gct cca	11	2021AP4	DO	gct cca
6	1310AP5	DO 2314	gct cct	12	2021AP5	DO	gct cct

Фигура 11.

		Сигнальный пептид	Начало гена VH	Ошибки	SEQ ID NO
1308AP2		GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCCGCACCTCAGCTGGTGCAGTCTGGGGC			35
IGHV1-2*01_X07448.1_Homo		-----	CAGGTGCAGCTGGTGCAGTCTGGGGC	6	43
IGHV1-3*01_X62109.1_Homo		-----	CAGGTCCAGCTTGTGCAGTCTGGGGC	7	44
IGHV1-8*01_M99637.1_Homo		-----	CAGGTGCAGCTGGTGCAGTCTGGGGC	6	45
IGHV1-18*01_M99641.1_Homo		-----	CAGGTTCAGCTGGTGCAGTCTGGAGC	6	46
IGHV1-24*01_M99642.1_Homo		-----	CAGGTCCAGCTGGTACAGTCTGGGGC	7	47
IGHV1-45*01_X92209.1_Homo		-----	CAGATGCAGCTGGTGCAGTCTGGGGC	6	48
IGHV1-46*01_X92343.1_Homo		-----	CAGGTGCAGCTGGTGCAGTCTGGGGC	6	49
IGHV1-58*01_M29809.1_Homo		-----	CAAATGCAGCTGGTGCAGTCTGGGCC	6	50
IGHV1-69*01_L22582.1_Homo		-----	CAGGTGCAGCTGGTGCAGTCTGGGGC	6	51
IGHV1-69-2*01_KF698734.1_Homo		-----	GAGGTCCAGCTGGTACAGTCTGGGGC	6	52
IGHV1-69D*01_KC713934.1_Homo		-----	CAGGTGCAGCTGGTGCAGTCTGGGGC	6	53
Консенсус			CAGGTGCAGCTGGTGCAGTCTGGGGC		54

Фигура 12.

		Ошибки	SEQ ID NO
2020AP2	GTCCTCGCAACTGCGGCCAGCCGGCCATGGCCGCACCTCAGCTGGTACAGTCTGGGGC		42
IGHV1-2*01_X07448.1_Homo	-----CAGGTGCAGCTGGTGCAGTCTGGGGC	7	43
IGHV1-3*01_X62109.1_Homo	-----CAGGTCCAGCTTGTGCAGTCTGGGGC	8	44
IGHV1-8*01_M99637.1_Homo	-----CAGGTGCAGCTGGTGCAGTCTGGGGC	7	45
IGHV1-18*01_M99641.1_Homo	-----CAGGTTTCAGCTGGTGCAGTCTGGAGC	7	46
IGHV1-24*01_M99642.1_Homo	-----CAGGTCCAGCTGGTACAGTCTGGGGC	6	47
IGHV1-45*01_X92209.1_Homo	-----CAGATGCAGCTGGTGCAGTCTGGGGC	7	48
IGHV1-46*01_X92343.1_Homo	-----CAGGTGCAGCTGGTGCAGTCTGGGGC	7	49
IGHV1-58*01_M29809.1_Homo	-----CAAATGCAGCTGGTGCAGTCTGGGCC	7	50
IGHV1-69*01_L22582.1_Homo	-----CAGGTGCAGCTGGTGCAGTCTGGGGC	7	51
IGHV1-69-2*01_KF698734.1_Homo	-----GAGGTCCAGCTGGTACAGTCTGGGGC	5	52
IGHV1-69D*01_KC713934.1_Homo	-----CAGGTGCAGCTGGTGCAGTCTGGGGC	7	53
Консенсус	CAGGTGCAGCTGGTGCAGTCTGGGGC		54

Фигура 13.

			Ошибки	SEQ ID NO
2018AP2	GTCCTCGCAACTGCGGCCAGCCGGCCATGGCC	GCACCTCAGCTGGTGCAGTCTGG		40
IGHV1-2*01_X07448.1_Homo	-----	CAGGTGCAGCTGGTGCAGTCTGG	6	55
IGHV1-3*01_X62109.1_Homo	-----	CAGGTCCAGCTTGTGCAGTCTGG	7	56
IGHV1-8*01_M99637.1_Homo	-----	CAGGTGCAGCTGGTGCAGTCTGG	6	57
IGHV1-18*01_M99641.1_Homo	-----	CAGGTTTCAGCTGGTGCAGTCTGG	5	58
IGHV1-24*01_M99642.1_Homo	-----	CAGGTCCAGCTGGTACAGTCTGG	7	59
IGHV1-45*01_X92209.1_Homo	-----	CAGATGCAGCTGGTGCAGTCTGG	6	60
IGHV1-46*01_X92343.1_Homo	-----	CAGGTGCAGCTGGTGCAGTCTGG	6	61
IGHV1-58*01_M29809.1_Homo	-----	CAAATGCAGCTGGTGCAGTCTGG	5	62
IGHV1-69*01_L22582.1_Homo	-----	CAGGTGCAGCTGGTGCAGTCTGG	6	63
IGHV1-69-2*01_KF698734.1_Homo	-----	GAGGTCCAGCTGGTACAGTCTGG	6	64
IGHV1-69D*01_KC713934.1_Homo	-----	CAGGTGCAGCTGGTGCAGTCTGG	6	65
Консенсус		CAGGTGCAGCTGGTGCAGTCTGG		66

Фигура 14.

	Сигнальный пептид	Начало гена VH	SEQ ID NO
1308AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC	GCACCTCAGCTGGTGCAGTCTGGGGC	35
2018AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC	GCACCTCAGCTGGTGCAGTCTGG---	40
2020AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCC	GCACCTCAGCTGGTACAGTCTGGGGC	42

Фигура 15.

			SEQ ID NO
1310AP5	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCCGCTCCTACCTTGAAGGAGTCTGGTC		67
IGHV2-5*01_X62111.1_Homo	-----CAGATCACCTTGAAGGAGTCTGGTC	6	68
IGHV2-26*01_M99648.1_Homo	-----CAGGTACCTTGAAGGAGTCTGGTC	6	69
IGHV2-70*01_L21969.1_Homo	-----CAGGTACCTTGAAGGAGTCTGGTC	7	70
IGHV2-70D*04_KC713935.1_Homo	-----CAGGTACCTTGAAGGAGTCTGGTC	6	71
Консенсус	CAGGTACCTTGAAGGAGTCTGGTC		72

Фигура 16.

		Ошибки	SEQ ID NO
0508AP	GTCCTCGCAACTGCGGCCAGCCGGCCATGGCCGCGCCGAGCTGGTGGAG		16
IGHV3-11*01_M99652.1_Homo	-----CAGGTGCAGCTGGTGGAG	4	73
IGHV3-30*01_M77325.1_Homo	-----CAGGTGCAGCTGGTGGAG	4	74
IGHV3-30-3*01_X92283.1_Homo	-----CAGGTGCAGCTGGTGGAG	4	75
IGHV3-30-5*01_AC244456_Homo	-----CAGGTGCAGCTGGTGGAG	4	76
IGHV3-33*01_AB019439_Homo	-----CAGGTGCAGCTGGTGGAG	4	77
IGHV3-NL1*01_HM855939.1_Homo	-----CAGGTGCAGCTGGTGGAG	4	78
IGHV3-13*01_X92217.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	79
IGHV3-15*01_X92216.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	80
IGHV3-20*01_M99657.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	81
IGHV3-21*01_Z14073.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	82
IGHV3-23*01_M99660.1_Homo	-----GAGGTGCAGCTGTTGGAG	4	83
IGHV3-23D*01_AC244492_Homo	-----GAGGTGCAGCTGTTGGAG	4	84
IGHV3-43*01_M99672.1_Homo	-----GAAGTGCAGCTGGTGGAG	4	85
IGHV3-43D*03_KC713950.1_Homo	-----GAAGTGCAGCTGGTGGAG	4	86
IGHV3-9*01_M99651.1_Homo	-----GAAGTGCAGCTGGTGGAG	4	87
IGHV3-48*01_M99675.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	88
IGHV3-49*01_M99676.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	89
IGHV3-53*01_M99679.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	90
IGHV3-64*01_M99682.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	91
IGHV3-64D*06_KC713941.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	92
IGHV3-66*01_X92218.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	93
IGHV3-7*01_M99649.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	94
IGHV3-72*01_X92206.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	95
IGHV3-73*01_X70197.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	96
IGHV3-74*01_L33851.1_Homo	-----GAGGTGCAGCTGGTGGAG	3	97
Консенсус	GAGGTGCAGCTGGTGGAG		98

Фигура 17.

		Ошибки	SEQ ID NO
2021AP2	GTCCTCGCAACTGCGGCCAGCCGGCCATGGCCGCACCTCAGCTGTTGGAGTCTGGGG		99
IGHV3-7*01_M99649.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGGGG	6	100
IGHV3-9*01_M99651.1_Homo	-----GAAGTGCAGCTGGTGGAGTCTGGGG	5	101
IGHV3-11*01_M99652.1_Homo	-----CAGGTGCAGCTGGTGGAGTCTGGGG	7	102
IGHV3-13*01_X92217.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGGGG	6	103
IGHV3-15*01_X92216.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGGGG	6	104
IGHV3-20*01_M99657.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGGGG	6	105
IGHV3-21*01_Z14073.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGGGG	6	106
IGHV3-23*01_M99660.1_Homo	-----GAGGTGCAGCTGTTGGAGTCTGGGG	5	107
IGHV3-23D*01_AC244492_Homo	-----GAGGTGCAGCTGTTGGAGTCTGGGG	5	108
IGHV3-30*01_M77325.1_Homo	-----CAGGTGCAGCTGGTGGAGTCTGGGG	7	109
IGHV3-30-3*01_X92283.1_Homo	-----CAGGTGCAGCTGGTGGAGTCTGGGG	7	110
IGHV3-30-5*01_AC244456_Homo	-----CAGGTGCAGCTGGTGGAGTCTGGGG	7	111
IGHV3-33*01_AB019439_Homo	-----CAGGTGCAGCTGGTGGAGTCTGGGG	7	112
IGHV3-43*01_M99672.1_Homo	-----GAAGTGCAGCTGGTGGAGTCTGGGG	5	113
IGHV3-43D*03_KC713950.1_Homo	-----GAAGTGCAGCTGGTGGAGTCTGGGG	5	114
IGHV3-48*01_M99675.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGGGG	6	115
IGHV3-49*01_M99676.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGGGG	6	116
IGHV3-53*01_M99679.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGGAG	6	117
IGHV3-64*01_M99682.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGGGG	6	118
IGHV3-64D*06_KC713941.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGGGG	6	119
IGHV3-66*01_X92218.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGGGG	6	120
IGHV3-72*01_X92206.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGGGG	6	121
IGHV3-73*01_X70197.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGGGG	6	122
IGHV3-74*01_L33851.1_Homo	-----GAGGTGCAGCTGGTGGAGTCCGGGG	7	123
IGHV3-NL1*01_HM855939.1_Homo	-----CAGGTGCAGCTGGTGGAGTCTGGGG	7	124
Консенсус	GAGGTGCAGCTGGTGGAGTCTGGGG		125

Фигура 18.

		Ошибки	SEQ ID NO
2018AP2	GTCCTCGCAACTGCGGCCAGCCGGCCATGGCCGCACCTCAGCTGGTGGAGTCTGG		40
IGHV3-7*01_M99649.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGG	5	126
IGHV3-9*01_M99651.1_Homo	-----GAAGTGCAGCTGGTGGAGTCTGG	4	127
IGHV3-11*01_M99652.1_Homo	-----CAGGTGCAGCTGGTGGAGTCTGG	6	128
IGHV3-13*01_X92217.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGG	5	129
IGHV3-15*01_X92216.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGG	5	130
IGHV3-20*01_M99657.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGG	5	131
IGHV3-21*01_Z14073.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGG	5	132
IGHV3-23*01_M99660.1_Homo	-----GAGGTGCAGCTGTTGGAGTCTGG	6	133
IGHV3-23D*01_AC244492_Homo	-----GAGGTGCAGCTGTTGGAGTCTGG	6	134
IGHV3-30*01_M77325.1_Homo	-----CAGGTGCAGCTGGTGGAGTCTGG	6	135
IGHV3-30-3*01_X92283.1_Homo	-----CAGGTGCAGCTGGTGGAGTCTGG	6	136
IGHV3-30-5*01_AC244456_Homo	-----CAGGTGCAGCTGGTGGAGTCTGG	6	137
IGHV3-33*01_AB019439_Homo	-----CAGGTGCAGCTGGTGGAGTCTGG	6	138
IGHV3-43*01_M99672.1_Homo	-----GAAGTGCAGCTGGTGGAGTCTGG	4	139
IGHV3-43D*03_KC713950.1_Homo	-----GAAGTGCAGCTGGTGGAGTCTGG	4	140
IGHV3-48*01_M99675.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGG	5	141
IGHV3-49*01_M99676.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGG	5	142
IGHV3-53*01_M99679.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGG	5	143
IGHV3-64*01_M99682.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGG	5	144
IGHV3-64D*06_KC713941.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGG	5	145
IGHV3-66*01_X92218.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGG	5	146
IGHV3-72*01_X92206.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGG	5	147
IGHV3-73*01_X70197.1_Homo	-----GAGGTGCAGCTGGTGGAGTCTGG	5	148
IGHV3-74*01_L33851.1_Homo	-----GAGGTGCAGCTGGTGGAGTCCGG	6	149
IGHV3-NL1*01_HM855939.1_Homo	-----CAGGTGCAGCTGGTGGAGTCTGG	6	150
Консенсус	GAGGTGCAGCTGGTGGAGTCTGG		151

Фигура 19.

	Сигнальный пептид	Начало гена VH	SEQ ID NO
0508AP	GTCCTCGCAACTGCGGGCCAGCCGGCCATGGCC	GCGCCGCAGCTGGTGGAG-----	16
2018AP2	GTCCTCGCAACTGCGGGCCAGCCGGCCATGGCC	GCACCTCAGCTGGTGGAGTCTGG--	40
2021AP2	GTCCTCGCAACTGCGGGCCAGCCGGCCATGGCC	GCACCTCAGCTGTTGGAGTCTGGGG	99

Фигура 20.

		Ошибки	SEQ ID NO
1312AP2	GTCCTCGCAACTGCGGGCCAGCCGGCCATGGCCGCACCTCAGCTGCAGGAGTCGGGCC		37
IGHV4-4*01_X05713.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	152
IGHV4-28*01_X05714.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	153
IGHV4-30-2*01_L10089.1_Homo	-----CAGCTGCAGCTGCAGGAGTCCGGCT	7	154
IGHV4-30-4*01_Z14238.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	155
IGHV4-31*01_L10098.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	156
IGHV4-34*01_AB019439_Homo	-----CAGGTGCAGCTACAGCAGTGGGGCG	10	157
IGHV4-38-2*01_Z12367.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	158
IGHV4-39*01_L10094.1_Homo	-----CAGCTGCAGCTGCAGGAGTCGGGCC	5	159
IGHV4-59*01_L10088.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	160
IGHV4-61*01_M29811.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	161
Консенсус	CAGGTGCAGCTGCAGGAGTCGGGCC		162

Фигура 21.

		Ошибки	SEQ ID NO
2019AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCCGCACCTCAGCTGCAGGAGTCGGGCC		41
IGHV4-4*01_X05713.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	152
IGHV4-28*01_X05714.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	153
IGHV4-30-2*01_L10089.1_Homo	-----CAGCTGCAGCTGCAGGAGTCCGGCT	7	154
IGHV4-30-4*01_Z14238.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	155
IGHV4-31*01_L10098.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	156
IGHV4-34*01_AB019439_Homo	-----CAGGTGCAGCTACAGCAGTGGGGCG	10	157
IGHV4-38-2*01_Z12367.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	158
IGHV4-39*01_L10094.1_Homo	-----CAGCTGCAGCTGCAGGAGTCGGGCC	5	159
IGHV4-59*01_L10088.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	160
IGHV4-61*01_M29811.1_Homo	-----CAGGTGCAGCTGCAGGAGTCGGGCC	6	161
Консенсус	CAGGTGCAGCTGCAGGAGTCGGGCC		162

Фигура 22.

		Ошибки	SEQ ID NO
1313AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCCGCACCTCAGCTGGTGCAGTCTGGAG		38
IGHV5-10-1*01_X92227.1_Homo	-----GAAGTGCAGCTGGTGCAGTCTGGAG	4	163
IGHV5-51*01_M99686.1_Homo	-----GAGGTGCAGCTGGTGCAGTCTGGAG	5	164
Консенсус	GAAGTGCAGCTGGTGCAGTCTGGAG		165

Фигура 23.

		Ошибки	SEQ	ID NO
IGHV6-1*01_J04097_Homo	-----CAGGTACAGCTGCAGCAGTCAGGTCC			166
0508AP	GTCTCTGCAACTGCGGGCCAGCCGGCCATGGCCGCGCCGAGCTGGTGGAG-----	8		16
1310AP5	GTCTCTGCAACTGCGGGCCAGCCGGCCATGGCCGCTCCTACCTTGAAGGAGTCTGGTC-	13		67
1312AP2	GTCTCTGCAACTGCGGGCCAGCCGGCCATGGCCGCACCTCAGCTGCAGGAGTCGGGCC-	9		37
2019AP2	GTCTCTGCAACTGCGGGCCAGCCGGCCATGGCCGCACCTCAGCTGCAGGAGTCGGGCC-	9		41
2018AP2	GTCTCTGCAACTGCGGGCCAGCCGGCCATGGCCGCACCTCAGCTGGTGGAGTCTGG---	10		40
2021AP2	GTCTCTGCAACTGCGGGCCAGCCGGCCATGGCCGCACCTCAGCTGTTGGAGTCTGGGG-	12		99
1313AP2	GTCTCTGCAACTGCGGGCCAGCCGGCCATGGCCGCACCTCAGCTGGTGCAGTCTGGAG-	11		38
1314AP2	GTCTCTGCAACTGCGGGCCAGCCGGCCATGGCCGCACCTCAGCTGGTGCAGTCTGGGTC	12		39
1308AP2	GTCTCTGCAACTGCGGGCCAGCCGGCCATGGCCGCACCTCAGCTGGTGCAGTCTGGGGC	11		35
2020AP2	GTCTCTGCAACTGCGGGCCAGCCGGCCATGGCCGCACCTCAGCTGGTACAGTCTGGGGC	12		42

Фигура 24.

		Mismatches	SEQ ID NO
1314AP2	GTCCTCGCAACTGCGGCCCAGCCGGCCATGGCCGCACCTCAGCTGGTGCAATCTGGGTC		39
IGHV7-4-1*01_L10057.1_Homo	-----CAGGTGCAGCTGGTGCAATCTGGGTC	6	167
Consensus	CA TCAGCTGGTGCAATCTGGGTC		168

FIG. 25