

(19)



**Евразийское
патентное
ведомство**

(21) **202492312** (13) **A1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ**

(43) Дата публикации заявки
2024.11.06

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

(22) Дата подачи заявки
2023.03.07

(54) **СИСТЕМА МОДУЛЬНОГО ВЕКТОРА (MODVEC): ПЛАТФОРМА ДЛЯ СОЗДАНИЯ ВЕКТОРОВ ЭКСПРЕССИИ НОВОГО ПОКОЛЕНИЯ**

(31) 63/317,949

(32) 2022.03.08

(33) US

(86) PCT/US2023/063848

(87) WO 2023/172903 2023.09.14

(71) Заявитель:

ЭМДЖЕН ИНК. (US)

(72) Изобретатель:

Чэнь Фуи, Мок Марисса, Белуски
Эдвард Дж. (US)

(74) Представитель:

Медведев В.Н. (RU)

(57) Настоящее изобретение относится к способу оптимизации уровней экспрессии многоцепочечного белка, где белок содержит по меньшей мере две различные полипептидные цепи. Получение многоцепочечных белков ставит сложные задачи, связанные со спариванием/фолдингом новых четвертичных структур, состоящих из нескольких полипептидных цепей, особенно при спаривании тяжелых и легких цепей антител в полиспецифическом формате. Особое значение для эффективной и правильной сборки молекул с множеством полипептидных цепей имеет экспрессия различных цепей в правильном соотношении внутри клетки. Настоящее изобретение направлено на решение данных проблем.

202492312

A1

A1

202492312

ОПИСАНИЕ ИЗОБРЕТЕНИЯ

2420-582022EA/042

СИСТЕМА МОДУЛЬНОГО ВЕКТОРА (MODVEC): ПЛАТФОРМА ДЛЯ СОЗДАНИЯ ВЕКТОРОВ ЭКСПРЕССИИ НОВОГО ПОКОЛЕНИЯ ССЫЛКА НА РОДСТВЕННУЮ ЗАЯВКУ

[1] Для настоящей заявки испрашивается приоритет по предварительной заявке на патент США № 63/317949, поданной 8 марта 2022 г., полное содержание которой настоящим включено посредством ссылки в данный документ.

ВКЛЮЧЕНИЕ ПОСРЕДСТВОМ ССЫЛКИ

[2] Включенный посредством ссылки во всей своей полноте перечень последовательностей в формате XML, идентифицируемый по следующим свойствам файла: имя файла: A-2671-WO01-SEC_Sequence_Listing.XML; размер файла: 8686 байт; создан 6 марта 2023 г.

ОБЛАСТЬ ТЕХНИКИ, К КОТОРОЙ ОТНОСИТСЯ ИЗОБРЕТЕНИЕ

[1] Настоящее изобретение относится к способу получения стандартизированных кассет экспрессии. Настоящее изобретение также относится к способу рекомбинации таких стандартизированных кассет экспрессии *in vivo* в клетке-хозяине.

ПРЕДПОСЫЛКИ ИЗОБРЕТЕНИЯ

[2] Полиспецифические антитела и антителоподобные конструкции обладают рядом характеристик, которые привлекательны для разработчиков терапевтических молекул. Клинический потенциал полиспецифических антител, которые нацелены на несколько целей одновременно, таких как биспецифические и триспецифические антитела, демонстрирует большие перспективы для нацеливания на сложные заболевания. Однако создание таких молекул сопряжено с большими трудностями по причине спаривания/фолдинга новых четвертичных структур, состоящих из нескольких полипептидных цепей, при трансфекции в одну клетку, особенно при спаривании тяжелых и легких цепей антител. Особое значение для эффективной и правильной сборки молекул с множеством полипептидных цепей имеет экспрессия различных цепей в правильном соотношении внутри клетки.

[3] Настоящее изобретение решает данную проблему с помощью системы модульных векторов (ModVec) для быстрого создания разнообразных векторов экспрессии и библиотек векторов для получения рекомбинантных белков. ModVec обеспечивает быстрое и высокоэффективное создание как простых, так и сложных векторных конструкций, позволяя оптимизировать векторы экспрессии для отдельных молекул, для конкретных крупномолекулярных соединений и/или для различных хозяев для экспрессии.

[4] Одной из основных проблем при получении многоцепочечных молекул является разное качество продуктов, обусловленное несбалансированным соотношением экспрессии цепей, и данное соотношение является специфическим для молекулы, а также специфическим для процесса развития клетки-хозяина, в том числе элементов вектора

экспрессии. В настоящее время не существует подхода, позволяющего определить оптимальное специфическое для молекулы соотношение экспрессии цепей с высокой производительностью. Библиотека векторов ModVec предоставляет платформу для быстрого определения наилучшей конфигурации экспрессии для отдельной многоцепочечной молекулы. ModVec можно применять в более общих задачах для сборки сложных больших молекул ДНК для целей, отличных от получения векторов экспрессии для терапевтических средств на основе антител.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

[5] В одном аспекте настоящее изобретение направлено на способ экспрессии многоцепочечного белка, содержащего по меньшей мере две различные полипептидные цепи, включающий

[6] а. обеспечение наличия полинуклеотидных последовательностей, которые кодируют по меньшей мере два различных полипептида, где указанные полинуклеотидные последовательности фланкированы с 5'- и 3'-сторон сайтом расщепления эндонуклеазой рестрикции IIS типа, за которым следует сайт ее распознавания;

[7] б. обеспечение наличия двух или более наборов последовательностей элементов, где каждый набор последовательностей элементов вместе содержит по меньшей мере одну функциональную кассету экспрессии,

[8] где каждая последовательность элемента фланкирована с 5'- и 3'-сторон сайтом расщепления эндонуклеазой рестрикции IIS типа, за которым следует сайт ее распознавания,

[9] где сайты распознавания эндонуклеазой рестрикции IIS типа и сайты расщепления выбраны таким образом, чтобы наборы последовательностей элементов могли быть собраны в функциональную кассету экспрессии;

[10] с. обеспечение наличия по меньшей мере одного исходного вектора-остова, где исходный вектор-остов предусматривает плазмиду, содержащую по меньшей мере один сайт расщепления эндонуклеазой рестрикции IIS типа, за которым следует сайт ее распознавания и последовательности для поддержания вектора в бактериальных клетках;

[11] d. сборку полинуклеотидных последовательностей, которые кодируют по меньшей мере два различных полипептида, с двумя или более наборами последовательностей элементов для получения функциональных кассет экспрессии в исходном векторе-остове с помощью способа, основанного на применении расщепления ферментом рестрикции IIS типа и лигирования посредством сайтов расщепления и "липких" концов, в результате чего получают лигированные векторы, содержащие по меньшей мере две функциональные кассеты экспрессии, которые способны экспрессировать по меньшей мере две различные полипептидные цепи;

[12] e. трансфекцию клеток млекопитающих лигированными векторами;

[13] f. сравнение уровней полученного предпочтительного продукта и уровней примесей для каждого из по меньшей мере двух полипептидов и

[14] g. выявление оптимального соотношения экспрессии двух полипептидов для трансфицированных клеток млекопитающих и выявление трансфицированных клеток млекопитающих, которые характеризуются оптимальным соотношением экспрессии по меньшей мере двух полипептидных цепей.

[15] В одном варианте осуществления способ дополнительно включает

[16] h. выявление расположения элементов вектора, которое обеспечивает оптимальное соотношение экспрессии по меньшей мере двух полипептидных цепей.

[17] В одном аспекте настоящее изобретение направлено на вектор, характеризующийся расположением элементов вектора, которое обеспечивает оптимальное соотношение экспрессии по меньшей мере двух полипептидных цепей.

[18] В одном варианте осуществления последовательности элементов содержат по меньшей мере две промоторные последовательности и по меньшей мере две последовательности поли-А.

[19] В одном варианте осуществления уровни полученного предпочтительного продукта измеряют с помощью способа, выбранного из группы, состоящей из катионообменной хроматографии (в восстанавливающих или невосстанавливающих условиях), масс-спектрометрии, любого другого хроматографического разделения или их комбинации.

[20] В одном варианте осуществления ни один из сайтов расщепления эндонуклеазой рестрикции IIS типа при расщеплении эндонуклеазой рестрикции IIS типа не образует 5'-"липкие" концы из четырех нуклеотидов, выбранные из группы, состоящей из GTAA, TCCA и CACA.

[21] В одном варианте осуществления сайт расщепления эндонуклеазой рестрикции IIS типа выбран из группы, состоящей из

[22] AGGT, AGTA, ATCA, CAGT, CCAT, GAAT, GAGG, GGCA, GGTC, TAGC, TCTT, GGAG и CCAC.

[23] В одном варианте осуществления клетки млекопитающих выбраны из группы, состоящей из клеток CHO, клеток CHOK1, DXB-11, DG-44, COS-7, HEK293, BHK, TM4, CV1, VERO-76, HELA, MDCK, BRL 3A, W138, Нер G2, MMT 060562, клеток TRI, клеток MRC 5, клеток FS4 и клеток миеломы млекопитающих.

[24] В одном варианте осуществления оптимальное соотношение экспрессии по меньшей мере двух полипептидов выбрано из группы, состоящей из 1:1, 1:2 и 1:3.

[25] В одном варианте осуществления многоцепочечный белок содержит первую тяжелую цепь антитела, первую легкую цепь антитела, вторую тяжелую цепь антитела и вторую легкую цепь антитела, где первая тяжелая цепь антитела соединяется с первой легкой цепью антитела для связывания первого антигена или эпитопа, а вторая тяжелая цепь антитела соединяется со второй легкой цепью антитела для связывания второго антигена или эпитопа, где оптимальное соотношение экспрессии первой тяжелой цепи антитела, первой легкой цепи антитела, второй тяжелой цепи антитела и второй легкой цепи антитела составляет 1:1:1:1.

[26] В одном варианте осуществления многоцепочечный белок содержит первую тяжелую цепь антитела, вторую тяжелую цепь антитела и общую легкую цепь антитела, где первая тяжелая цепь антитела соединяется с общей легкой цепью антитела для связывания первого антигена или эпитопа, а вторая тяжелая цепь антитела соединяется с общей легкой цепью антитела для связывания второго антигена или эпитопа, где оптимальное соотношение экспрессии первой тяжелой цепи антитела, второй тяжелой цепи антитела и общей легкой цепи антитела составляет 1:1:2.

[27] В одном варианте осуществления многоцепочечный белок содержит тяжелую цепь антитела, первую легкую цепь антитела, модифицированную тяжелую цепь антитела и вторую легкую цепь антитела,

[28] где модифицированная тяжелая цепь антитела содержит от N-конца к C-концу одну из нижеуказанных структур, выбранных из следующей группы:

[29] VH-CH1-связывающий домен (BD)-шарнир-CH2-CH3,

[30] BD-VH-CH1-шарнир-CH2-CH3 и

[31] VH-CH1-шарнир-CH2-CH3-BD,

[32] где BD выбран либо как одноцепочечный Fv (scFv), либо как одноцепочечный Fab (scFab);

[33] где тяжелая цепь антитела соединяется с первой легкой цепью антитела для связывания первого антигена или эпитопа, а VH модифицированной тяжелой цепи антитела соединяется со второй легкой цепью антитела для связывания второго антигена или эпитопа, где BD связывается с третьим антигеном или эпитопом, и

[34] где оптимальное соотношение экспрессии первой тяжелой цепи антитела, первой легкой цепи, модифицированной тяжелой цепи антитела и второй легкой цепи антитела составляет 1:1:1:1.

[35] В одном варианте осуществления многоцепочечный белок содержит тяжелую цепь антитела, модифицированную тяжелую цепь антитела и общую легкую цепь антитела,

[36] где модифицированная тяжелая цепь антитела содержит от N-конца к C-концу одну из нижеуказанных структур, выбранных из следующей группы:

[37] VH-CH1-связывающий домен (BD)-шарнир-CH2-CH3,

[38] BD-VH-CH1-шарнир-CH2-CH3 и

[39] VH-CH1-шарнир-CH2-CH3-BD,

[40] где BD выбран либо как одноцепочечный Fv (scFv), либо как одноцепочечный Fab (scFab);

[41] где тяжелая цепь антитела соединяется с общей легкой цепью антитела для связывания первого антигена или эпитопа, а VH модифицированной тяжелой цепи антитела соединяется с общей легкой цепью антитела для связывания первого антигена или эпитопа, где BD связывается со вторым антигеном или эпитопом, и

[42] где оптимальное соотношение экспрессии первой тяжелой цепи антитела, модифицированной тяжелой цепи антитела и общей легкой цепи антитела составляет

1:1:2.

КРАТКОЕ ОПИСАНИЕ ГРАФИЧЕСКИХ МАТЕРИАЛОВ

[43] На фигурах 1A-1E изображена платформа модульных векторов для высокоэффективного конструирования векторов ("ModVec"). Фигура 1A) Пример последовательности модуля ДНК. Каждый модуль (изображен буквами N и серым цветом) фланкирован адаптерами Golden Gate (область в рамке), которые содержат сайт распознавания для *BsmBI* (выделен жирным шрифтом). Разрезаемые посредством *BsmBI* сайты обозначены знаком ^, а "липкие" концы, полученные в результате расщепления посредством *BsmI*, обозначены подчеркиванием. Фигура 1B) Сборку вектора экспрессии осуществляют посредством реакции Golden Gate с применением *BsmBI* с элементами вектора, фланкированными заданными комплементарными "липкими" концами, а также остовом вектора экспрессии. Фигура 1C) Схема сборки ModVec размером 11 т. о., состоящей из 11 частей (включая остов вектора). ITR: инвертированные концевые повторы транспозона *piggyBac*. CDS: кодирующая последовательность. Фигура 1D) Все 19 конструкций были успешно собраны с помощью ModVec за один раунд клонирования. 96% колоний (109 из 114) характеризовались правильно собранными конструкциями. Фигура 1E) Схема сборки библиотеки векторов в одной пробирке. Фигура 1F) Эффективное создание библиотеки векторов в одной пробирке с помощью ModVec; при оценке 350 колоний были получены 69 из 72 возможных конструкций.

[44] На фигурах 2A-2D изображено, что выбор промотора оказывает значимое влияние на возможность получения. Фигура 2A) Схема конфигураций векторов. Как ген LC, так и ген HC находились под контролем одного и того же промотора: либо промотора 1, либо промотора 2. Влияние промотора на возможность получения стандартных mAb A, bsAb B и bsAb C показано на фигуре 2B, фигуре 2C и фигуре 2D соответственно.

[45] На фигурах 3A-3H изображено повышение выработки четырехцепочечного гетеро-IgG с помощью скрининга библиотеки конфигураций векторов. Фигура 3A) Схемы конфигураций векторов для трех масштабируемых пулов. Фигура 3B) Соотношение LC1:LC2 очищенных с помощью ProA образцов из пулов 905, 910 и 911 перед масштабированием для дополнительного анализа. Фигура 3C) Выход при очистке с ProA для масштабированных пулов 905, 910 и 911. Фигура 3D) Степень извлечения после катионообменной очистки для трех пулов. Фигура 3E) Окончательный выход правильно собранного гетеро-IgG-D после катионного обмена. На фигуре 3F, фигуре 3G и фигуре 3H соответственно показаны катионообменные хроматограммы для очищенных с помощью ProA образцов из пулов 905, 910 и 911. Идентичность каждого пика определяли с помощью масс-спектрометрии.

[46] На фигурах 4A-4H изображено повышение выработки асимметричной триспецифической молекулы с помощью скрининга библиотеки конфигураций векторов. Фигура 4A) Схемы структуры библиотеки конфигураций векторов. Были протестированы три различные группы конфигураций: одинарный LC, двойной LC A и двойной LC B. В каждой конфигурации для контроля транскрипции полипептидных цепей применяли три

различных промотора: промотор 1, промотор 2 и промотор 3. Фигура 4B) Выход при очистке с ProA для всех конфигураций векторов. Средний выход конфигураций векторов при очистке с ProA в группах с двойным LC A, двойным LC B и одинарным LC составлял $117,2 \pm 43,4$ мг/л, $112,6 \pm 39,2$ мг/л и $40,7 \pm 20,73$ мг/л соответственно. Значения выражены как среднее значение \pm стандартное отклонение. Фигура 4C) Качество продуктов в отношении очищенных с помощью ProA образцов от всех конфигураций векторов. Средний % MP по mMSE конфигураций векторов при двойном LC A, двойном LC B, одинарном LC составлял $75,6 \pm 14,9\%$, $72,4 \pm 16,6\%$ и $71,7 \pm 21,6\%$ соответственно. Фигура 4D) Окончательный выход правильно собранных tsAb E всех конфигураций векторов. Средний окончательный выход конфигураций векторов в группах с двойным LC A, двойным LC B и одинарным LC составлял $88,7 \pm 33,6$ мг/л, $81,4 \pm 35,6$ мг/л и $28,5 \pm 16,8$ мг/л соответственно. Фигура 4E) Выход при очистке с ProA из стабильных пулов, полученных из стандартных и оптимальных конфигураций с помощью стандартного протокола разработки линии клеток. Выход при очистке с ProA из оптимального пула характеризовался увеличением на 124% по сравнению с выходом при очистке с ProA из стандартного пула. Фигура 4F) Качество продуктов в отношении полученных с помощью ProA образцов из стандартного и оптимального пула было сопоставимым. Фигура 4G) Окончательный выход tsAb E составлял 114 мг/л из стандартного пула и 257 мг/л из оптимального пула. Оптимальный пул характеризовался увеличением на 125% окончательного выхода по сравнению со стандартным пулом. Фигура 4H) Схемы оптимальной конфигурации вектора, выявленной путем скрининга библиотеки, и конфигурации вектора платформы авторов изобретения.

[47] На фигурах 5A-5D изображен скрининг библиотеки конфигураций векторов для трудноэкспрессируемого гетеро-IgG. Фигура 5A) Схемы применяемых конфигураций векторов. Экспрессия гетеро-IgG D предусматривала совместную трансфекцию двух бицистронных векторов, содержащих гены всех четырех полипептидных цепей. Векторы конфигурации 1 содержат гены LC и HC из одного и того же mAb в одном и том же бицистронном векторе. Векторы конфигурации 2 содержат гены LC и HC в отдельном бицистронном векторе. Фигура 5B) Выход при очистке с ProA для всех восстановленных стабильных пулов клеток CHO; восстановлено 29 из 44 пулов. Фигура 5C) Соотношения LC1:LC2 очищенных с помощью ProA образцов из всех пулов. Пулы, отмеченные звездочками, были масштабированы для дополнительного анализа. Фигура 5D) Качество продуктов всех пулов, измеренное по основному пику MSE в невозстанавливающих условиях и основному пику SEC.

[48] На фигурах 6A-6G изображено, что положение гена в векторе экспрессии может оказывать значимое влияние на возможность получения. Фигура 6A) Схема конфигурации DA2 и DB2. В обеих конфигурациях для управления экспрессией всех полипептидных цепей применяли промотор 2. Фигура 6B) Пул DA2 характеризовался более высоким выходом при очистке с ProA, чем пул DB2. Фигура 6C) Качество продуктов в отношении очищенных с помощью ProA образцов из пулов DA2 и DB2 было

сопоставимым. Фигура 6D) Пул DA2 характеризовался более высоким окончательным выходом tsAb E, чем пул DB2. Фигура 6E) Электрофореграммы (микрокапиллярный электрофорез в невосстанавливающих условиях) очищенных с помощью ProA образцов из пула DA2. Относительная площадь пика для половины mAb1 составляла 10,1, а для половины mAb2 составляла 7,4 для очищенных с помощью ProA образцов. Фигура 6F) Электрофореграммы (микрокапиллярный электрофорез в невосстанавливающих условиях) очищенных с помощью ProA образцов из пула DB2. Относительная площадь пика для половины mAb1 составляла 5,8, а для половины mAb2 составляла 9,8. Фигура 6G) Относительное соотношение пиков половины mAb1:половины mAb2 для очищенных с помощью ProA образцов из пулов DA2 составляло 1,4, тогда как данное значение для очищенных с помощью ProA образцов для пула DB2 составляло 0,6.

ПОДРОБНОЕ ОПИСАНИЕ ИЗОБРЕТЕНИЯ

[49] Способ по настоящему изобретению позволяет получать представляющие интерес кассеты экспрессии из наборов последовательностей элементов путем сборки конструкций из фрагментов нуклеиновых кислот посредством односторонних "липких" концов, образованных на обоих концах фрагментов с применением эндонуклеаз рестрикции II типа. В настоящем изобретении можно применять ферменты рестрикции II типа. Сайт распознавания эндонуклеазой рестрикции II типа представляет собой сайт распознавания эндонуклеазой рестрикции, распознающей двунитевую ДНК и расщепляющей двунитевую ДНК в сайте расщепления, который находится за пределами сайта распознавания на двунитевой ДНК. Эндонуклеаза рестрикции II типа осуществляет расщепление таким образом, что в зависимости от конкретной эндонуклеазы рестрикции II типа получают "липкие" концы длиной от 3 до 6 нуклеотидов. Как правило, в способе по настоящему изобретению можно применять ферменты, дающие начало 4-нуклеотидным "липким" концам. Однако также можно применять эндонуклеазы II типа, образующие более длинные односторонние "липкие" концы. Нуклеотидная область, образующая "липкие" концы при расщеплении, в данном документе называется сайтом расщепления. Поскольку нуклеотиды сайта расщепления не являются частью сайта распознавания, их можно выбирать по желанию, не нарушая при этом расщепляющую активность эндонуклеазы рестрикции II типа.

[50] Для реализации на практике настоящего изобретения можно применять любой фермент рестрикции II типа, который обеспечивает "липкие" концы, достаточные для эффективного лигирования в сайтах его расщепления. Подборка таких ферментов представлена на веб-странице REBASE (rebase.neb.com/cgi-bin/asymmlist) и в обзоре Szybalsky et al. (1991, Gene, 100:13-26). Наиболее предпочтительными являются следующие эндонуклеазы рестрикции II типа: Bsal, BbsI, BsmBI, SapI, BspMI, AarI, Esp3I, BpiI и HgaI. Многие из упомянутых эндонуклеаз рестрикции доступны от New England Biolabs. Источники данных ферментов также можно найти на упомянутой выше веб-странице REBASE. В настоящем изобретении можно применять ферменты рестрикции II типа с асимметричными сайтами распознавания (например, показанными на данной веб-

странице), которые имеют сайт расщепления за пределами сайта распознавания и обеспечивают при расщеплении "липкие" концы с по меньшей мере тремя, предпочтительно 4 или более нуклеотидными остатками (например, BsmBI, BN736I; BpuAI, VpaK321, SfaNI и т. д.). В одном варианте осуществления эндонуклеаза рестрикции II типа представляет собой BsmBI.

[51] Рекомендуются, чтобы сайт распознавания содержал не менее 4, более предпочтительно по меньшей мере 6 или более пар оснований для минимизации вероятности обнаружения такого сайта в представляющей интерес части последовательности. Также можно применять нуклеазы рестрикции II типа с сайтами распознавания размером 5 п. о. (например, SfaNI). Эндонуклеазы рестрикции II типа, которые образуют односторонние "липкие" концы длиной 4 нуклеотида на концах расщепленных фрагментов, теоретически могут образовывать концы с 256 возможными последовательностями. Были сконструированы ферменты рестрикции II типа с еще более длинными сайтами распознавания, например, содержащими десять или более пар оснований. В одном варианте осуществления сайт распознавания представляет собой 5'-CGTCTC-3'. В одном варианте осуществления при расщеплении эндонуклеазой рестрикции IIS длина 5'-"липкого" конца составляет четыре нуклеотида, но он не выбран из группы, состоящей из GTAA, TCCA и CACA. В одном варианте осуществления при расщеплении эндонуклеазой рестрикции IIS длина 5'-"липкого" конца составляет четыре нуклеотида, и он выбран из группы, состоящей из AGGT, AGTA, ATCA, CAGT, CCAT, GAAT, GAGG, GGCA, GGTC, TAGC, TCTT, GGAG и CCAC.

[52] Примеры лигаз, которые необходимо применять в настоящем изобретении, включают ДНК-лигазу T4, ДНК-лигазу E. coli, ДНК-лигазу Taq, все из которых коммерчески доступны от New England Biolabs.

[53] В одном аспекте настоящее изобретение направлено на способ экспрессии многоцепочечного белка, содержащего по меньшей мере две различные полипептидные цепи, включающий

[54] а. обеспечение наличия полинуклеотидных последовательностей, которые кодируют по меньшей мере два различных полипептида, где указанные полинуклеотидные последовательности фланкированы с 5'- и 3'-сторон сайтом расщепления эндонуклеазой рестрикции IIS типа, за которым следует сайт ее распознавания;

[55] б. обеспечение наличия двух или более наборов последовательностей элементов, где каждый набор последовательностей элементов вместе содержит по меньшей мере одну функциональную кассету экспрессии,

[56] где каждая последовательность элемента фланкирована с 5'- и 3'-сторон сайтом расщепления эндонуклеазой рестрикции IIS типа, за которым следует сайт ее распознавания,

[57] где сайты распознавания эндонуклеазой рестрикции IIS типа и сайты расщепления выбраны таким образом, чтобы наборы последовательностей элементов

могли быть собраны в функциональную кассету экспрессии;

[58] с. обеспечение наличия по меньшей мере одного исходного вектора-остова, где исходный вектор-остов предусматривает плазмиду, содержащую по меньшей мере один сайт расщепления эндонуклеазой рестрикции PIS типа, за которым следует сайт ее распознавания и последовательности для поддержания вектора в бактериальных клетках;

[59] d. сборку полинуклеотидных последовательностей, которые кодируют по меньшей мере два различных полипептида, с двумя или более наборами последовательностей элементов для получения функциональных кассет экспрессии в исходном векторе-остове с помощью способа, основанного на применении расщепления ферментом рестрикции PIS типа и лигирования посредством сайтов расщепления и "липких" концов, в результате чего получают лигированные векторы, содержащие по меньшей мере две функциональные кассеты экспрессии, которые способны экспрессировать по меньшей мере две различные полипептидные цепи;

[60] e. трансфекцию клеток млекопитающих лигированными векторами;

[61] f. сравнение уровней полученного предпочтительного продукта и уровней примесей для каждого из по меньшей мере двух полипептидов и

[62] g. выявление оптимального соотношения экспрессии двух полипептидов для трансфицированных клеток млекопитающих и выявление трансфицированных клеток млекопитающих, которые характеризуются оптимальным соотношением экспрессии по меньшей мере двух полипептидных цепей.

[63] В одном варианте осуществления способ дополнительно включает

[64] h. выявление расположения элементов вектора, которое обеспечивает оптимальное соотношение экспрессии по меньшей мере двух полипептидных цепей.

[65] В одном аспекте настоящее изобретение направлено на вектор, характеризующийся расположением элементов вектора, которое обеспечивает оптимальное соотношение экспрессии по меньшей мере двух полипептидных цепей.

[66] Каждый набор последовательностей элементов, как правило, может быть собран в виде кассеты экспрессии. В контексте настоящего изобретения подразумевается, что выражение "кассета экспрессии" обозначает последовательность нуклеиновой кислоты, которая направляет клеточный аппарат на синтез РНК и белка. Обычно кассета экспрессии будет содержать кодирующую последовательность и последовательности, контролирующие экспрессию такой кодирующей последовательности. Обычно кассета экспрессии может содержать по меньшей мере промотор, открытую рамку считывания и терминирующую последовательность. Другие последовательности элементов включают контрольную последовательность, инсуляторную последовательность, баркодовую последовательность ДНК, последовательности праймера, промоторные последовательности, последовательности поли-А и последовательности IRES, а также последовательность селективного маркера млекопитающих. В одном варианте осуществления последовательности элементов содержат по меньшей мере две промоторные последовательности и по меньшей мере две последовательности поли-А.

[67] Термин "контрольные последовательности" в данном документе определяется как включающий все компоненты, которые необходимы или преимущественны для получения mRNA или полипептида либо в условиях *in vitro*, либо в клетке-хозяине. Каждая контрольная последовательность может быть нативной или чужеродной по отношению к последовательности нуклеиновой кислоты, кодирующей полипептид. Такие контрольные последовательности включают без ограничения лидерную последовательность Шайна-Дальгарно, оптимальные последовательности инициации трансляции (которые описаны в Kozak, 1991, J. Biol. Chem. 266:19867-19870), последовательность полиаденилирования, пропептидную последовательность, препропептидную последовательность, промотор, сигнальную последовательность и сигнал терминации транскрипции. Как минимум, контрольные последовательности обычно включают промотор и сигнал окончания транскрипции (терминатор или сигнал терминации). Обычно также могут присутствовать сигналы начала и окончания трансляции. Контрольные последовательности могут быть оптимизированы в соответствии с их конкретной целью.

[68] Термин "промотор" в данном документе определяется как последовательность ДНК, которая связывает РНК-полимеразу и направляет полимеразу к правильному, находящемуся ниже сайту начала транскрипции последовательности нуклеиновой кислоты, кодирующей биологическое соединение, для инициации транскрипции. РНК-полимераза эффективно катализирует сборку матричной РНК, комплементарной соответствующей нити ДНК кодирующей область. Термин "промотор" также следует понимать как включающий 5' -некодирующую область (между промотором и сайтом начала трансляции) для трансляции после транскрипции в mRNA, цис-действующие элементы контроля транскрипции, такие как энхансеры, и другие нуклеотидные последовательности, способные взаимодействовать с факторами транскрипции.

[69] Способ по настоящему изобретению обычно осуществляют таким образом, чтобы элементы кассеты экспрессии собирались во исходном векторе-остове таким образом, чтобы они находились в функциональной связи. Термины "функциональная связь", или "функционально связанный", или тому подобное в данном документе определяются как конфигурация, в которой контрольная последовательность соответствующим образом размещена в положении относительно кодирующей последовательности последовательности ДНК таким образом, чтобы контрольная последовательность управляла продуцированием mRNA или полипептида.

[70] "Инсуляторная последовательность" или "инсуляторы" представляют собой сегменты нуклеиновой кислоты, которые снижают или исключают влияние транскрипции с соседних областей на сегмент нуклеиновой кислоты, с которым связан инсулятор. Инсуляторы предпочтительно расположены выше других контрольных последовательностей и/или ниже генов. Инсуляторы предпочтительно размещены между различными генами, единицами транскрипции или генетическими доменами с тем, чтобы уменьшить или предотвратить интерференцию от соседних последовательностей

экспрессии.

[71] Функция "энхансерной последовательности" или "энхансеров" заключается в усилении транскрипции с промоторов, расположенных вблизи энхансера. Энхансеры могут функционировать как выше, так и ниже гена, и в любой ориентации.

[72] "Баркодовую последовательность ДНК" или "баркоды" можно применять для выявления молекул нуклеиновой кислоты, например, когда секвенирование может выявить определенный баркод, связанный с представляющей интерес молекулой нуклеиновой кислоты. В некоторых случаях для выявления молекулы нуклеиновой кислоты можно использовать специфичное для последовательности событие, при этом по меньшей мере часть баркода распознается в специфичном для последовательности событии, например, по меньшей мере часть баркода может участвовать в реакции лигирования или удлинения цепи. Таким образом, баркод позволяет выявить, отобрать или амплифицировать связанные с ним молекулы ДНК.

[73] "IRES" или "внутренний сайт связывания рибосомы" означает сайт, который обеспечивает внутреннее связывание рибосомы, достаточное для инициирования трансляции в анализе на кэп-независимую трансляцию, таком как описанный в данном документе анализ с использованием бицистронного репортера. Наличие IRES обеспечивает кэп-независимую трансляцию связанной последовательности, кодирующей белок, которая не транслировалась бы в противном случае.

[74] "Ген селектируемого маркера" или "селектируемый маркер" кодирует белок, необходимый для выживаемости и роста клетки-хозяина, растущей в селективной среде для культивирования. Типичные гены маркеров для отбора кодируют белки, которые (а) придают устойчивость к антибиотикам или другим токсинам, например, ампициллину, тетрациклину или канамицину в случае прокариотических клеток-хозяев; (b) восполняют ауксотрофные недостатки у клетки; или (с) обеспечивают необходимые питательные вещества, не доступные из комплексной или определенной среды. Специфические селектируемые маркеры представляют собой ген устойчивости к канамицину, ген устойчивости к ампициллину и ген устойчивости к тетрациклину. Преимущественно для отбора как прокариотических, так и эукариотических клеток-хозяев также можно применять ген устойчивости к неомицину.

[75] Для амплификации гена, который будет экспрессироваться, можно применять другие селектируемые гены. Амплификация представляет собой процесс, при котором гены, которые требуются для продуцирования белка, крайне важного для роста или выживания клеток, многократно повторяются в тандеме в хромосомах последовательных поколений рекомбинантных клеток. Примеры подходящих селектируемых маркеров для клеток млекопитающих включают гены дигидрофолатредуктазы (DHFR) и тимидинкиназы, не содержащие промоторов. Трансформанты клеток млекопитающих помещают в условия давления отбора, где только лишь трансформанты адаптированы к выживанию вследствие присутствия в векторе гена селектируемого маркера. Селекционное давление устанавливают посредством культивирования

трансформированных клеток в условиях, при которых концентрация средства для селекции в среде последовательно увеличивается, обеспечивая тем самым амплификацию как селектируемого гена, так и ДНК, которая кодирует другой ген, такой как один или несколько компонентов конструкций на основе многоцепочечных белков, описанных в данном документе. В результате с амплифицированной ДНК синтезируются повышенные количества полипептида.

[76] Соответственно, элемент в контексте настоящего изобретения представляет собой любой компонент кассеты экспрессии. Набор элементов представляет собой группу элементов, которые вместе могут образовывать кассету экспрессии. Способ по настоящему изобретению предусматривает обеспечение наличия двух наборов последовательностей элементов. Это означает, что необходимо обеспечить наличие достаточного количества элементов для того, чтобы можно было получить по меньшей мере две различные кассеты экспрессии. Это подразумевает, что необходимо обеспечить наличие по меньшей мере двух различных разновидностей по меньшей мере одного элемента. То есть один промотор в сочетании с двумя ORF и двумя сигналами терминации составляет два набора последовательностей элементов для целей настоящего изобретения.

[77] В способе по настоящему изобретению обычно по меньшей мере два набора последовательностей элементов содержат промоторный элемент, элемент открытой рамки считывания и элемент сигнала терминации.

[78] В способе согласно настоящему изобретению один или несколько наборов элементов могут содержать "частичные" последовательности элементов, такие как UTR, сигнальные пептиды и разделенные открытые рамки считывания.

[79] Каждый набор последовательностей элементов представлен в такой форме, чтобы набор мог быть собран в функциональную кассету экспрессии в исходном векторе-остове. Затем обычно каждый элемент с обеих сторон фланкируют сайтом расщепления эндонуклеазой рестрикции II типа, за которым следует сайт ее распознавания, при этом сайты распознавания эндонуклеазой рестрикции II типа и сайты расщепления выбраны таким образом, чтобы наборы последовательностей элементов могли быть собраны в функциональную кассету экспрессии. Таким образом, каждая последовательность элемента и фланкирующая последовательность обычно содержит в порядке от одного конца к другому: сайт распознавания эндонуклеазой рестрикции II типа, сайт расщепления ею, последовательность элемента, сайт расщепления эндонуклеазой рестрикции II типа, сайт ее распознавания.

[80] Соответственно, наборы элементов получают или обеспечивают их наличие в подходящем векторе с сайтами распознавания эндонуклеазой рестрикции II типа и стандартизированными сайтами расщепления (предпочтительно из 4 п. о.), выбранными таким образом, что после сборки, например с использованием однореакторного подхода, такого как клонирование Golden Gate, образовывается функциональная кассета экспрессии.

[81] Получают или обеспечивают наличие набора исходных векторов-остовов.

Такие векторы содержат левые и правые соединительные последовательности, подходящие для сборки с помощью гомологии последовательностей.

[82] Выбирают подмножество последовательностей элементов вместе с исходными векторами-остовами (bbn). Их можно собрать, например, с помощью клонирования Golden Gate, что приведет к получению функциональных кассет экспрессии, входящих в состав исходных векторов-остовов.

[83] Иллюстративные клетки-хозяева включают прокариотические, дрожжевые клетки или клетки высших эукариот. Прокариотические клетки-хозяева включают эубактерии, такие как грамотрицательные или грамположительные организмы, например, *Enterobacteriaceae*, такие как *Escherichia*, например, *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, например, *Salmonella typhimurium*, *Serratia*, например, *Serratia marcescans* и *Shigella*, а также *Bacillus*, например, *B. subtilis* и *B. licheniformis*, *Pseudomonas* и *Streptomyces*. Эукариотические микроорганизмы, такие как мицелиальные грибы или дрожжи, являются подходящими хозяевами для клонирования или экспрессии рекомбинантных полипептидов. *Saccharomyces cerevisiae* или обычные пекарские дрожжи наиболее часто применяются среди низших эукариотических микроорганизмов-хозяев. Однако ряд других родов, видов и штаммов являются общедоступными и применимыми в данном документе, такие как *Pichia*, например, *P. pastoris*, *Schizosaccharomyces pombe*; *Kluyveromyces*, *Yarrowia*; *Candida*; *Trichoderma reesia*; *Neurospora crassa*; *Schwanniomyces*, например, *Schwanniomyces occidentalis*, и мицелиальные грибы, такие как, например, *Neurospora*, *Penicillium*, *Tolyocladium* и хозяева *Aspergillus*, такие как *A. nidulans* и *A. niger*.

[84] Клетки-хозяева для экспрессии гликозилированных антигенсвязывающих белков могут быть получены из многоклеточных организмов. Примеры клеток беспозвоночных включают клетки растений и насекомых. Были выявлены многочисленные бакуловирусные штаммы и варианты и соответствующие перmissive клетки-хозяева насекомых из таких хозяев, как *Spodoptera frugiperda* (гусеница), *Aedes aegypti* (комар), *Aedes albopictus* (комар), *Drosophila melanogaster* (плодовая мушка) и *Bombyx mori*. Общедоступен ряд вирусных штаммов для трансфекции таких клеток, например, вариант L-1 NPV *Autographa californica* и штамм NPV Bm-5 *Bombyx mori*.

[85] Клетки-хозяева позвоночных также являются подходящими хозяевами, и рекомбинантное получение антигенсвязывающих белков из таких клеток стало обычной процедурой. Линии клеток млекопитающих, доступные в качестве хозяев для экспрессии, хорошо известны из уровня техники и включают без ограничения immortalized линии клеток, доступные из Американской коллекции типовых культур (ATCC), включая без ограничения клетки яичника китайского хомячка (CHO), включая клетки линии CHOK1 (ATCC CCL61), DXB-11, DG-44 и клетки яичника китайского хомячка/-DHFR (CHO, Urlaub *et al*, Proc. Natl. Acad. Sci. USA 77: 4216, 1980); линию клеток CV1 почки обезьяны, трансформированную с помощью SV40 (COS-7, ATCC CRL 1651); эмбриональную линию клеток почки человека (клетки линии 293 или 293,

субклонированные для роста в суспензионной культуре (Graham *et al.*, J. Gen Virol. 36: 59, 1977); клетки почки новорожденного хомяка (ВНК, ATCC CCL 10); клетки Сертоли мыши (ТМ4, Mather, Biol. Reprod. 23: 243-251, 1980); клетки почки обезьяны (CV1 ATCC CCL 70); клетки почки африканской зеленой мартышки (VERO-76, ATCC CRL-1587); клетки карциномы шейки матки человека (HELA, ATCC CCL 2); клетки почки собаки (MDCK, ATCC CCL 34); клетки печени крысы линии buffalo (BRL 3A, ATCC CRL 1442); клетки легких человека (W138, ATCC CCL 75); клетки гепатомы человека (Hep G2, HB 8065); клетки опухоли молочной железы мыши (ММТ 060562, ATCC CCL51); клетки линии TRI (Mather *et al.*, Annals N.Y Acad. Sci. 383: 44-68, 1982); клетки линии MRC 5 или клетки линии FS4; клетки миеломы млекопитающих и ряд других линий клеток. В определенных вариантах осуществления линии клеток могут быть выбраны путем определения того, какие линии клеток обладают высокими уровнями экспрессии и конститутивно продуцируют многоцепочечные белки по настоящему изобретению. В другом варианте осуществления может быть отобрана линия клеток из В-лимфоцитарной линии дифференцировки, которая не продуцирует свое собственное антитело, но обладает способностью продуцировать и секретировать гетерологичное антитело. Клетки линии СНО являются клетками-хозяевами в некоторых вариантах осуществления для экспрессии многоцепочечных белков по настоящему изобретению.

[86] Клетки-хозяева трансформируют или трансфицируют с помощью описанных выше нуклеиновых кислот или векторов для получения многоцепочечных белков и культивируют в стандартных питательных средах, модифицированных в зависимости от необходимости индуцирования промоторов, отбора трансформантов или амплификации генов, кодирующих требуемые последовательности. Кроме того, новые векторы и трансфицированные линии клеток с несколькими копиями транскрипционных единиц, разделенных селективным маркером, особенно полезны для экспрессии антигенсвязывающих белков. Таким образом, настоящее изобретение также предусматривает способ получения многоцепочечных белков, описанных в данном документе, включающий культивирование клетки-хозяина, содержащей один или несколько векторов экспрессии, описанных в данном документе, в культуральной среде в условиях, обеспечивающих экспрессию многоцепочечных белков, кодируемых одним или несколькими векторами экспрессии; и извлечение многоцепочечных белков из культуральной среды.

[87] Клетки-хозяева, применяемые для получения антигенсвязывающих белков по настоящему изобретению, можно культивировать в различных средах. Коммерчески доступные среды, такие как Ham's F10 (Sigma), минимальная питательная среда ((MEM), (Sigma), RPMI-1640 (Sigma) и модифицированная по Дульбекко среда Игла ((DMEM), Sigma), подходят для культивирования клеток-хозяев. Кроме того, любую из сред, описанных в Ham *et al.*, Meth. Enz. 58: 44, 1979; Barnes *et al.*, Anal. Biochem. 102: 255, 1980, патентах США №№ 4767704, 4657866, 4927762, 4560655 или 5122469, WO90103430, WO 87/00195 или патенте США № 30985 можно применять в качестве культуральной среды

для клеток-хозяев. Любую из этих сред можно при необходимости дополнить гормонами и/или другими факторами роста (такими как инсулин, трансферрин или эпидермальный фактор роста), солями (такими как хлорид натрия, кальций, магний и фосфат), буферами (такими как HEPES), нуклеотидами (такими как аденозин и тимидин), антибиотиками (такими как лекарственное средство Gentamycin™), микроэлементами (определяемыми как неорганические соединения, обычно присутствующие в конечных концентрациях в микромолярном диапазоне) и глюкозой или эквивалентным источником энергии. Любые другие необходимые добавки также могут быть включены в соответствующих концентрациях, которые должны быть известны специалистам в данной области техники. Условия культивирования, такие как температура, pH и другие, являются такими, которые ранее применялись для клетки-хозяина, выбранной для экспрессии, и будут очевидны специалисту в данной области техники.

[88] В ходе культивирования клеток-хозяев многоцепочечные белки можно получать внутриклеточно, в периплазматическом пространстве, или они могут секретироваться непосредственно в среду. Если антигенсвязывающий белок продуцируется внутриклеточно, то в качестве первой стадии удаляют твердые остатки, либо клетки-хозяева, либо лизированные фрагменты, например путем центрифугирования или ультрафильтрации. Биспецифический антигенсвязывающий белок можно очищать с применением, например, гидроксиапатитной хроматографии, катионо- или анионообменной хроматографии или аффинной хроматографии с применением представляющего(-их) интерес антигена(-ов) или белка А или белка G в качестве аффинного лиганда. Белок А можно использовать для очистки белков, которые содержат полипептиды на основе тяжелых цепей $\gamma 1$, $\gamma 2$ или $\gamma 4$ человека (Lindmark *et al.*, J. Immunol. Meth. 62: 1-13, 1983). Белок G рекомендуется для всех мышинных изоформ и для $\gamma 3$ человека (Guss *et al.*, EMBO J. 5: 1567-1575, 1986). Матрица, к которой присоединяют аффинный лиганд, наиболее часто представляет собой агарозу, но также доступны другие матрицы. Механически стабильные матрицы, такие как стекло с контролируемым размером пор или поли(стиролдивинил)бензол, обеспечивают более высокие скорости потока и меньшее время обработки, чем те, которые достигаются при использовании агарозы. В случае если белок содержит домен СН3, для очистки можно применять смолу Bakerbond АВХ™ (J. T. Baker, Phillipsburg, N.J.). Другие методики очистки белков, такие как осаждение этанолом, обращенно-фазовая HPLC, хроматофокусирование, SDS-PAGE и осаждение сульфатом аммония, также возможны в зависимости от конкретных многоцепочечных белков, которые необходимо извлечь.

[89] Применяемый в данном документе термин "антитело" относится к тетрамерному белку-иммуноглобулину, содержащему два полипептида легкой цепи (приблизительно 25 кДа каждый) и два полипептида тяжелой цепи (приблизительно 50-70 кДа каждый). Термин "легкая цепь" или "легкая цепь иммуноглобулина" относится к полипептиду, содержащему от аминоконца к карбоксильному концу одну переменную область легкой цепи (VL) иммуноглобулина и один константный домен легкой цепи (CL)

иммуноглобулина. Константный домен легкой цепи (CL) иммуноглобулина может быть каппа- (κ) или лямбда-(λ)-типа. Термин "тяжелая цепь" или "тяжелая цепь иммуноглобулина" относится к полипептиду, содержащему от аминоконца к карбоксильному концу одну переменную область тяжелой цепи (VH) иммуноглобулина, константный домен 1 тяжелой цепи (CH1) иммуноглобулина, шарнирную область иммуноглобулина, константный домен 2 тяжелой цепи (CH2) иммуноглобулина, константный домен 3 тяжелой цепи (CH3) иммуноглобулина и необязательно константный домен 4 тяжелой цепи (CH4) иммуноглобулина. Тяжелые цепи классифицируют как мю- (μ), дельта- (Δ), гамма- (γ), альфа- (α) и эpsilon- (ϵ) цепи, и они определяют изотип антитела как IgM, IgD, IgG, IgA и IgE соответственно. Классы антител IgG и IgA дополнительно делят на подклассы, а именно IgG1, IgG2, IgG3 и IgG4, а также IgA1 и IgA2 соответственно. Тяжелые цепи в антителах IgG, IgA и IgD содержат три домена (CH1, CH2 и CH3), тогда как тяжелые цепи в антителах IgM и IgE содержат четыре домена (CH1, CH2, CH3 и CH4). Константные домены тяжелой цепи иммуноглобулина могут происходить из иммуноглобулина любого изотипа, включая подтипы. Цепи антитела связаны друг с другом посредством межполипептидных дисульфидных связей между доменом CL и доменом CH1 (т. е. между легкой и тяжелой цепью) и между шарнирными областями тяжелых цепей антитела.

[90] В человеческом антителе CH1 означает область, содержащую аминокислотную последовательность в пределах положений 118-215 согласно EU-индексу. Область из аминокислот, характеризующаяся высокой гибкостью, называемая "шарнирной областью", присутствует между CH1 и CH2. CH2 представляет собой область, содержащую аминокислотную последовательность в пределах положений 231-340 согласно EU-индексу, и CH3 представляет собой область, содержащую аминокислотную последовательность в пределах положений 341-446 согласно EU-индексу.

[91] "CL" представляет собой константную область легкой цепи. В случае κ -цепи человеческого антитела CL представляет собой область, содержащую аминокислотную последовательность в пределах положений 108-214 согласно EU-индексу. В λ -цепи CL представляет собой область, содержащую аминокислотную последовательность в пределах положений 108-215.

[92] Схемы нумерации, представленные как EU-индексом, как например согласно Kabat *et al.*, Sequences of Proteins of Immunological Interest, 5th Ed. Public Health Service, National Institutes of Health, Bethesda, MD (1991), так и нумерацией АНо (Honegger A. and Plückthun A. *J Mol Biol.* 2001 Jun 8;309(3):657-70), могут использоваться в настоящем изобретении. Положения аминокислот и определяющих комплементарность областей (CDR), а также каркасных областей (FR) данного антитела можно идентифицировать с использованием любой системы. Например, положения 39, 44, 183, 356, 357, 360, 370, 392, 399 и 409 в тяжелой цепи согласно EU являются эквивалентными положениям 46, 51, 230, 484, 485, 491, 501, 528, 535 и 551 в тяжелой цепи согласно АНо соответственно.

[93] В одном варианте осуществления оптимальное соотношение экспрессии по меньшей мере двух полипептидов выбрано из группы, состоящей из 1:1, 1:2 и 1:3.

[94] В одном варианте осуществления многоцепочечный белок содержит первую тяжелую цепь антитела, первую легкую цепь антитела, вторую тяжелую цепь антитела и вторую легкую цепь антитела, где первая тяжелая цепь антитела соединяется с первой легкой цепью антитела для связывания первого антигена или эпитопа, а вторая тяжелая цепь антитела соединяется со второй легкой цепью антитела для связывания второго антигена или эпитопа, где оптимальное соотношение экспрессии первой тяжелой цепи антитела, первой легкой цепи антитела, второй тяжелой цепи антитела и второй легкой цепи антитела составляет 1:1:1:1.

[95] В одном варианте осуществления многоцепочечный белок содержит первую тяжелую цепь антитела, вторую тяжелую цепь антитела и общую легкую цепь антитела, где первая тяжелая цепь антитела соединяется с общей легкой цепью антитела для связывания первого антигена или эпитопа, а вторая тяжелая цепь антитела соединяется с общей легкой цепью антитела для связывания второго антигена или эпитопа, где оптимальное соотношение экспрессии первой тяжелой цепи антитела, второй тяжелой цепи антитела и общей легкой цепи антитела составляет 1:1:2.

[96] В одном варианте осуществления многоцепочечный белок содержит тяжелую цепь антитела, первую легкую цепь антитела, модифицированную тяжелую цепь антитела и вторую легкую цепь антитела,

[97] где модифицированная тяжелая цепь антитела содержит от N-конца к С-концу одну из нижеуказанных структур, выбранных из следующей группы:

[98] VH-CH1-связывающий домен (BD)-шарнир-CH2-CH3,

[99] BD-VH-CH1-шарнир-CH2-CH3 и

[100] VH-CH1-шарнир-CH2-CH3-BD,

[101] где BD выбран либо как одноцепочечный Fv (scFv), либо как одноцепочечный Fab (scFab);

[102] где тяжелая цепь антитела соединяется с первой легкой цепью антитела для связывания первого антигена или эпитопа, а VH модифицированной тяжелой цепи антитела соединяется со второй легкой цепью антитела для связывания второго антигена или эпитопа, где BD связывается с третьим антигеном или эпитопом, и

[103] где оптимальное соотношение экспрессии первой тяжелой цепи антитела, первой легкой цепи, модифицированной тяжелой цепи антитела и второй легкой цепи антитела составляет 1:1:1:1.

[104] В одном варианте осуществления многоцепочечный белок содержит тяжелую цепь антитела, модифицированную тяжелую цепь антитела и общую легкую цепь антитела,

[105] где модифицированная тяжелая цепь антитела содержит от N-конца к С-концу одну из нижеуказанных структур, выбранных из следующей группы:

[106] VH-CH1-связывающий домен (BD)-шарнир-CH2-CH3,

[107] BD-VH-CH1-шарнир-CH2-CH3 и

[108] VH-CH1-шарнир-CH2-CH3-BD,

[109] где BD выбран либо как одноцепочечный Fv (scFv), либо как одноцепочечный Fab (scFab);

[110] где тяжелая цепь антитела соединяется с общей легкой цепью антитела для связывания первого антигена или эпитопа, а VH модифицированной тяжелой цепи антитела соединяется с общей легкой цепью антитела для связывания первого антигена или эпитопа, где BD связывается со вторым антигеном или эпитопом, и

[111] где оптимальное соотношение экспрессии первой тяжелой цепи антитела, модифицированной тяжелой цепи антитела и общей легкой цепи антитела составляет 1:1:2.

[112] "Связывающий домен" или "BD" может обычно содержать вариабельную область легкой цепи (VL) антитела и вариабельную область тяжелой цепи (VH) антитела, однако он не обязательно должен содержать обе из них. Fd-фрагменты, например, имеют две области VH и часто сохраняют в некоторой мере антигенсвязывающую функцию интактного антигенсвязывающего домена. Дополнительные примеры формата фрагментов антител, вариантов антител или связывающих доменов включают (1) Fab-фрагмент, моновалентный фрагмент с доменами VL, VH, CL и CH1; (2) F(ab')₂-фрагмент, бивалентный фрагмент с двумя Fab-фрагментами, соединенными дисульфидным мостиком в шарнирной области; (3) Fd-фрагмент с двумя доменами VH и CH1; (4) Fv-фрагмент с доменами VL и VH из одного плеча антитела, (5) dAb-фрагмент (Ward et al., (1989) Nature 341: 544-546), который имеет домен VH; (6) выделенная определяющая комплементарность область (CDR) и (7) одноцепочечный Fv (scFv), причем последний является предпочтительным (например, получен из библиотеки scFv). Катионообменная хроматография ("СЕХ") является формой ионообменной хроматографии (ИЕХ), которую применяют для разделения молекул на основе их суммарного поверхностного заряда. Более конкретно, в катионообменной хроматографии применяют отрицательно заряженную ионообменную смолу, обладающую аффинностью к молекулам, характеризующимся положительным поверхностным зарядом. Катионообменную хроматографию применяют как в препаративных, так и в аналитических целях, и она позволяет разделять широкий диапазон молекул - от аминокислот и нуклеотидов до крупных белков. В данном случае авторы настоящего изобретения сосредотачиваются на препаративной катионообменной хроматографии белков. СЕХ можно проводить в восстанавливающих и невосстанавливающих условиях.

[113] Масс-спектрометрия ("MS") представляет собой аналитическую методику, применяемую для измерения соотношения массы к заряду ионов. Результаты представляют в виде масс-спектра, графика интенсивности в виде зависимости отношения массы к заряду.

[114] Раскрытие каждого литературного источника, изложенного в данном документе, включено в данный документ посредством ссылки во всей своей полноте.

[115] Настоящее изобретение дополнительно проиллюстрировано последующими примерами.

[116] ПРИМЕРЫ

[117] ModVec сконструирована с исключительной универсальностью с помощью модульного "слотового" подхода для сборки одной или нескольких кассет экспрессии в ряде возможных конфигураций, при этом делая возможным комбинаторное исследование разнообразия последовательностей в каждом слоте. Система ModVec разделена на слоты, где каждый слот может содержать последовательность ДНК с уникальными "липкими" концами Golden Gate. Данные последовательности ДНК могут быть элементами вектора, в том числе элементами, которые необходимы для репликации вектора (ColE Ori, pMB1 Ori...), устойчивости к антибиотикам (Amp, Kan, Cm...), элементами, которые необходимы для инфицирования вирусом (LTR), транспозиции (ITR), эписомальной репликации (OriP), направленной интеграции (loxP, Frt, attB/P), и элементами для экспрессии генов млекопитающих (например, промоторами, энхансерами, внутренними сайтами связывания рибосомы (IRES), самопроцессирующимся вирусным пептидом 2A, сигналом поли-А, контрольной последовательностью, инсуляторными последовательностями, в том числе MAR и UCSE и т. д.).

[118] В качестве доказательства идеи ModVec была впервые использована для сборки сложных векторов экспрессии у млекопитающих. Были созданы два остова ModVec, в том числе pMVP5, который можно применять в обычных линиях клеток млекопитающих, применяемых в исследовании (НЕК, СНО), и остов вектора, который был разработан специально для производственных линий клеток СНО Amgen. Данные остовы вектора содержат лишь минимальные элементы вектора, необходимые для репликации и отбора с помощью антибиотиков в *E. coli* для поддержания плазмиды.

[119] В первом эксперименте концепцию ModVec тестировали путем сборки вектора с функциональной последовательностью ДНК в каждом слоте: сборка ДНК размером 8 т. о. из 14 слотов (включая остов вектора). Он содержал три ЕС; ЕС1 и ЕС2 были бицистронными, а ЕС3 был моноцистронным. Он содержал три различных промотора для управления экспрессией 5 кодирующих последовательностей. Было продемонстрировано, что применение кольцевых плазмид в качестве вводных элементов для реакций GG повышает эффективность, особенно для многокомпонентной сборки Golden Gate (Potapov, Ong et al. 2018). Поэтому были созданы векторы из "частей", в которых элементы вектора, для которых ожидалось, что они будут часто повторно использоваться в каждом слоте, были клонированы в вектор pGG_Cm_BsaI сначала с помощью реакции GG с BsaI, а затем подтверждены секвенированием по Сэнгеру. В ходе данного процесса была построена "библиотека частей" элементов вектора, таких как промоторы, IRES, сигналы поли-А, контрольные последовательности, инсуляторы и т. д., с подходящими для ModVec "липкими" концами в формате, который будет легко хранить и при необходимости пополнять. Используя комбинацию кольцевых "частей" и синтезированной линейной ДНК, для первой сборки из 14 слотов использовали 20 мкл

смеси для реакции GG с 1 мкл каждого элемента вектора (в концентрации 40-70 нг/мкл). В качестве остова вектора использовали pMVP5. Смесь для реакции GG была трансформирована в *E. coli*, а для секвенирования по Сэнгеру были отобраны 32 колонии. Был успешно получен конструируемый ДНК-вектор размером 8 т. о., однако эффективность правильной сборки составляла лишь 3,125% (правильной оказалась 1 из 32 отобранных колоний). Из внутренних, а также внешних данных (Kanigowska, Shen et al. 2016) было видно, что эффективность реакции GG может быть выше при меньшем объеме реакции (миниатюрная реакция GG). Затем производили сборку 14-слотовой системы при меньших объемах реакции: 2 мкл, 1 мкл и 0,5 мкл с помощью акустического устройства для манипуляции с жидкостями ЕСНО, которое обеспечивает точный перенос жидкости при данных небольших объемах. Данные реакции были пропорционально уменьшены в масштабе по сравнению с первоначальной реакцией объемом 20 мкл. Миниатюрная реакция GG повышала эффективность данной сборки ModVec из 14 частей до 16% (правильными были 13 из 78 колоний).

[120] В результате тщательного анализа как успешных, так и неудачных попыток первых сборок PoC ModVec было установлено, что конкретный набор из 11 "липких" концов, полученных в реакции GG, всегда правильно лигировался в данном контексте. Затем тестировали, может ли использование лишь данного набора "липких" концов повысить эффективность ModVec. ModVec использовали для сборки 10-слотовой сборки с данными оптимизированными "липкими" концами в реакции GG с объемом 1 мкл. Было обнаружено, что более 95% клонов содержали правильно собранные векторы (правильными были 548 из 576 колоний).

[121] 95% эффективность сборки, полученная при оптимизации конструкции и в условиях ModVec, делает возможной создание библиотеки векторов экспрессии в одной пробирке. Функционально схожие элементы вектора, которые содержали одинаковые "липкие" концы, смешивают для получения "библиотек слотов". 2 различных промотора для управления 12 CDS в EC1 объединяли с 3 различными промоторами для управления экспрессией маркера EC3 для отбора, в том числе мостиком, который устраняет EC2, что посредством комбинаторной сборки могло обеспечивать 72 различных вектора экспрессии. После трансформации отбирали 350 колоний, а после проверки результатов секвенирования получали в общей сложности 312 контигов. Данные 312 контигов соответствовали 69 из 72 возможных векторов, что демонстрирует низкие зависящие от последовательностей смещения при сборке Modvec и указывает на возможность реализации стратегии использования библиотеки с помощью данной системы.

[122] **Материалы и способы**

[123] *Создание векторов*

[124] Все фрагменты ДНК, в том числе элементы вектора и LC и HC антитела, были синтезированы компанией Twist Bioscience или амплифицированы методом ПЦР с существующих матриц и клонированы в остовы вектора с помощью стандартного протокола сборки Golden Gate (Engler et al., 2008). Для получения векторов из частей

каждая реакционная смесь Golden Gate содержала 2 мкл фрагментов ДНК (5 нг/мкл) и 2 мкл вектора pGG (20 нг/мкл), 1 мкл буфера FastDigest (Thermo Fisher, B64) с 5 мкМ АТФ (Thermo Fisher, R0441), 0,5 мкл лигазы T4 (Thermo Fisher, EL0014) и 0,5 мкл BsaI (Thermo Fisher, ER0291), а также 3 мкл H₂O. Реакцию Golden Gate проводили при 37°C в течение 2 мин. и при 16°C в течение 3 мин. в течение 15 циклов на термоциклере. Для дезактивации всех ферментов проводили заключительную инкубацию продолжительностью 5 мин. при 85°C. Миниатюризованная реакция Golden Gate и одnoreакторная реакция клонирования библиотеки векторов были организованы с использованием устройства для манипуляции с жидкостями ECHO 525 (Labcyte) для выполнения переноса жидкости в нанолитровом масштабе. Объем каждого фрагмента ДНК и остова вектора был пропорционально уменьшен для миниатюризованной реакции Golden Gate. Для сборки ModVec 15 мкл реакционной смеси Golden Gate, включающей 2 мкл векторов из частей pGG (40-70 нг/мкл), 2 мкл остовов вектора экспрессии (50 нг/мкл), 1,5 мкл буфера FastDigest (Thermo Fisher, B64) с 5 мкМ АТФ (Thermo Fisher, R0441), 0,75 мкл лигазы T4 (Thermo Fisher, EL0014) и 0,75 мкл BsmBI (Thermo Fisher, FD0454), а также переменным количеством H₂O для получения 15 мкл, смешивали с помощью устройства для манипуляции с жидкостями TECAN. Инкубацию при 37°C продлевали до 3 мин., а также инкубацию при 16°C продлевали до 5 мин. для каждого цикла Golden Gate, а количество циклов увеличивали до 20 для сборки ModVec. Плазмидную ДНК получали с применением набора Qiagen miniprep (Qiagen, 27104). После подтверждения результатов секвенирования плазмидную ДНК, кодирующую HC и LC, смешивали в определенных соотношениях.

[125] **Культивирование клеток и экспрессия белков**

[126] Для получения стабильных пулов клеток посредством случайной интеграции в геном 25 мкг ДНК были методом электропорации введены в принадлежащие компании заявителя запатентованные суспензионные клетки яичников китайского хомячка (CHO) с использованием систем электропорации Bio-Rad Gene Pulser® Xcell. После электропорации всю трансфицированную массу высевали в запатентованную восстановительную среду. Каждые 2-3 дня до восстановления отслеживали плотность жизнеспособных клеток и их жизнеспособность с помощью счетчика VI-CELL® (Beckman Coulter), а также производили замену среды. Восстановление определяли как > 85% жизнеспособности с помощью VI-CELL®. Восстановленные клетки использовали для посева 50-мл партий подпитываемых культур во встряхиваемых колбах, которые собирали через 10 дней. Для получения стабильных пулов клеток с использованием векторов экспрессии на основе транспозазы *piggyBac* 4 мкг ДНК использовали для трансфекции запатентованных суспензионных клеток CHO с нокаутом по глутаминсинтетазе (CHO GS KO) с использованием липофектамина LTX (Thermo Fisher, 11668030) в количестве 4×10^6 жизнеспособных клеток на мл. Трансфицированные клетки переносили через 48-72 часа после трансфекции в среду для отбора с метионинсульфоксимином (MSX). Восстановленные клетки использовали для посева 4-мл

партий подпитываемых культур в 24-луночных блоках для культивирования в количестве 1×10^6 клеток на мл, которые собирали через 10 дней. В ходе получения отслеживали плотность жизнеспособных клеток и их жизнеспособность с помощью счетчика VI-CELL® (Beckman Coulter), а также в день 3, 6 и 8 производили замену среды. В день 10 измеряли плотность жизнеспособных клеток и их жизнеспособность, а кондиционированные среды из данных полученных партий использовали для определения титра с помощью устройства ForteBio OCTET® Red, оснащенного биодатчиками с белком А.

[127] ***Высокоэффективная очистка белка***

[128] Для очистки белка использовали систему KingFisher® Flex (Thermo Fisher) с магнитными микроносителями с ProA (GE Life Sciences), как описано ранее (Gong et al., 2021). Вкратце, за день до сбора биомассы 100 мкл магнитных микроносителей с ProA добавляли к 4 мл культур CHO GS KO. Микроносители собирали и подвергали очистке с помощью системы Kingfisher с 24-зондовой магнитной головкой. После промывки 3 раза фосфатно-солевым буферным раствором и 2 раза водой Milli-Q белки элюировали посредством 500 мкл 100 mM ацетата натрия с pH 3,6 в течение 10 мин., а затем нейтрализовали добавлением 10 мкл 3 М Трис, pH 11,0.

[129] ***Крупномасштабная очистка с помощью ProA с последующей катионообменной хроматографией***

[130] Белки, экспрессированные в 50-мл культурах CHO, очищали так, как описано ранее (Gong et al., 2021; Yoo et al., 2014), с помощью аффинного захвата на ProA (1 мл HiTrap MabSelect SuRe, GE Life Sciences, № по каталогу GE11-0034-93), элюировали посредством 100 mM ацетата натрия, pH 3,6, с последующей немедленной заменой буфера на 10 mM ацетат натрия, 150 mM NaCl, pH 5,2, с применением колонки HiTrap Desalting объемом 5 мл (GE Life Sciences, № по каталогу GE17-1408-01).

[131] Катионообменную хроматографию проводили так, как описано ранее (Gong et al., 2021). Вкратце, 1,5-1,8 мл очищенных с помощью ProA образцов разводили посредством 20 мл 20 mM MES, pH 6,2, и загружали на 1-мл катионообменную колонку (SP-HP HiTrap, GE Life Sciences, № по каталогу GE29-0513-24) со скоростью 1 мл/мин. После промывки колонки посредством 8 объемов колонки того же буфера со скоростью 1 мл/мин. белки элюировали линейным градиентом 0-400 mM NaCl в течение 40 объемов колонки со скоростью 0,4 мл/мин. Фракции со степенью чистоты 90% и выше, определенной методом эксклюзионной хроматографии (SEC), объединяли, а их концентрацию определяли с помощью микропланшетного ридера Multiscan GO (Thermo Fisher) так, как описано ранее (Winters et al., 2015). Окончательные объединенные в пулы образцы анализировали методами SEC, микрокапиллярного электрофореза в невосстанавливающих условиях (nrMCE) и жидкостной хроматографии-масс-спектрометрии (LC-MS).

[132] ***Микрокапиллярный электрофорез в невосстанавливающих условиях и эксклюзионная хроматография***

[133] Качество продуктов в виде очищенных материалов анализировали с помощью микрокапиллярного электрофореза в восстанавливающих и невосстанавливающих условиях и SEC так, как описано ранее (Gong et al., 2021; Guo et al., 2021). Для nгMSE 6 мкл белка смешивали с 21 мкл буфера для образцов (8,4 мМ TrisHCl, pH 7,0, 7,98% глицерина, 2,38 мМ EDTA, 2,8% SDS и 2,4 мМ йодацетамида), нагревали при 85°C в течение 10 минут, а затем анализировали с помощью прибора Caliper LabChip GXII Touch (PerkinElmer). Для аналитической SEC образцы белка анализировали на приборе для HPLC Acquity® (Waters) с применением колонки ВЕН (200 Å, 1,7 микрона, 4,7×300 мм) с 100 мМ фосфата натрия, pH 6,9, 50 мМ NaCl, 7,5% этанола в качестве рабочего буфера со скоростью потока 0,45 мл/мин.

[134] ***Жидкостная хроматография-масс-спектрометрия и RapidFire-масс-спектрометрия***

[135] Для определения MW очищенных образцов проводили LC-MS интактных образцов в невосстанавливающих условиях так, как сообщалось ранее (Campuzano et al., 2019; Spahr et al., 2018). К 30 мкл каждого очищенного образца добавляли аликвоту объемом 30 мкл 1% трифторуксусной кислоты (TFA). Затем примерно 15 мкг каждого образца впрыскивали в устройство для UPLC Agilent 1290, при этом элюат из колонки был напрямую соединен с времяпролетным масс-спектрометром с электрораспылительным источником ионов (ESI-TOF) Agilent 6224 (Agilent Technologies). Для хроматографического разделения использовали колонку Zorbax® для сверхэффективной жидкостной хроматографии (UPLC) RRHD 300SB-C8, 2,1×50 мм, с размером частиц 1,8 мкм (Agilent Technologies). Колонку нагревали до 70°C со скоростью потока 0,5 мл/мин. В качестве хроматографических растворителей применяли водный раствор "А" (0,1% TFA в H₂O) и органический раствор "В" (0,1% TFA в 90% n-пропаноле). Использованный градиент был изократическим при соотношении 80% А/20% В в течение 4 мин., 28% А/72% В в течение 2 мин., 10% А/90% В в течение 0,5 мин. и, наконец, 5% А/95% В в течение 0,5 мин. Метод MS позволяет сканировать m/z [1000-7000] с получением 0,7 спектра/с. Полученные спектры суммировали, а затем подвергали деконволюции с помощью либо программного обеспечения Agilent Mass Hunter Qualitative Analysis (версии В.07.00), либо модуля Intact Program от Protein Metrics (PMI Intact).

[136] Систему для RapidFire-MS использовали для образцов, полученных в результате высокоэффективного скрининга библиотеки конфигураций векторов. К 50 мкл предоставленного образца добавляли равный объем 0,1% вес./вес. муравьиной кислоты и 20 мкл данного раствора вводили для анализа в устройство для RapidFire-MS. Картридж для SPE представлял собой 4-мкл картридж/колонку PLRP, 1000 Å. Подвижные фазы представляли собой 10% n-пропанол, содержащий 0,1% муравьиной кислоты, и 90% n-пропанол, содержащий 0,1% муравьиной кислоты. Все данные обрабатывали с помощью PMI Intact.

[137] **Результаты**

[138] ***Платформа Modular Vector (ModVec) для высокоэффективной сборки***

сложных векторов

[139] Для обеспечения высокоэффективного конструирования векторов была разработана система модульных векторов для млекопитающих на основе сборки GG (ModVec), позволяющая создавать разнообразные векторы экспрессии и библиотеки векторов для получения рекомбинантных белков. ModVec обеспечивает высокоэффективное создание как простых, так и сложных векторных конструкций с сохранением оптимизации векторов экспрессии для отдельных молекул, для конкретных крупномолекулярных соединений и/или для различных хозяев для экспрессии.

[140] ModVec сконструирована с исключительной универсальностью для сборки одной или нескольких кассет экспрессии в ряде возможных конфигураций, при этом делая возможным комбинаторное исследование разнообразия последовательностей в каждом модуле. Каждый модуль содержит последовательность ДНК или библиотеки последовательностей ДНК с тщательно сконструированными "липкими" концами в реакции GG (фигура 1B). Данные последовательности ДНК могут представлять собой элементы вектора, обычно применяемые при рекомбинантной экспрессии, такие как последовательности, которые необходимы для репликации вектора (точка начала репликации) и устойчивости к антибиотикам, элементы, которые необходимы для инфицирования вирусом, транспозиции, эписомальной репликации, направленной интеграции, или элементы, обычно применяемые для экспрессии генов млекопитающих. Данные последовательности ДНК фланкированы адаптерами GG (выделенная рамкой область на фигура 1A), которые содержат, например, сайты рестрикции *BsmBI*, и их можно получить посредством синтеза ДНК или амплификации методом ПЦР. Был разработан набор комплементарных, небеспорядочных четырехнуклеотидных "липких" концов, фланкированных сайтами рестрикции *BsmBI*, для обеспечения бесшовной и однонаправленной сборки элементов ДНК в рациональные кассеты экспрессии (фигура 1B).

[141] Концепция ModVec была впервые протестирована путем сборки 11-модульных бицистронных конструкций размером 11 т. о. в реакции GG объемом 15 мкл (фигура 1C). Было собрано 19 конструкций, и для проверки последовательности инокулировали по шесть колоний для каждой конструкции. Все 19 бицистронных конструкций были успешно собраны в ходе одного раунда клонирования. Более 95% (109/114) клонов содержали правильно собранные векторы (фигура 1D). Из данных результатов видно, что ModVec является эффективным подходом для сборки больших сложных фрагментов ДНК, когда совместимы наборы "липких" концов в реакции GG (т. е. не являются беспорядочными в отношении друг друга). 95% эффективность правильной сборки, полученная с помощью оптимизированной ModVec, делает возможной создание библиотеки векторов экспрессии в одной пробирке. Функционально схожие элементы вектора, которые содержали одинаковые "липкие" концы, смешивали для получения "библиотек модулей" (фигура 1E). В качестве демонстрации библиотек для конструирования векторов в одной пробирке два разных промотора, управляющих 12

кодирующими последовательностями (CDS 1 - CDS 12), объединяли с тремя разными промоторами, управляющими экспрессией CDS 13, что посредством комбинаторной сборки позволяло получить 72 разных вектора экспрессии. Миниатюризованная реакция GG была организована с использованием устройства для манипуляции с жидкостями ECHO, что, как было показано, повышает эффективность сборки GG (Kanigowska et al., 2016). По 80 нл каждой из 25 частей смешивали в 2,5 мкл жидкости для реакции сборки GG. После 30 циклов сборки в реакции GG весь объем вводили методом трансформации в клетки, а клетки высевали. После инкубации в течение ночи на агаровой пластинке выросло около 1000 колоний, 350 из которых были перенесены посредством инокуляции в жидкую культуру для проверки последовательности методом секвенирования по Сэнгеру. В общей сложности было получено 312 контигов, которые соответствовали 69 из 72 возможных векторов (фигура 1F).

[142] Из данных результатов видно, что ModVec является эффективной платформой клонирования для HT-сборки сложных векторов экспрессии и библиотеки конфигураций экспрессии. Таким образом, теперь стало возможным улучшить выработку и качество продуктов в виде bsAb посредством высокоэффективного скрининга конфигураций векторов для выявления конфигураций векторов, которые характеризуются сбалансированным уровнем экспрессии полипептидной цепи в bsAb. Затем систему клонирования ModVec использовали для изучения различных стратегий модуляции транскрипции с целью улучшения выработки и качества продуктов в виде сложных многоцепочечных полиспецифических антител, получаемых рекомбинантным методом в клетках млекопитающих.

[143] ***Влияние промоторов на выработку двухцепочечных моноклональных антител (mAb) и симметричных биспецифических антител (bsAb)***

[144] Тестировали влияние промоторов на выработку симметричных двухцепочечных молекул А, В и С (фигуры 2B-2D). Молекула А представляет собой каноническое моноклональное антитело, молекулы В и С представляют собой симметричные биспецифические антитела. Были получены бицистронные векторы, в которых гены LC и HC для каждой молекулы управлялись одним и тем же промотором (промотором 1 или промотором 2) (фигура 2A). Промотор 1 и промотор 2 представляют собой производные от промотора huCMV, но промотор 2 характеризуется более высокой активностью, чем промотор 1. Для эффективной интеграции в геном и быстрого получения стабильных пулов клеток применяли вектор экспрессии на основе транспозазы *piggyBac*. Внутренние наблюдения позволяют предположить, что векторы экспрессии на основе транспозазы *piggyBac* обладают очень высокой репродуцируемостью, поэтому для каждой конфигурации получали один стабильный пул клеток яичников китайского хомячка (CHO). Независимо от типа молекулы, пулы, трансфицированные конструкциями с применением промотора 2, характеризовались более высоким выходом после аффинной хроматографии с белком А (ProA), чем пулы, трансфицированные конструкциями с применением промотора 1 (фигуры 2B-2D). Выход молекулы А при очистке с ProA

составлял 403 мг/л в пуле, трансфицированных конструкциями с применением промотора 1, и 521 мг/л в пуле, трансфицированных конструкциями с применением промотора 2, что означает выход на 30% больше при переключении с промотора 1 на промотор 2. Более выраженное увеличение наблюдали для молекул В и С, где переключение с промотора 1 на промотор 2 приводило к увеличению выхода при очистке с ProA на 83% и 100% соответственно. Характеристики восстановления и роста клеток также были сопоставимы среди пулов, трансфицированных всеми конструкциями. Для данных трех симметричных молекул одновременное изменение активности промотора генов как HC, так и LC не оказывало значимого влияния на качество продуктов по результатам измерения процента содержания основных разновидностей молекул в эксклюзионной хроматографии (% основного пика SEC, фигуры 2B-2D, средняя панель). В результате при применении промотора 2 было достигнуто увеличение на 28%, 122% и 96% окончательного выхода молекул А, В, С. Более резкое увеличение выработки наблюдали для молекул, которые изначально характеризовались более низкими выходами (молекул В и С). Из результатов видно, что модуляция активности промотора является эффективным подходом для повышения выработки mAb и bsAb, и что ее можно наиболее эффективно применять в отношении трудноэкспрессируемых белков.

[145] Затем стратегии конструирования векторов расширяли с целью повышения как выработки, так и качества продуктов в виде гетеро-IgG - более сложного четырехцепочечного bsAb.

[146] Улучшение качества продуктов и выработки для гетеро-IgG посредством баланса цепей

[147] Гетеро-IgG является наиболее распространенным биспецифическим форматом по причине его антителоподобной структуры (Labrijn et al., 2019), но его также сложнее получать, чем mAb, по причине возможности появления множественных примесей, производных от продукта, вызванных неправильным спариванием LC-HC и HC-HC (Brinkmann & Kontermann, 2017). Многие стратегии конструирования белков, такие как мутация с заменой пар заряженных аминокислот (CPM) (Dillon et al., 2017; Gunasekaran et al., 2010), "выступы-во-впадины" (KiH) (Ridgway et al., 1996), сконструированный домен с заменой нитей (SEED) (Davis et al., 2010) и общая легкая цепь (Krah et al., 2017; Shiraiwa et al., 2019), использовали для содействия правильному спариванию цепей и предотвращению неправильного спаривания цепей, тем самым улучшая возможность получения гетеро-IgG. Однако различия в уровне экспрессии каждой полипептидной цепи, входящей в состав гетеро-IgG в клетках млекопитающих, могут привести к образованию связанных с продуктом примесей и снизить выход правильных bsAb. Например, в стабильных пулах CHO молекулы гетеро-IgG-D, которые были трансфицированы принадлежащими заявителю стандартными векторами экспрессии, наблюдали до 60-кратного увеличение продуцирования LC1 по сравнению с LC2 и секрецию соответственно больших количеств неправильно спаренных разновидностей молекул с двумя копиями LC1 (Guo et al., 2021). Была выдвинута

гипотеза, что выход правильно спаренных гетеро-IgG в сравнении с неправильно спаренными примесями, связанными с продуктом, можно увеличить путем балансировки экспрессии каждой полипептидной цепи ближе к соотношению 1:1:1:1 посредством конструирования векторов. По результатам успеха авторов изобретения с двухцепочечными молекулами был принят подход конструирования векторов для балансировки экспрессии полипептидных цепей с применением промоторов с различной активностью.

[148] С помощью ModVec были получены 11 конфигураций векторов для гетеро-IgG-D. Каждая конфигурация вектора содержала два котрансфицируемых бицистронных вектора, при этом CDS двух полипептидов в данном векторе управлялись либо одними и теми же, либо разными промоторами (промотор 1, промотор 2 и промотор 3, фигура 5A). В некоторых конфигурациях векторов положения кодирующих последовательностей LC или HC были изменены местами внутри бицистронной кассеты (фигура 5A, конфигурация 2). Проводили 4 трансфекции на каждую конфигурацию вектора в клетках CHO и в общей сложности получали 44 пула. Стабильные пулы получали путем случайной интеграции в геном с метаболическим отбором. Пулы из различных конфигураций векторов имели разные показатели восстановления: из 44 трансфицированных пулов 15 не восстановились. Пулы, которые не смогли восстановиться, зачастую были трансфицированы конфигурациями, в которых LC1/LC2 и HC1/HC2 находились в одном и том же векторе (конфигурация 2 на фигуре 5A). Помимо различий в восстановлении стабильных пулов CHO библиотека конфигураций векторов характеризовалась широким диапазоном выхода после аффинной очистки с ProA (фигура 5B), качества продуктов, измеренного по соотношению LC1:LC2 (фигура 5C), и процента пика для основных разновидностей молекул (% MP) в микрокапиллярном электрофорезе в невозстанавливающих условиях (nrMCE) и эксклюзионной хроматографии (SEC) (фигура 5D).

[149] Для масштабированного получения и дополнительного анализа были выбраны три пула: два пула характеризовались соотношением LC1:LC2, близким к идеальному значению, равному 1 (фигура 3B, пулы 910 и 911), а один пул характеризовался искаженным соотношением $LC1:LC2 > 5$ (фигура 3B, пул 905). Конфигурации векторов, использованные для получения данных стабильных пулов, проиллюстрированы на фигуре 3A. Кондиционированную среду данных трех масштабированных пулов очищали с помощью аффинного захвата с ProA с последующим катионным обменом (CEX) с применением смолы на основе сефарозы SP. Все три пула характеризовались схожим выходом при очистке с ProA (фигура 3C). Степень извлечения основных разновидностей молекул гетеро-IgG-D из пула 905 после катионного обмена составляла 12%, тогда как в пулах 910 и 911 степень извлечения составляла 46% и 40% соответственно (фигура 3D). Следовательно, выход правильно собранного гетеро-IgG-D после катионного обмена был выше для пулов 910 и 911 (115 и 75 мг/мл соответственно), чем для пула 905 (34 мг/л) (фигура 3E).

[150] Помимо получения различных относительных уровней основных разновидностей молекул, пулы, трансфицированные различными конфигурациями векторов, также характеризовались разным профилем связанных с продуктом примесей, как показано на катионообменных хроматограммах (фигуры 3F-3H). Основная разновидность молекул в очищенных с помощью ProA образцах пула 905 была выявлена с помощью масс-спектрометрии как неправильно спаренный гетеротетрамер с 2 копиями LC1 (2x LC1, т. е. LC1:HC1:LC1:HC2, основной пик на фигуре 3F), в то время как правильно собранный гетеро-IgG-D был второстепенной разновидностью молекул (второстепенный пик на фигуре 3F). Это поясняет низкий выход после очистки и степень извлечения для пула 905 (фигуры 3D-3E) и согласуется с высоким соотношением LC1:LC2 (фигура 3B). Это также указывает на то, что уровень экспрессии LC1 и LC2 был несбалансированным при использовании промотора 1 для управления экспрессией обоих LC. Не наблюдались антителоподобные гомодимеры из пула 905, которые могли бы возникнуть в результате неправильного спаривания HC1:HC2; основной проблемой пула СНО, трансфицированного данной конфигурацией вектора, было неправильное спаривание LC-НС. При переключении с промотора, управляющего как LC1, так и LC2, на промотор 2 (пул 910, фигура 3A) 2x LC1 становился второстепенной разновидностью молекул, а правильно собранный гетеро-IgG-D становился основной разновидностью молекул (фигура 3G) в материалах, очищенных с ProA, что приводило к более высокому извлечению основных разновидностей молекул после катионного обмена (фигуры 3D-3E). Интересно, что замена обоих промоторов LC на промотор 2 приводит к балансировке экспрессии LC1 и LC2. Ранее авторами настоящего изобретения сообщалось, что LC2 секретировался клетками млекопитающих в низких количествах (Guo et al., 2021). Возможно, что переключение обоих промоторов LC на промотор 2 оказывало большее влияние на "дефицитный по экспрессии" LC2, что приводило к более сбалансированному соотношению LC1:LC2. Авторы настоящего изобретения трансфицировали > 10 повторностей пулов СНО посредством промотора, использованного для получения пула 905, и постоянно наблюдали высокое соотношение LC1:LC2 (5-60) и высокие уровни разновидностей молекул 2x LC1. Более того, в пуле 911 при изменении промотора LC1 на промотор 3, который характеризуется промежуточной активностью транскрипции между промотором 1 и промотором 2, в то время как промотор LC2 оставался неизменным (промотор 2), правильно собранный гетеро-IgG-D оставался основной разновидностью молекулы. Однако преобладающая примесь изменилась на половину антитела 1 (LC1:HC1) вместо 2x LC1 (фигура 3H). Изменение основной примеси с 2x LC1 в пулах 905 и 910 на половину антитела 1 в пуле 911 может быть следствием изменений в промоторах гена LC, влияющих на экспрессию генов LC и HC с бицистронных кассет.

[151] ***Конструирование высокоэффективных векторов для асимметричной трехцепочечной триспецифической молекулы на основе гетеро-IgG***

[152] Затем тестировали, можно ли сконструировать высокоэффективные векторы для получения сложного трехцепочечного асимметричного триспецифического антитела E

(tsAb E). В тестах на транзиентную трансфекцию tsAb E вырабатывался в титрах, которые были в 5-10 раз ниже ожидаемых уровней для mAb. Вдохновленные успехом скрининга конфигурации из 11 векторов для гетеро-IgG, было решено применить подход конструирования векторов для повышения выработки tsAb E, но при этом включить большее количество конфигураций векторов в конструкцию библиотеки. Было высказано предположение, что тестирование большего количества стратегий модуляции экспрессии полипептидных цепей увеличит вероятность нахождения конфигураций со стехиометрически оптимальными соотношениями для всех трех полипептидных цепей асимметричного tsAb E. Помимо промоторов с различной активностью варьировали дозу гена (количество копий гена LC), а также цистронное расположение генов LC и HC в качестве переменных в конструкции библиотеки. Окончательная библиотека содержала 189 конфигураций векторов, которые включали 27 моноцистронных векторов с одной копией гена LC (одинарный LC), 81 бицистронный вектор с двумя копиями гена LC в первом цистроне каждого вектора (двойной LC A) и 81 бицистронный вектор с двумя копиями гена LC во втором цистроне каждого вектора (двойной LC B) (фигура 4A). Для управления экспрессией каждой полипептидной цепи комбинаторным образом авторы настоящего изобретения использовали три промотора (промотор 1, промотор 2 и промотор 3); поэтому в данной конструкции библиотеки векторов гены в бицистронном векторе могли необязательно управляться одними и теми же или разными промоторами.

[153] Клетки СНО подвергали трансфекции в 24-луночных культуральных планшетах с применением запатентованной системы для высокоэффективной экспрессии, которая основана на опосредованной транспозазой интеграции с метаболическим отбором. Для каждой конфигурации вектора получали один пул. Все пулы характеризовались сопоставимыми профилями восстановления и роста (данные не представлены), что является обычным для данной системы экспрессии. Как показано на фигура 4B, имела место разница на порядок в выработке данного tsAb в зависимости от конфигурации вектора; выход всех пулов при очистке с ProA варьировал в диапазоне от 20 мг/л до 220 мг/л, что наглядно демонстрировало степень влияния конфигурации вектора на выход. Пулы, трансфицированные конфигурациями двойного LC A ($117,2 \pm 43,4$ мг/л) или двойного LC B ($112,6 \pm 39,2$ мг/л), характеризовались более высоким средним выходом при очистке с ProA, чем пулы, трансфицированные конфигурацией одинарного LC ($40,7 \pm 20,73$ мг/л) (*t-критерий* Стьюдента, двойной LC A относительно одинарного LC $p < 0,0001$, двойной LC B относительно одинарного LC $p < 0,0001$; значения выражены как среднее значение \pm стандартное отклонение). Пулы двойного LC A и пулы двойного LC B характеризовались сопоставимым выходом при очистке с ProA (*t-критерий* Стьюдента, двойной LC относительно двойного LC B, $p=0,48$). Предварительную оценку качества продукта производили путем проведения микрокапиллярного электрофореза в невосстанавливающих условиях (nrMCE) на каждом пуле при очистке с ProA; средние процентные значения пика для основных разновидностей молекул (% MP) между пулами, полученными из данных трех групп

конфигураций векторов, были сопоставимы (фигура 4С, двойной LC А МР: $75,6 \pm 14,9\%$, двойной LC В МР: $72,4 \pm 16,6\%$, одинарный LC МР: $71,7 \pm 21,6\%$; двойной LC А относительно двойного LC В $p=0,19$, двойной LC А относительно одинарного LC $p=0,34$, двойной LC В относительно одинарного LC $p=0,88$, *t*-критерий Стьюдента). В результате, как показано на фигура 4D, средний выход правильно собранного tsAb Е был выше в пулах, полученных из конфигураций двойного LC А ($88,7 \pm 33,6$ мг/л) или двойного LC В ($81,4 \pm 35,6$ мг/л), чем в пулах, полученных из конфигураций одинарного LC ($28,5 \pm 16,8$ мг/л). Выход правильных разновидностей молекул был сопоставим среди пулов, полученных из двойного LC А и двойного LC В (двойной LC А относительно двойного LC В $p=0,48$, двойной LC А относительно двойного LC $p < 0,0001$, двойной LC В относительно двойного LC $p < 0,0001$, *t*-критерий Стьюдента). Из данных результатов видно влияние количества копий гена LC на выработку сложных терапевтических молекул. Это согласуется с предыдущими выводами о том, что доза гена, определяемая количеством копий гена, может оказывать большое влияние на выработку и что избыток LC в целом полезен для выработки (Gonzalez et al., 2002).

[154] Поскольку показатели группы двойного LC А и группы двойного LC В были сопоставимы, сравнивали показатели отдельных конфигураций из данных двух групп с целью понимания влияния положения гена на выработку и качество продуктов. Для дополнительного анализа были выбраны конфигурации DA2 (конфигурация двойного LC А с промотором 2, управляющим экспрессией всех полипептидных цепей) и DB2 (конфигурация двойного LC В с промотором 2, управляющим экспрессией всех полипептидных цепей) (фигура 6А). В DA2 гены LC находились в первом цистроне, а гены HC находились во втором цистроне, в то время как в DB2 положения генов LC и генов HC были изменены местами. Как показано на фигуре 6В, выход пула DA2 при очистке с ProA составлял 155 мг/л. Изменение местами положений генов LC и HC путем помещения генов HC в первый цистрон, а генов LC во второй цистрон снижало общую выработку. Выход пула DB2 при очистке с ProA составлял 113 мг/л. Это согласуется с эффектом положения генов в мультицистронных векторах экспрессии, где уровни экспрессии генов после первого гена, как правило, снижены относительно уровня экспрессии первого гена (Eszterhas et al., 2002; Patel et al., 2021). В данном случае снижение экспрессии генов LC при одновременном повышении экспрессии генов HC приводит к общему снижению выработки, что снова соответствует ожиданию, что высокий уровень экспрессии LC полезен для общей выработки (Gonzalez et al., 2002). Качество продуктов, измеренное с помощью МР и пре-МР по nrMCE, было сопоставимым среди образцов при очистке с ProA, очищенных из пулов DA2 и DB2 (МР по nrMCE, 80% для DA2 и 81% для DB2; пре-МР по nrMCE, 18% для DA2 и 19% для DB2; фигура 6С). Масс-спектрометрия подтверждала, что МР по nrMCE был правильно собранным tsAb Е, а пре-МР по nrMCE представляли собой половину mAb1 (LC:HC1) и половину mAb2 (LC:HC2, фигуры 6Е и 6F).

[155] Несмотря на то, что уровни пре-МР по nrMCE были сопоставимы среди

образцов при очистке с ProA из пулов DA2 и DB2, относительное содержание половины mAb1 и mAb2 различалось (фигуры 6E и 6F). Относительная площадь пика половины mAb1 составляла 10,1, а половины mAb2 составляла 7,4 на электрофореграмме ngMSE очищенных с помощью ProA образцов из пула DA2 (фигура 6E). Относительная площадь пика для половины mAb1 составляла 5,8, а для половины mAb2 составляла 9,8 для очищенных с помощью ProA образцов из пула DB2 (фигура 6F). Было рассчитано относительное соотношение пиков половина mAb1:половина mAb2, и результаты представлены на фигуре 6G. Относительное соотношение пиков половина mAb1:половина mAb2 составляло 1,4 в очищенных с помощью ProA образцах из пула DA2 и 0,6 в очищенных с помощью ProA образцах из пула DB2. Данные результаты дополнительно демонстрируют, что положение гена LC или HC в векторах экспрессии может значимо влиять как на выработку основных разновидностей молекул, так и на профиль примесей, полученных из клеток СНО. Данные результаты также подтверждают гипотезу авторов изобретения о том, что объединение в комбинацию промоторов с различной активностью, количества копий гена и положения гена является эффективным подходом для модуляции уровней экспрессии полипептидной цепи и может повысить вероятность выявления конфигураций векторов, которые обеспечивают высокую выработку и качество продуктов для многоцепочечных белковых терапевтических средств.

[156] По результатам скрининга библиотеки векторов были отобраны лучшие конфигурации векторов для перехода к разработке линии клеток (CLD) для tsAb E и данные векторы были подвергнуты сравнению с вектором стандартной платформы (показано на фигуре 4H). В стандартной конфигурации вектора использован промотор 1, который управляет экспрессией всех полипептидных цепей, и обеспечивает объединение в комбинацию бицистронного вектора с геном LC в первом цистроне и HC1 во втором цистроне с моноцистронным вектором, содержащим HC2. В отличие от этого, оптимальная конфигурация, выявленная с помощью библиотеки векторов, состояла из двух бицистронных векторов, каждый из которых экспрессировал копию гена LC в первом цистроне с либо HC1, либо HC2 во втором цистроне, при этом промотор 2 управлял экспрессией обоих генов LC, а промотор 1 управлял экспрессией генов HC1 и HC2 (фигура 4H). Стабильные пулы СНО, полученные из стандартной конфигурации, обеспечивали получение tsAb E с выходом при очистке с ProA, составляющим 119 мг/л, тогда как оптимальная конфигурация из библиотеки векторов обеспечивала получение 267 мг/л после очистки с ProA, что приводило к увеличению на 124% (фигура 4E). Качество продуктов, измеряемое с помощью ngMSE, было сопоставимым между очищенными с ProA образцами из обоих пулов (фигура 4F). В результате окончательный выход правильно собранного tsAb E составлял 114 мг/л в стандартном пуле и 257 мг/л в оптимальном пуле (фигура 4G). Оптимальная конфигурация вектора, выявленная посредством скрининга библиотеки конфигураций векторов, позволяла на 125% увеличить выработку по сравнению со стандартной конфигурацией вектора авторов настоящего изобретения для данного перспективного и сложного для экспрессии

терапевтического средства на основе tsAb.

[157] **Литературные источники**

[158] Bhoskar, P., Belongia, B., Smith, R., Yoon, S., Carter, T., & Xu, J. (2013). Free light chain content in culture media reflects recombinant monoclonal antibody productivity and quality. *Biotechnol Prog*, 29(5), 1131-1139. <https://doi.org/10.1002/btpr.1767>

[159] Blanco, N., Williams, A. J., Tang, D., Zhan, D., Misaghi, S., Kelley, R. F., & Simmons, L. C. (2020). Tailoring translational strength using Kozak sequence variants improves bispecific antibody assembly and reduces product-related impurities in CHO cells. *Biotechnol Bioeng*, 117(7), 1946-1960. <https://doi.org/10.1002/bit.27347>

[160] Brinkmann, U., & Kontermann, R. E. (2017). The making of bispecific antibodies. *MAbs*, 9(2), 182-212. <https://doi.org/10.1080/19420862.2016.1268307>

[161] Campuzano, I. D. G., Robinson, J. H., Hui, J. O., Shi, S. D., Netirojjanakul, C., Nshanian, M., . . . Bern, M. (2019). Native and Denaturing MS Protein Deconvolution for Biopharma: Monoclonal Antibodies and Antibody-Drug Conjugates to Polydisperse Membrane Proteins and Beyond. *Anal Chem*, 91(15), 9472-9480. <https://doi.org/10.1021/acs.analchem.9b00062>

[162] Carver, J., Ng, D., Zhou, M., Ko, P., Zhan, D., Yim, M., . . . Hu, Z. (2020). Maximizing antibody production in a targeted integration host by optimization of subunit gene dosage and position. *Biotechnol Prog*, 36(4), e2967. <https://doi.org/10.1002/btpr.2967>

[163] Cazier, A. P., & Blazek, J. (2021). Advances in promoter engineering: Novel applications and predefined transcriptional control. *Biotechnol J*, 16(10), e2100239. <https://doi.org/10.1002/biot.202100239>

[164] Davis, J. H., Aperlo, C., Li, Y., Kurosawa, E., Lan, Y., Lo, K. M., & Huston, J. S. (2010). SEEDbodies: fusion proteins based on strand-exchange engineered domain (SEED) CH3 heterodimers in an Fc analogue platform for asymmetric binders or immunofusions and bispecific antibodies. *Protein Eng Des Sel*, 23(4), 195-202. <https://doi.org/10.1093/protein/gzp094>

[165] Deshaies, R. J. (2020). Multispecific drugs herald a new era of biopharmaceutical innovation. *Nature*, 580(7803), 329-338. <https://doi.org/10.1038/s41586-020-2168-1>

[166] Dillon, M., Yin, Y., Zhou, J., McCarty, L., Ellerman, D., Slaga, D., . . . Carter, P. J. (2017). Efficient production of bispecific IgG of different isotypes and species of origin in single mammalian cells. *MAbs*, 9(2), 213-230. <https://doi.org/10.1080/19420862.2016.1267089>

[167] Dong, E., Lam, C., Tang, D., Louie, S., Yim, M., Williams, A. J., . . . Misaghi, S. (2021). Concurrent transfection of randomized transgene configurations into targeted integration CHO host is an advantageous and cost-effective method for expression of complex molecules. *Biotechnol J*, 16(4), e2000230. <https://doi.org/10.1002/biot.202000230>

[168] Eisenhut, P., Mebrahtu, A., Moradi Barzadd, M., Thalen, N., Klanert, G., Weinguny, M., . . . Rockberg, J. (2020). Systematic use of synthetic 5'-UTR RNA structures to tune protein translation improves yield and quality of complex proteins in mammalian cell factories. *Nucleic Acids Res*, 48(20), e119. <https://doi.org/10.1093/nar/gkaa847>

- [169] Engler, C., Kandzia, R., & Marillonnet, S. (2008). A one pot, one step, precision cloning method with high throughput capability. *PLoS One*, 3(11), e3647. <https://doi.org/10.1371/journal.pone.0003647>
- [170] Eszterhas, S. K., Bouhassira, E. E., Martin, D. I., & Fiering, S. (2002). Transcriptional interference by independently regulated genes occurs in any relative arrangement of the genes and is influenced by chromosomal integration position. *Mol Cell Biol*, 22(2), 469-479. <https://doi.org/10.1128/MCB.22.2.469-479.2002>
- [171] Fan, G., Wang, Z., Hao, M., & Li, J. (2015). Bispecific antibodies and their applications. *J Hematol Oncol*, 8, 130. <https://doi.org/10.1186/s13045-015-0227-0>
- [172] Gong, D., Riley, T. P., Bzymek, K. P., Correia, A. R., Li, D., Spahr, C., . . . Garces, F. (2021). Rational selection of building blocks for the assembly of bispecific antibodies. *MAbs*, 13(1), 1870058. <https://doi.org/10.1080/19420862.2020.1870058>
- [173] Gonz lez, R., Andrews, B. A., & Asenjo, J. A. (2002). Kinetic model of BiP- and PDI-mediated protein folding and assembly. *J Theor Biol*, 214(4), 529-537. <https://doi.org/10.1006/jtbi.2001.2478>
- [174] Gunasekaran, K., Pentony, M., Shen, M., Garrett, L., Forte, C., Woodward, A., . . . Yan, W. (2010). Enhancing antibody Fc heterodimer formation through electrostatic steering effects: applications to bispecific molecules and monovalent IgG. *J Biol Chem*, 285(25), 19637-19646. <https://doi.org/10.1074/jbc.M110.117382>
- [175] Guo, C., Chen, F., Xiao, Q., Catterall, H. B., Robinson, J. H., Wang, Z., . . . Hubert, R. (2021). Expression liabilities in a four-chain bispecific molecule. *Biotechnol Bioeng*, 118(10), 3744-3759. <https://doi.org/10.1002/bit.27850>
- [176] Ha, J. H., Kim, J. E., & Kim, Y. S. (2016). Immunoglobulin Fc Heterodimer Platform Technology: From Design to Applications in Therapeutic Antibodies and Proteins. *Front Immunol*, 7, 394. <https://doi.org/10.3389/fimmu.2016.00394>
- [177] Kaneyoshi, K., Kuroda, K., Uchiyama, K., Onitsuka, M., Yamano-Adachi, N., Koga, Y., & Omasa, T. (2019). Secretion analysis of intracellular "difficult-to-express" immunoglobulin G (IgG) in Chinese hamster ovary (CHO) cells. *Cytotechnology*, 71(1), 305-316. <https://doi.org/10.1007/s10616-018-0286-5>
- [178] Kanigowska, P., Shen, Y., Zheng, Y., Rosser, S., & Cai, Y. (2016). Smart DNA Fabrication Using Sound Waves: Applying Acoustic Dispensing Technologies to Synthetic Biology. *J Lab Autom*, 21(1), 49-56. <https://doi.org/10.1177/2211068215593754>
- [179] Krah, S., Schroter, C., Eller, C., Rhiel, L., Rasche, N., Beck, J., . . . Becker, S. (2017). Generation of human bispecific common light chain antibodies by combining animal immunization and yeast display. *Protein Eng Des Sel*, 30(4), 291-301. <https://doi.org/10.1093/protein/gzw077>
- [180] Kwaks, T. H., & Otte, A. P. (2006). Employing epigenetics to augment the expression of therapeutic proteins in mammalian cells. *Trends Biotechnol*, 24(3), 137-142. <https://doi.org/10.1016/j.tibtech.2006.01.007>
- [181] Labrijn, A. F., Janmaat, M. L., Reichert, J. M., & Parren, P. (2019). Bispecific

antibodies: a mechanistic review of the pipeline. *Nat Rev Drug Discov*, 18(8), 585-608. <https://doi.org/10.1038/s41573-019-0028-1>

[182] Ma, J., Mo, Y., Tang, M., Shen, J., Qi, Y., Zhao, W., . . . Qian, C. (2021). Bispecific Antibodies: From Research to Clinical Application. *Front Immunol*, 12, 626616. <https://doi.org/10.3389/fimmu.2021.626616>

[183] Magistrelli, G., Pontini, G., Poitevin, Y., Malinge, P., Bourguignon, J., Gauye, F., . . . Fischer, N. (2018). Tuning Relative Polypeptide Expression to Optimize Assembly, Yield and Downstream Processing of Bispecific Antibodies. *Antibodies (Basel)*, 7(3). <https://doi.org/10.3390/antib7030029>

[184] Mathias, S., Wippermann, A., Raab, N., Zeh, N., Handrick, R., Gorr, I., . . . Otte, K. (2020). Unraveling what makes a monoclonal antibody difficult-to-express: From intracellular accumulation to incomplete folding and degradation via ERAD. *Biotechnol Bioeng*, 117(1), 5-16. <https://doi.org/10.1002/bit.27196>

[185] McBurney, M. W., Mai, T., Yang, X., & Jardine, K. (2002). Evidence for repeat-induced gene silencing in cultured Mammalian cells: inactivation of tandem repeats of transfected genes. *Exp Cell Res*, 274(1), 1-8. <https://doi.org/10.1006/excr.2001.5443>

[186] Neves, D., Vos, S., Blank, L. M., & Ebert, B. E. (2019). Pseudomonas mRNA 2.0: Boosting Gene Expression Through Enhanced mRNA Stability and Translational Efficiency. *Front Bioeng Biotechnol*, 7, 458. <https://doi.org/10.3389/fbioe.2019.00458>

[187] Patel, Y. D., Brown, A. J., Zhu, J., Rosignoli, G., Gibson, S. J., Hatton, D., & James, D. C. (2021). Control of Multigene Expression Stoichiometry in Mammalian Cells Using Synthetic Promoters. *ACS Synth Biol*, 10(5), 1155-1165. <https://doi.org/10.1021/acssynbio.0c00643>

[188] Rajendra, Y., Peery, R. B., Houglan, M. D., Barnard, G. C., Wu, X., Fitchett, J. R., . . . Demarest, S. J. (2017). Transient and stable CHO expression, purification and characterization of novel hetero-dimeric bispecific IgG antibodies. *Biotechnol Prog*, 33(2), 469-477. <https://doi.org/10.1002/btpr.2414>

[189] Ridgway, J. B., Presta, L. G., & Carter, P. (1996). 'Knobs-into-holes' engineering of antibody CH3 domains for heavy chain heterodimerization. *Protein Eng*, 9(7), 617-621. <https://doi.org/10.1093/protein/9.7.617>

[190] Romanova, N., & Noll, T. (2018). Engineered and Natural Promoters and Chromatin-Modifying Elements for Recombinant Protein Expression in CHO Cells. *Biotechnol J*, 13(3), e1700232. <https://doi.org/10.1002/biot.201700232>

[191] Schaefer, W., Volger, H. R., Lorenz, S., Imhof-Jung, S., Regula, J. T., Klein, C., & Molhoj, M. (2016). Heavy and light chain pairing of bivalent quadroma and knobs-into-holes antibodies analyzed by UHR-ESI-QTOF mass spectrometry. *MAbs*, 8(1), 49-55. <https://doi.org/10.1080/19420862.2015.1111498>

[192] Schlatter, S., Stansfield, S. H., Dinnis, D. M., Racher, A. J., Birch, J. R., & James, D. C. (2005). On the optimal ratio of heavy to light chain genes for efficient recombinant antibody production by CHO cells. *Biotechnol Prog*, 21(1), 122-133.

<https://doi.org/10.1021/bp049780w>

[193] Shiraiwa, H., Narita, A., Kamata-Sakurai, M., Ishiguro, T., Sano, Y., Hironiwa, N., . . . Igawa, T. (2019). Engineering a bispecific antibody with a common light chain: Identification and optimization of an anti-CD3 epsilon and anti-GPC3 bispecific antibody, ERY974. *Methods*, 154, 10-20. <https://doi.org/10.1016/j.ymeth.2018.10.005>

[194] Spahr, C. S., Daris, M. E., Graham, K. C., Soriano, B. D., Stevens, J. L., & Shi, S. D. (2018). Discovery, characterization, and remediation of a C-terminal Fc-extension in proteins expressed in CHO cells. *MAbs*, 10(8), 1291-1300. <https://doi.org/10.1080/19420862.2018.1511197>

[195] Von Kreudenstein, T. S., Escobar-Cabrera, E., Lario, P. I., D'Angelo, I., Brault, K., Kelly, J., . . . Dixit, S. B. (2013). Improving biophysical properties of a bispecific antibody scaffold to aid developability: quality by molecular design. *MAbs*, 5(5), 646-654. <https://doi.org/10.4161/mabs.25632>

[196] Winters, C., van Wegen, E. E., Daffertshofer, A., & Kwakkel, G. (2015). Generalizability of the Proportional Recovery Model for the Upper Extremity After an Ischemic Stroke. *Neurorehabil Neural Repair*, 29(7), 614-622. <https://doi.org/10.1177/1545968314562115>

[197] Yoo, D., Provchy, J., Park, C., Schulz, C., & Walker, K. (2014). Automated high-throughput protein purification using an AKTApurifier and a CETAC autosampler. *J Chromatogr A*, 1344, 23-30. <https://doi.org/10.1016/j.chroma.2014.04.014>

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ экспрессии многоцепочечного белка, содержащего по меньшей мере две различные полипептидные цепи, включающий

а. обеспечение наличия полинуклеотидных последовательностей, которые кодируют по меньшей мере два различных полипептида, где указанные полинуклеотидные последовательности фланкированы с 5'- и 3'-сторон сайтом расщепления эндонуклеазой рестрикции PIS типа, за которым следует сайт ее распознавания;

б. обеспечение наличия двух или более наборов последовательностей элементов, где каждый набор последовательностей элементов вместе содержит по меньшей мере одну функциональную кассету экспрессии,

где каждая последовательность элемента фланкирована с 5'- и 3'-сторон сайтом расщепления эндонуклеазой рестрикции PIS типа, за которым следует сайт ее распознавания,

где сайты распознавания эндонуклеазой рестрикции PIS типа и сайты расщепления выбраны таким образом, чтобы наборы последовательностей элементов могли быть собраны в функциональную кассету экспрессии;

с. обеспечение наличия по меньшей мере одного исходного вектора-остова, где исходный вектор-остов предусматривает плазмиду, содержащую по меньшей мере один сайт расщепления эндонуклеазой рестрикции PIS типа, за которым следует сайт ее распознавания и последовательности для поддержания вектора в бактериальных клетках;

д. сборку полинуклеотидных последовательностей, которые кодируют по меньшей мере два различных полипептида, с двумя или более наборами последовательностей элементов для получения функциональных кассет экспрессии в исходном векторе-остове с помощью способа, основанного на применении расщепления ферментом рестрикции PIS типа и лигировании посредством сайтов расщепления и "липких" концов, в результате чего получают лигированные векторы, содержащие по меньшей мере две функциональные кассеты экспрессии, которые способны экспрессировать по меньшей мере две различные полипептидные цепи;

е. трансфекцию клеток млекопитающих лигированными векторами;

ф. сравнение уровней полученного предпочтительного продукта и уровней примесей для каждого из по меньшей мере двух полипептидов и

г. выявление оптимального соотношения экспрессии двух полипептидов для трансфицированных клеток млекопитающих и выявление трансфицированных клеток млекопитающих, которые характеризуются оптимальным соотношением экспрессии по меньшей мере двух полипептидных цепей.

2. Способ по п. 1, дополнительно включающий

h. выявление расположения элементов вектора, которое обеспечивает оптимальное соотношение экспрессии по меньшей мере двух полипептидных цепей.

3. Способ по любому из предыдущих пунктов, где последовательности элементов

выбраны из группы, состоящей из контрольной последовательности, инсуляторной последовательности, баркодовой последовательности ДНК, последовательностей праймера, промоторных последовательностей, последовательностей поли-А и последовательностей IRES, а также последовательности селективного маркера млекопитающих.

4. Способ по п. 3, где последовательности элементов содержат по меньшей мере две промоторные последовательности и по меньшей мере две последовательности поли-А.

5. Способ по любому из предыдущих пунктов, где уровни полученного предпочтительного продукта измеряют с помощью способа, выбранного из группы, состоящей из катионообменной хроматографии (в восстанавливающих или невосстанавливающих условиях), масс-спектрометрии или их комбинации.

6. Способ по любому из предыдущих пунктов, где ни один из сайтов расщепления эндонуклеазой рестрикции PIS типа при расщеплении эндонуклеазой рестрикции PIS типа не образует 5'-"липкие" концы из четырех нуклеотидов, выбранные из группы, состоящей из GTAA, TCCA и CACA.

7. Способ по любому из предыдущих пунктов, где сайт расщепления эндонуклеазой рестрикции PIS типа выбран из группы, состоящей из

AGGT, AGTA, ATCA, CAGT, CCAT, GAAT, GAGG, GGCA, GGTC, TAGC, TCTT, GGAG и CCAC.

8. Способ по любому из предыдущих пунктов, где клетки млекопитающих выбраны из группы, состоящей из клеток CHO, клеток CHOK1, DXB-11, DG-44, COS-7, HEK293, ВНК, TM4, CV1, VERO-76, HELA, MDCK, BRL 3A, W138, Hep G2, MMT 060562, клеток TRI, клеток MRC 5, клеток FS4 и клеток миеломы млекопитающих.

9. Способ по любому из предыдущих пунктов, где оптимальное соотношение экспрессии по меньшей мере двух полипептидов выбрано из группы, состоящей из 1:1, 1:2 и 1:3.

10. Способ по п. 1, где многоцепочечный белок содержит первую тяжелую цепь антитела, первую легкую цепь антитела, вторую тяжелую цепь антитела и вторую легкую цепь антитела, где первая тяжелая цепь антитела соединяется с первой легкой цепью антитела для связывания первого антигена или эпитопа, а вторая тяжелая цепь антитела соединяется со второй легкой цепью антитела для связывания второго антигена или эпитопа, где оптимальное соотношение экспрессии первой тяжелой цепи антитела, первой легкой цепи антитела, второй тяжелой цепи антитела и второй легкой цепи антитела составляет 1:1:1:1.

11. Способ по п. 1, где многоцепочечный белок содержит первую тяжелую цепь антитела, вторую тяжелую цепь антитела и общую легкую цепь антитела, где первая тяжелая цепь антитела соединяется с общей легкой цепью антитела для связывания первого антигена или эпитопа, а вторая тяжелая цепь антитела соединяется с общей легкой цепью антитела для связывания второго антигена или эпитопа, где оптимальное соотношение экспрессии первой тяжелой цепи антитела, второй тяжелой цепи антитела и

общей легкой цепи антитела составляет 1:1:2.

12. Способ по п. 1, где многоцепочечный белок содержит тяжелую цепь антитела, первую легкую цепь антитела, модифицированную тяжелую цепь антитела и вторую легкую цепь антитела,

где модифицированная тяжелая цепь антитела содержит от N-конца к С-концу одну из нижеуказанных структур, выбранных из следующей группы:

VH-CH1-связывающий домен (BD)-шарнир-CH2-CH3,

BD-VH-CH1-шарнир-CH2-CH3 и

VH-CH1-шарнир-CH2-CH3-BD,

где BD выбран либо как одноцепочечный Fv (scFv), либо как одноцепочечный Fab (scFab);

где тяжелая цепь антитела соединяется с первой легкой цепью антитела для связывания первого антигена или эпитопа, а VH модифицированной тяжелой цепи антитела соединяется со второй легкой цепью антитела для связывания второго антигена или эпитопа, где BD связывается с третьим антигеном или эпитопом, и

где оптимальное соотношение экспрессии первой тяжелой цепи антитела, первой легкой цепи, модифицированной тяжелой цепи антитела и второй легкой цепи антитела составляет 1:1:1:1.

13. Способ по п. 1, где многоцепочечный белок содержит тяжелую цепь антитела, модифицированную тяжелую цепь антитела и общую легкую цепь антитела,

где модифицированная тяжелая цепь антитела содержит от N-конца к С-концу одну из нижеуказанных структур, выбранных из следующей группы:

VH-CH1-связывающий домен (BD)-шарнир-CH2-CH3,

BD-VH-CH1-шарнир-CH2-CH3 и

VH-CH1-шарнир-CH2-CH3-BD,

где BD выбран либо как одноцепочечный Fv (scFv), либо как одноцепочечный Fab (scFab);

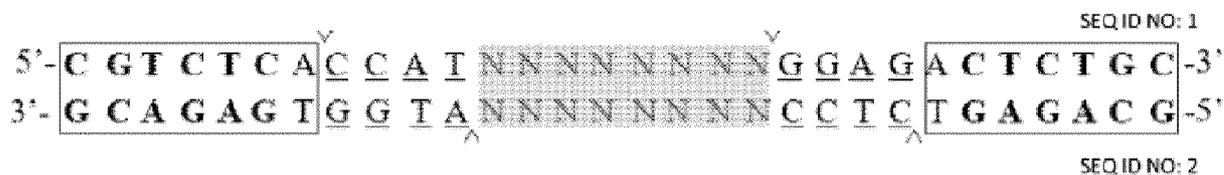
где тяжелая цепь антитела соединяется с общей легкой цепью антитела для связывания первого антигена или эпитопа, а VH модифицированной тяжелой цепи антитела соединяется с общей легкой цепью антитела для связывания первого антигена или эпитопа, где BD связывается со вторым антигеном или эпитопом, и

где оптимальное соотношение экспрессии первой тяжелой цепи антитела, модифицированной тяжелой цепи антитела и общей легкой цепи антитела составляет 1:1:2.

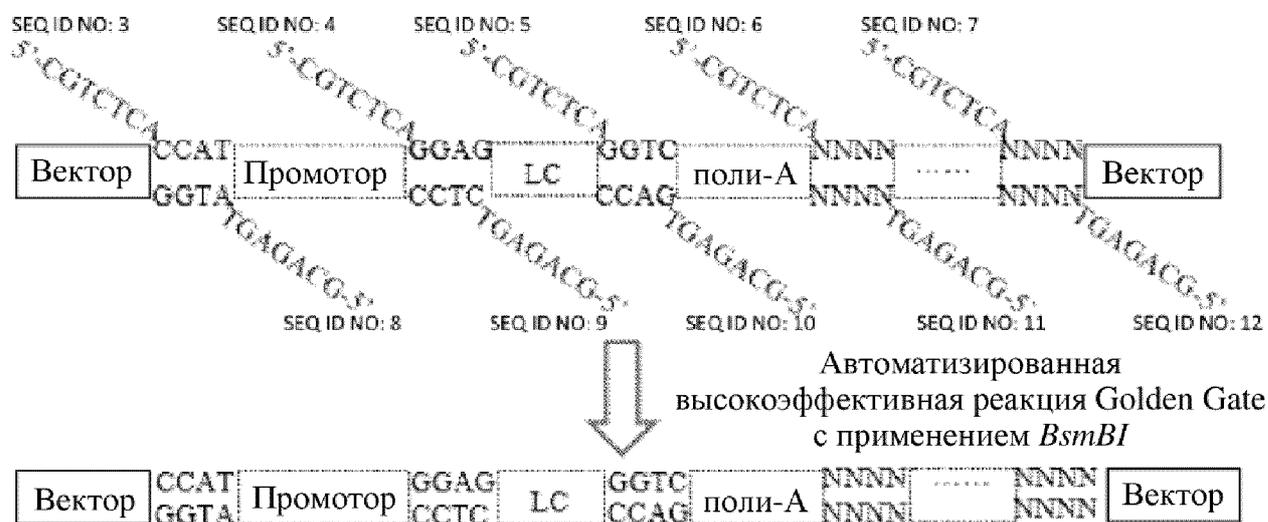
14. Вектор, характеризующийся расположением элементов вектора в соответствии с п. 2, стадия h.

По доверенности

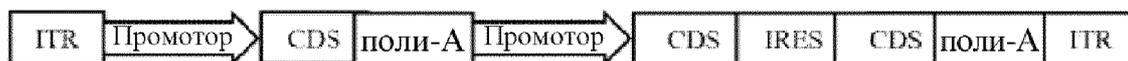
1/11



Фигура 1А



Фигура 1В



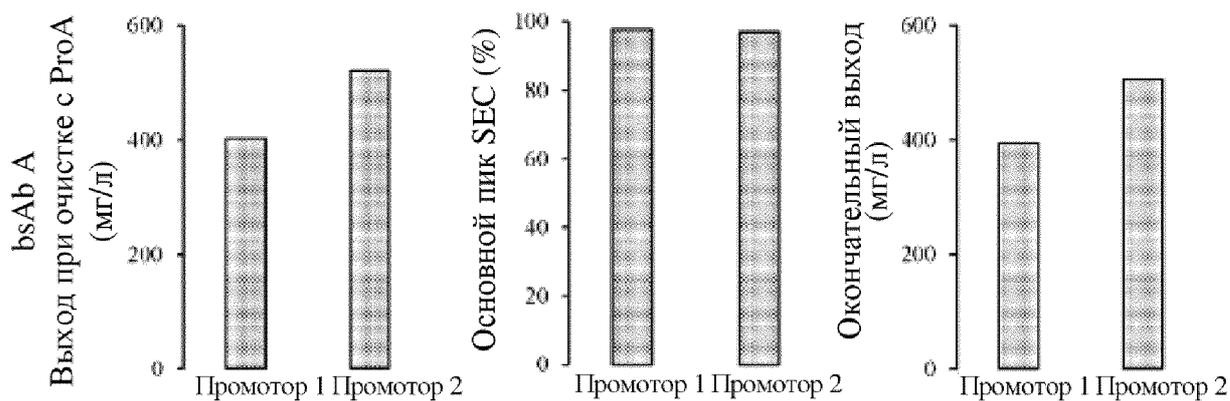
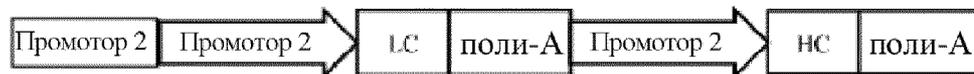
Сборка Golden Gate размером 11 т. о., состоящей из 11 частей

Фигура 1С

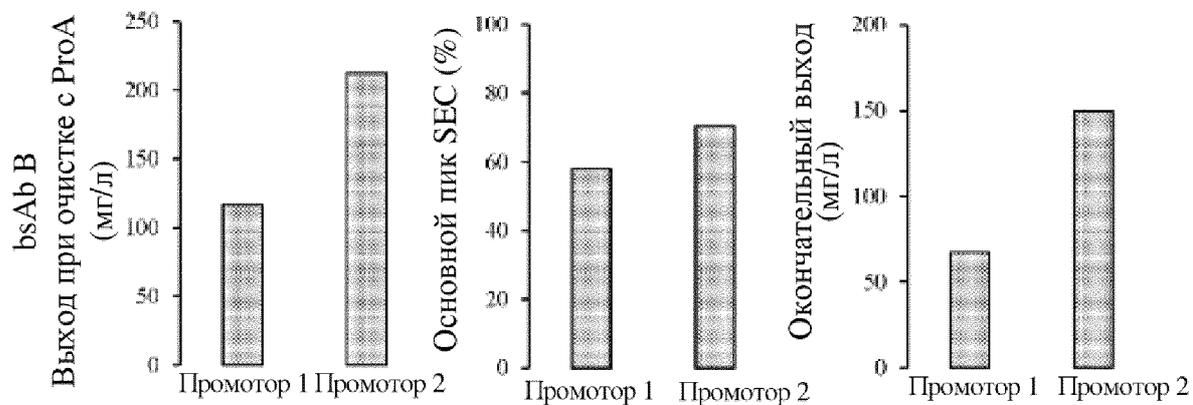
Фигура 2А



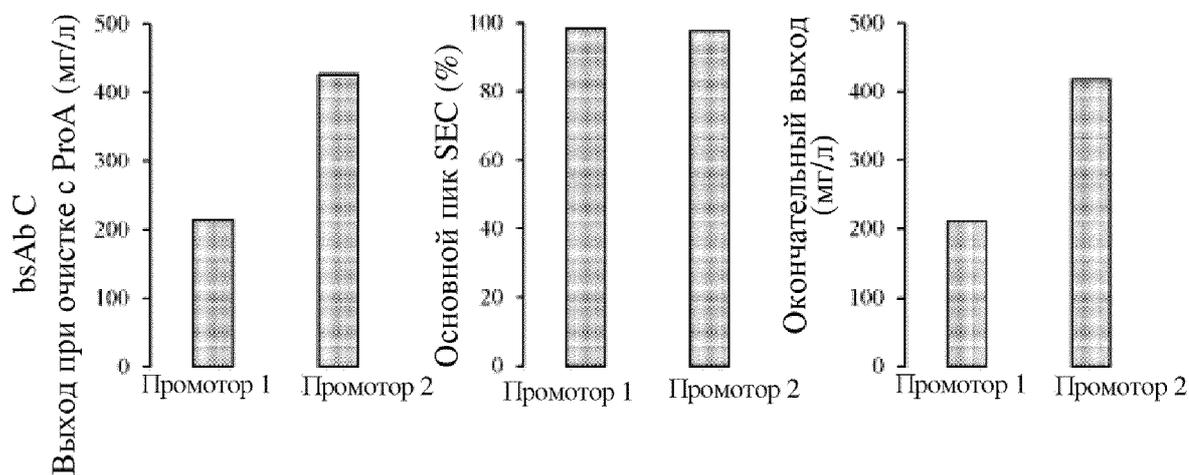
Фигура 2В



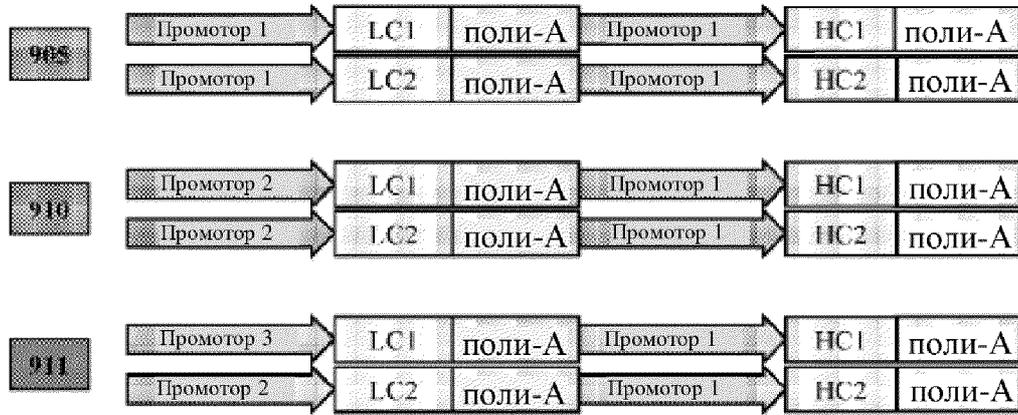
Фигура 2С



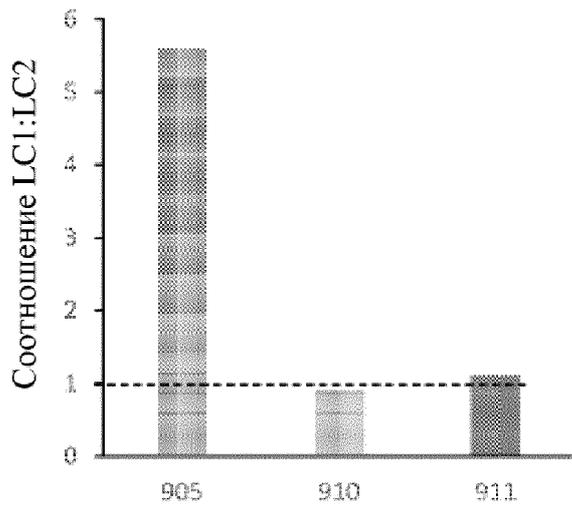
Фигура 2D



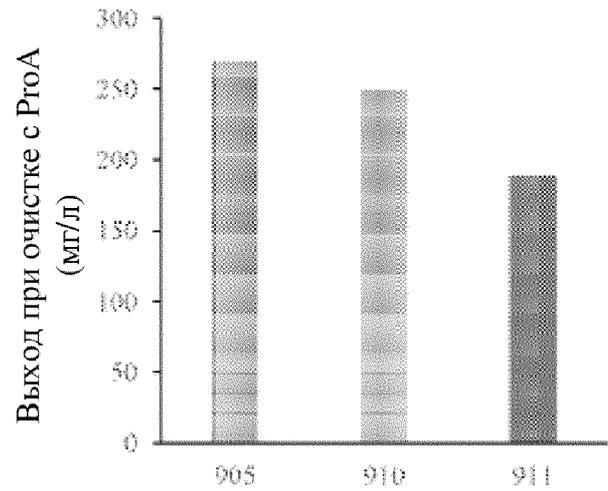
Фигура 3А



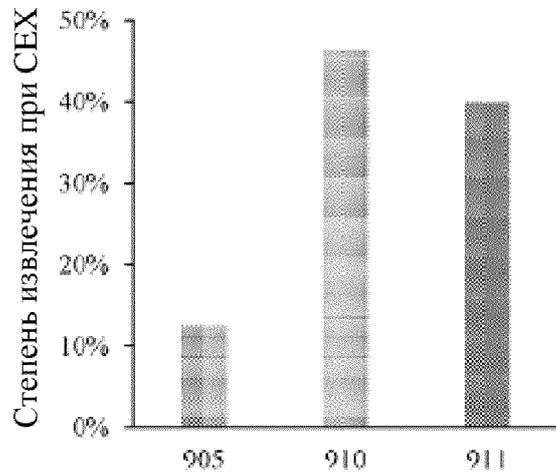
Фигура 3В



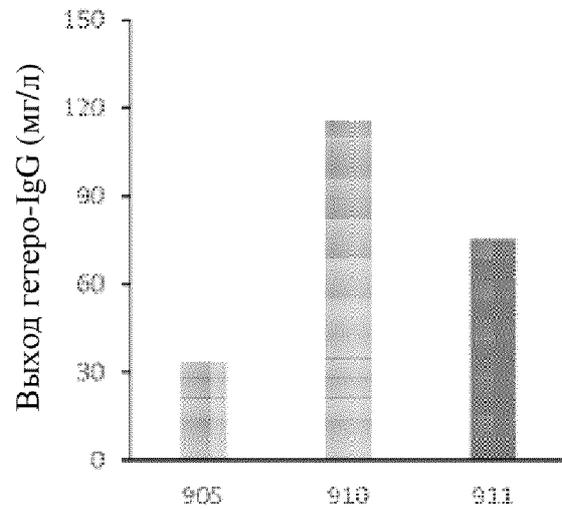
Фигура 3С



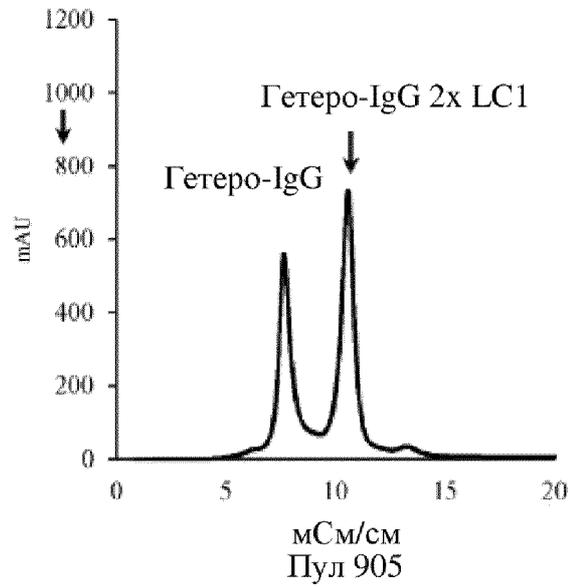
Фигура 3D



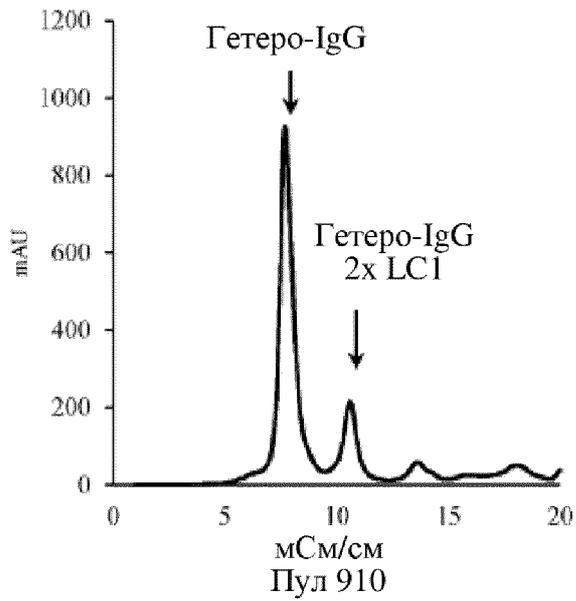
Фигура 3Е



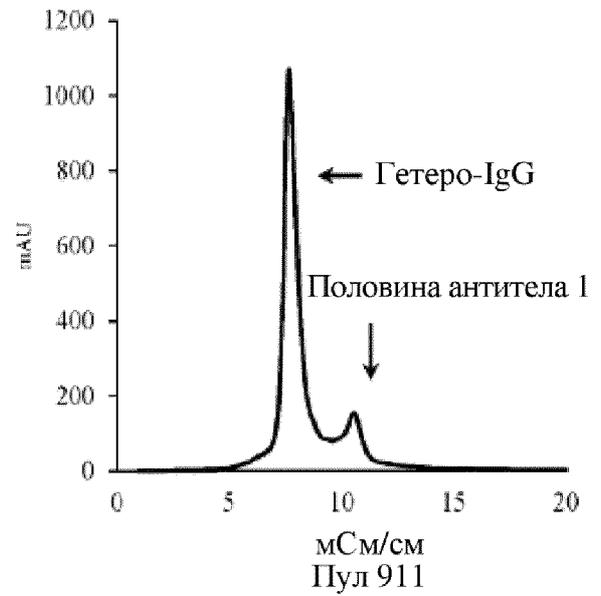
Фигура 3F



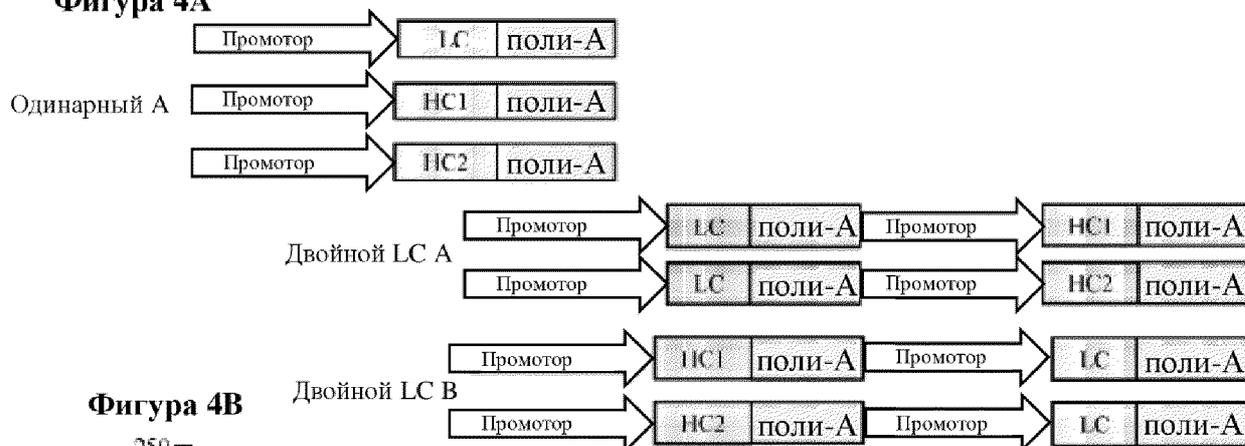
Фигура 3G



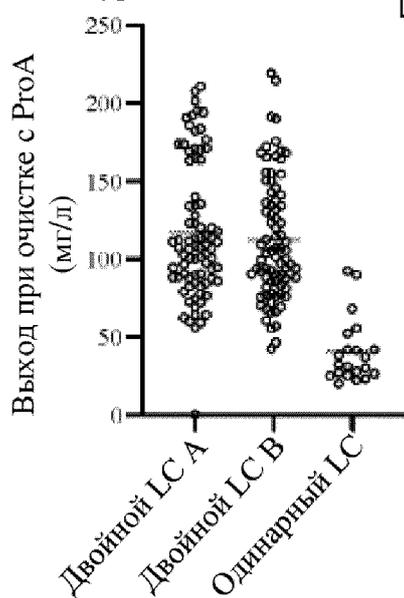
Фигура 3H



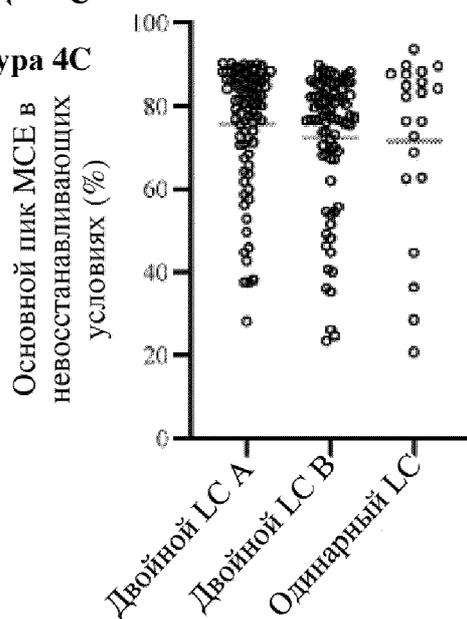
Фигура 4А



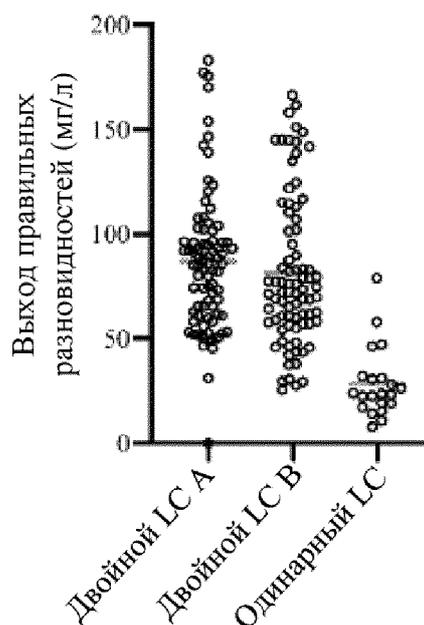
Фигура 4В

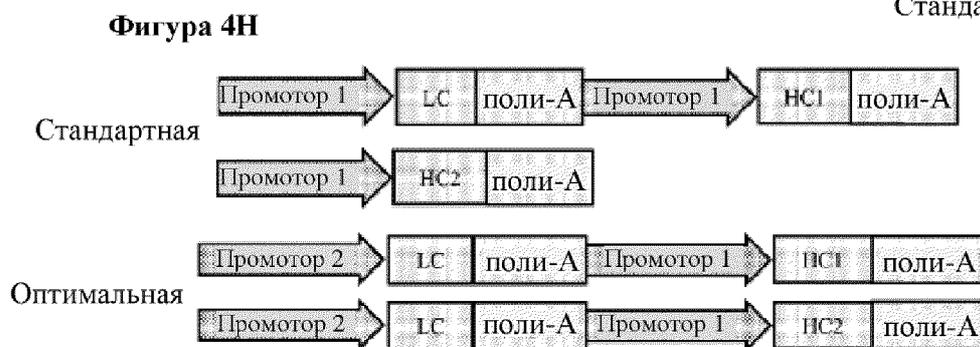
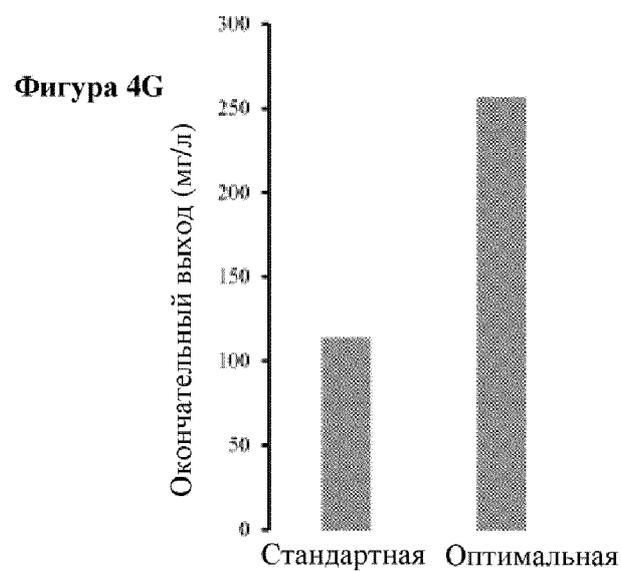
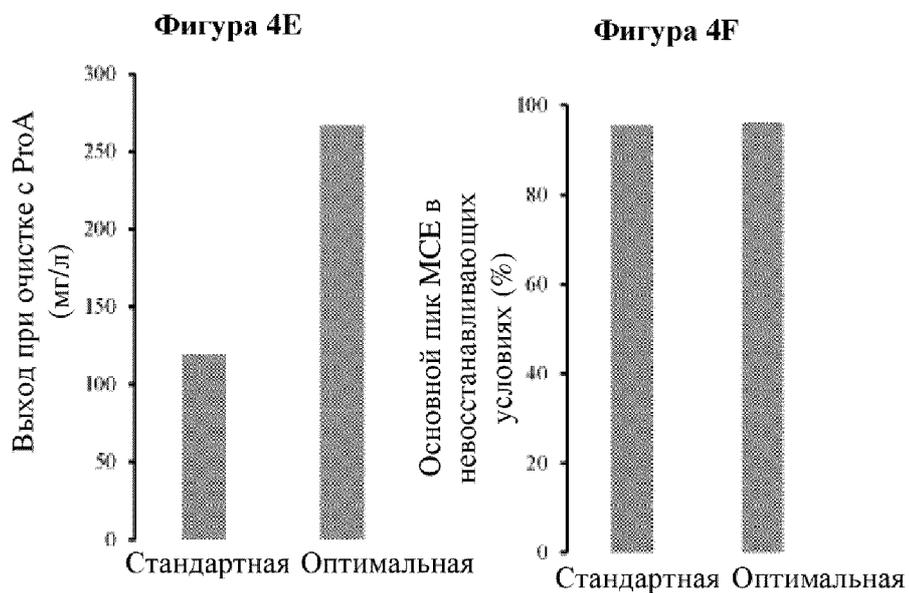


Фигура 4С

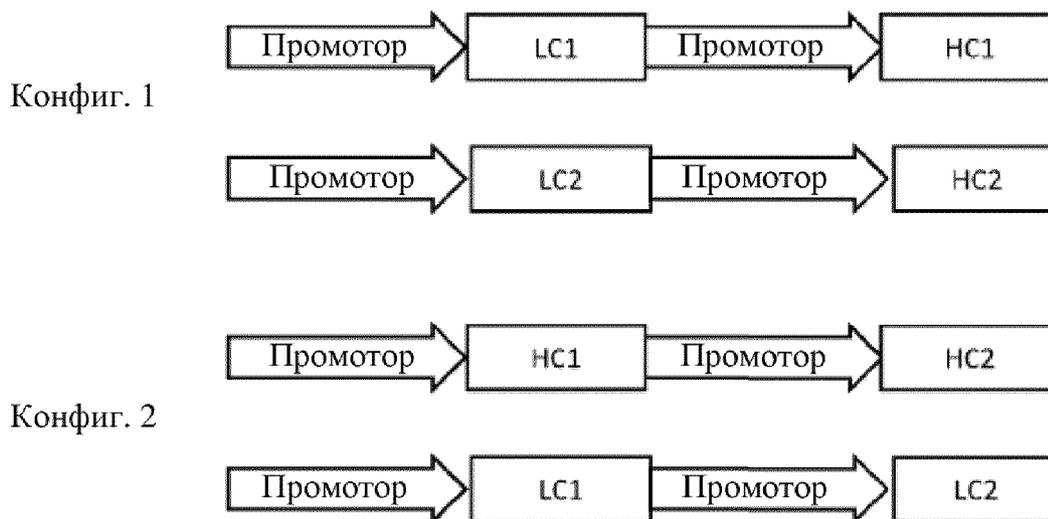


Фигура 4D

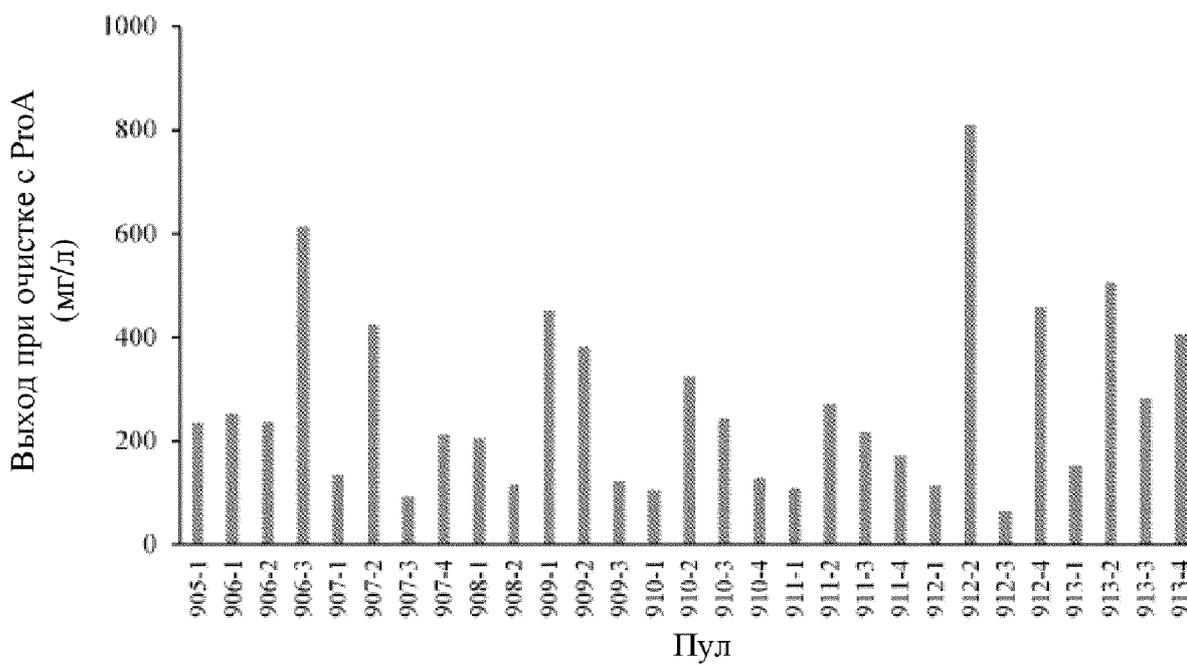




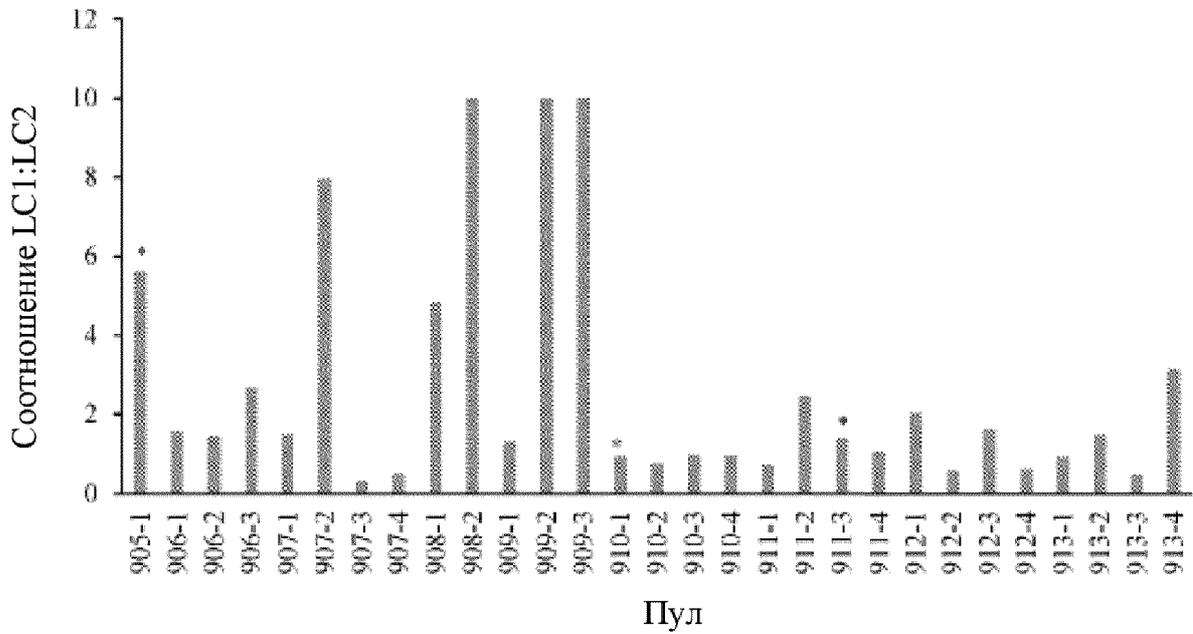
Фигура 5А



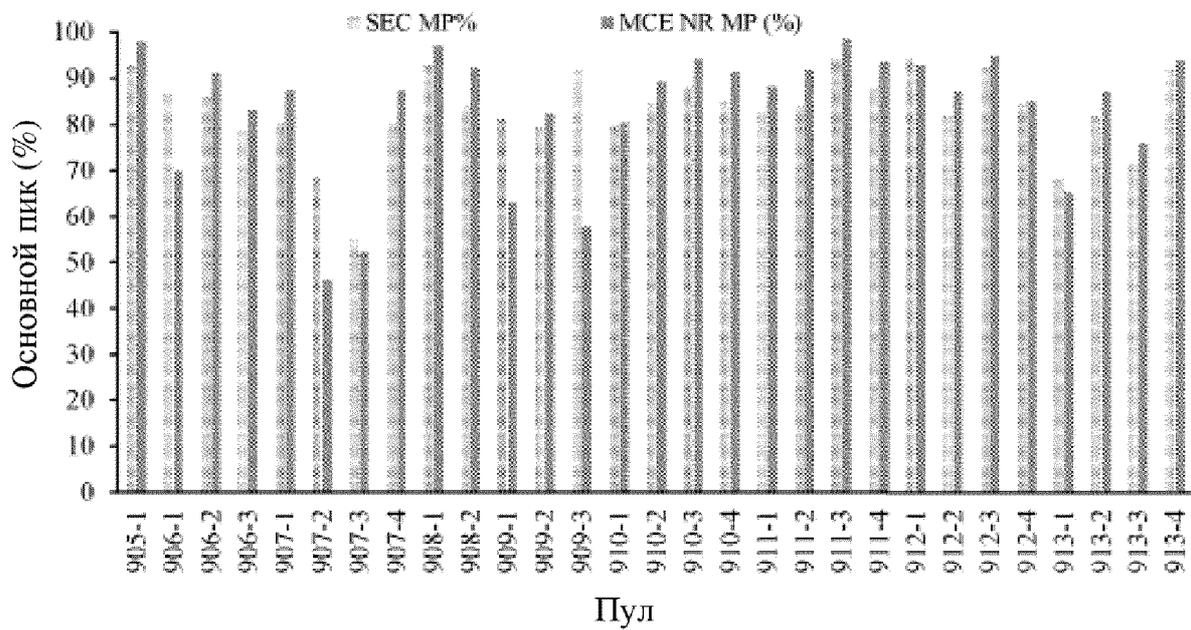
Фигура 5В



Фигура 5С



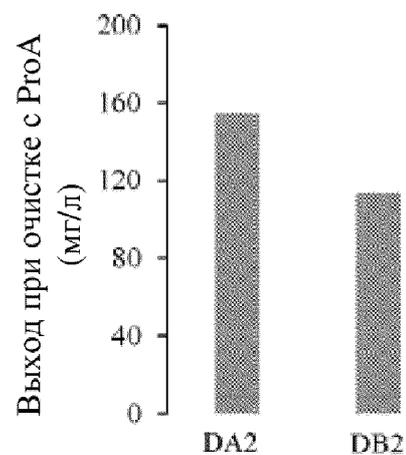
Фигура 5D



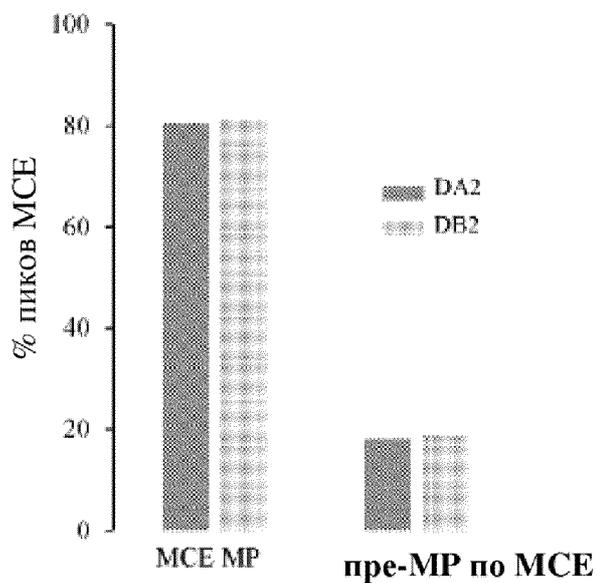
Фигура 6А



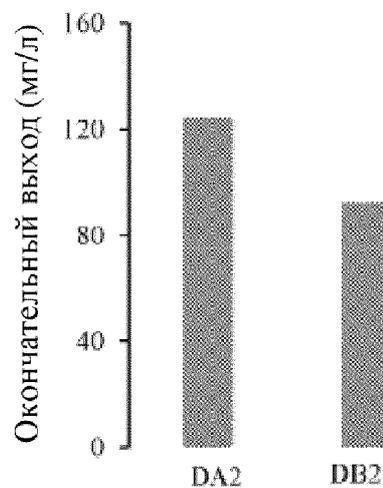
Фигура 6В



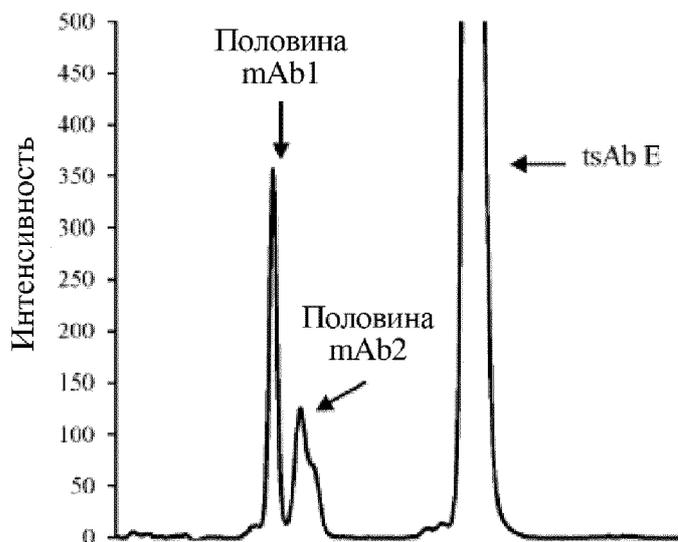
Фигура 6С



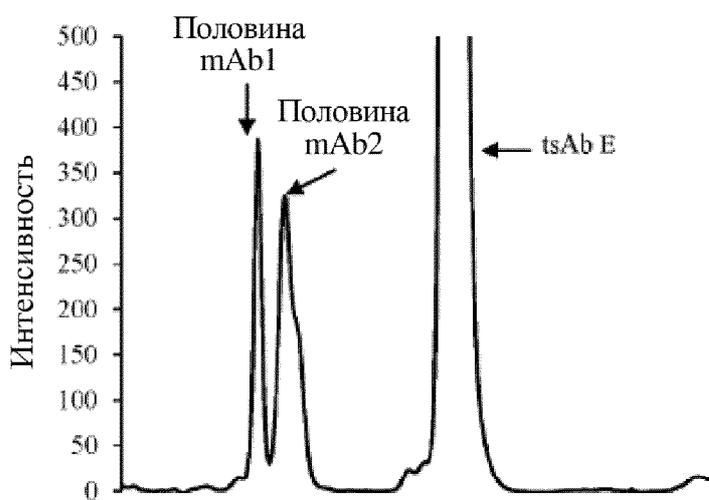
Фигура 6D



Фигура 6Е



Фигура 6F



Фигура 6G

