

(19)



**Евразийское
патентное
ведомство**

(21) **202292015** (13) **A2**

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

(43) Дата публикации заявки
2022.09.30

(22) Дата подачи заявки
2016.08.31

(51) Int. Cl. *C40B 20/04* (2006.01)
C40B 30/04 (2006.01)
C40B 40/08 (2006.01)
C40B 50/06 (2006.01)
C12N 5/0783 (2010.01)
C12N 5/0784 (2010.01)
C12N 15/62 (2006.01)
C12N 15/63 (2006.01)
C07K 14/705 (2006.01)

(54) **БИБЛИОТЕКИ МОДУЛЬНЫХ ПОЛИПЕПТИДОВ И СПОСОБЫ ИХ ПОЛУЧЕНИЯ И ПРИМЕНЕНИЯ**

(31) **62/212,999**

(32) **2015.09.01**

(33) **US**

(62) **201890611; 2016.08.31**

(71) Заявитель:
**ТЕ РИДЖЕНТС ОФ ТЕ
ЮНИВЕРСИТИ ОФ КАЛИФОРНИЯ
(US)**

(72) Изобретатель:
**Лим Вендел А., Койл Скотт М.,
Гордли Рассел М., Ройбал Коул Т. (US)**

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

(57) В настоящем изобретении предложены библиотеки синтетических модульных полипептидов и нуклеиновые кислоты, кодирующие указанные библиотеки синтетических модульных полипептидов. В настоящем изобретении также предложены способы получения библиотек синтетических модульных полипептидов и нуклеиновых кислот, кодирующих библиотеки синтетических модульных полипептидов. В настоящем изобретении также предложены способы скрининга библиотеки синтетических модульных полипептидов для идентификации выбранного фенотипа, связанного с элементом библиотеки синтетических модульных полипептидов, причем указанные способы можно применять в количественных исследованиях в условиях *in vitro* и в условиях *in vivo*.

A2

202292015

202292015

A2

БИБЛИОТЕКИ МОДУЛЬНЫХ ПОЛИПЕПТИДОВ И СПОСОБЫ ИХ ПОЛУЧЕНИЯ И ПРИМЕНЕНИЯ

Перекрестная ссылка

Настоящая заявка испрашивает приоритет согласно предварительной заявке на патент США №62/212999, поданной 1 сентября 2015 года, содержание которой полностью включено в настоящую заявку посредством ссылки.

Заявление об исследованиях, финансируемых из федерального бюджета

Настоящее изобретение было выполнено при государственной поддержке в рамках грантов EY016546, P50 GM081879, R01 CA196277, F32 GM006499 и R01 GM055040, присужденных Национальными институтами здравоохранения (США). Правительство США обладает определенными правами на настоящее изобретение.

Включение перечня последовательностей, предоставленного в виде текстового файла, посредством ссылки

Перечень последовательностей предоставлен в настоящей заявке в виде текстового файла «UCSF-518WO SeqList_ST25.txt», созданного 29 августа 2016 года и имеющего размер 145 килобайт. Содержимое текстового файла полностью включено в настоящую заявку посредством ссылки.

Введение

Многие эукариотические белки осуществляют свои функции с помощью модульных доменов или мотивов, которые контролируют или облегчают входные и выходные функции белка в целом. В настоящее время существуют способы перегруппировки и рекомбинации модульных доменов эукариотических белков, позволяющие создавать новые белки с измененными соотношениями входных/выходных функций и совершенно новыми общими функциями. Успешное конструирование отдельных модульных белков путем простой рекомбинации доменов, опосредующих входные и выходные функции, послужило доказательством принципа упомянутого модульного подхода к разработке новых белков.

Конструирование новых синтетических белков, например, для применения в терапевтических клетках, до настоящего времени осуществляли на основании подхода последовательного создания конструкций. Аналогичным образом, трансфекцию и скрининг таких синтетических белков также выполняли с использованием низкопроизводительного подхода, используя отдельные конструкции, которые подвергали скринингу на индивидуальной основе, или, в некоторых случаях, небольшое количество

отдельных конструкций подвергали одновременному скринингу. Одновременный скрининг, несмотря на то, что он является более эффективным, чем скрининг отдельных конструкций на индивидуальной основе, имеет ограниченную масштабируемость вследствие требования, заключающегося в том, что отдельные конструкции должны оставаться физически раздельными, чтобы облегчить окончательную идентификацию эффективных конструкций. Индивидуальный скрининг большого количества новых белков является обременительным при проведении количественного исследования в условиях *in vitro*, но становится еще более неприемлемо затруднительным и дорогостоящим при переходе испытаний к стадии количественных исследований, проводимых в моделях в условиях *in vivo*. Упомянутое индивидуальное получение и индивидуальный скрининг значительно ограничивают скорость разработки новых синтетических белков.

Краткое описание изобретения

В настоящем изобретении предложены библиотеки синтетических модульных полипептидов и нуклеиновые кислоты, кодирующие указанные библиотеки синтетических модульных полипептидов. В настоящем изобретении также предложены способы получения библиотек синтетических модульных полипептидов и нуклеиновых кислот, кодирующих библиотеки синтетических модульных полипептидов. В настоящем изобретении также предложены способы скрининга библиотеки синтетических модульных полипептидов для идентификации выбранного фенотипа, связанного с элементом библиотеки синтетических модульных полипептидов, причем указанные способы можно применять в количественных исследованиях в условиях *in vitro* и в условиях *in vivo*.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный способ включает: а) введение в клетки-хозяева библиотеки, содержащей нуклеиновые кислоты-штрихкоды, с получением гетерогенной популяции генетически модифицированных клеток-хозяев, причем указанная библиотека нуклеиновых кислот-штрихкодов содержит множество элементов, при этом каждый из множества элементов содержит нуклеотидную последовательность, кодирующую отличающийся синтетический модульный полипептид; и б) идентификацию генетически модифицированной клетки-хозяина в гетерогенной популяции, которая проявляет выбранный фенотип в ответ на стимул.

В настоящем изобретении также предложен способ, в котором синтетический модульный полипептид представляет собой полипептид химерного антигенного

рецептора (CAR) и в котором стимул представляет собой антигенпрезентирующую клетку, которая экспрессирует на своей поверхности антиген, который связан с CAR.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный стимул вступает в контакт с костимулирующей молекулой.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный синтетический модульный полипептид представляет собой модульный рецепторный полипептид.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный модульный рецепторный полипептид представляет собой химерный полипептид рецептора Notch.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный стимул представляет собой лиганд химерного полипептида рецептора Notch.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный синтетический модульный полипептид представляет собой модульный каркасный белок.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный синтетический модульный полипептид представляет собой модульный белок протеинкиназы или протеинфосфатазы.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный синтетический модульный полипептид представляет собой модульный транскрипционный регуляторный белок.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный синтетический модульный полипептид представляет собой модульный эпигенетический регуляторный белок.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный синтетический модульный полипептид представляет собой модульный белок

рекомбиназы или нуклеазы.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный выбранный фенотип имеет отличительную черту фенотипа, включающую два или более фенотипов.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, включающий секвенирование нуклеиновой последовательности-штрихкода из идентифицированной генетически модифицированной клетки-хозяина для идентификации синтетического модульного полипептида, связанного с фенотипом.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, включающий количественное определение синтетического модульного полипептида, связанного с фенотипом.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, включающий количественное определение отдельного модуля синтетических модульных полипептидов из библиотеки нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая связана с фенотипом.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем каждая нуклеотидная последовательность, кодирующая отличающийся синтетический модульный полипептид, содержит последовательность, кодирующую детектируемый репортер, который функционально соединен с нуклеотидной последовательностью, кодирующей синтетический модульный полипептид, причем указанный способ дополнительно включает разделение гетерогенной популяции генетически модифицированных клеток-хозяев на основании экспрессируемого детектируемого репортера.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем идентификацию выполняют в условиях *in vitro* или в условиях *ex vivo*.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный фенотип представляет собой один или более из: а) пролиферации; б) выработки цитокинов; в) экспрессии маркера клеточной поверхности; д) экспрессии репортерного

белка.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный способ включает библиотеку нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая содержит 100 или более уникальных элементов, и идентификация включает скрининг 100 или более уникальных элементов библиотеки для выбранного фенотипа.

В настоящем изобретении также предложен способ идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, причем указанный способ включает библиотеку нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, содержащую множество элементов, при этом каждый из множества элементов содержит нуклеотидную последовательность, кодирующую синтетический модульный полипептид, содержащий модульный домен, выбранный из группы, состоящей из: антигенсвязывающего домена, специфичного связывающего домена или белка специфичного партнера по связыванию, костимулирующего домена, коингибирующего домена, внутриклеточного сигнального домена, трансмембранного домена, домена каркасного белка, домена белка протеинкиназы, домена белка протеинфосфатазы, домена белка рецепторной тирозинкиназы, домена белка липидкиназы, домена белка липидфосфатазы, домена белка убиквитинилазы, домена белка деубиквитинилазы, домена белка SUMOилазы, домена белка ацетилазы, домена белка деацетилазы, домена белка метилазы, домена белка деметилазы, домена белка нуклеазы, домена белка рекомбиназы, домена белка транскрипционного фактора и их комбинаций.

В настоящем изобретении предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, причем указанная библиотека содержит: множество уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, при этом каждый уникальный полинуклеотид содержит: i) кодирующую область, кодирующую уникальный синтетический модульный полипептид, содержащую первую кодирующую последовательность, кодирующую первый модуль, соединенную с сохранением открытой рамки считывания со второй кодирующей последовательностью, кодирующей второй модуль, ii) область нуклеиновой последовательности-штрихкода, содержащую первую нуклеиновую последовательность-штрихкод, специфичную для первой кодирующей последовательности, соединенную со второй нуклеиновой последовательностью-штрихкодом, специфичной для второй кодирующей

последовательности, причем указанная первая и вторая нуклеиновые последовательности-штрихкоды расположены в обратном порядке от 5'-конца к 3'-концу по отношению к первой и второй кодирующим последовательностям; и при этом секвенирование каждой области нуклеиновой последовательности-штрихкода позволяет идентифицировать каждый уникальный синтетический модульный полипептид.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, в которой указанная первая и вторая кодирующие последовательности непосредственно соединены без каких-либо промежуточных не кодирующих нуклеотидов.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, в которой множество уникальных полинуклеотидов содержит по меньшей мере 1000 уникальных полинуклеотидов.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, в которой область нуклеиновой последовательности-штрихкода расположена в направлении 5'-конца относительно кодирующей области.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, в которой кодирующая область расположена в направлении 5'-конца относительно области нуклеиновой последовательности-штрихкода.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая содержит множество уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанный уникальный синтетический модульный полипептид содержит костимулирующий домен.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая содержит множество уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, который содержит первый и второй модули, причем указанный первый и второй модули содержат различные костимулирующие домены.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая включает множество

уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем каждый уникальный полинуклеотид дополнительно содержит промоторную последовательность, функционально соединенную с кодирующей областью и репортерной последовательностью, кодирующей детектируемый полипептид, причем указанный детектируемый полипептид представляет собой детектируемый оптическими способами полипептид, содержащий, например, детектируемый оптическими способами флуоресцентный полипептид.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая содержит множество уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, которая содержит кодирующую область, причем указанная кодирующая область дополнительно содержит третью кодирующую последовательность, кодирующую третий модуль, соединенную с сохранением открытой рамки считывания со второй кодирующей последовательностью, и область нуклеиновой последовательности-штрихкода содержит третью нуклеиновую последовательность-штрихкод, специфичную для третьей кодирующей последовательности, соединенную со второй нуклеиновой последовательностью-штрихкодом, причем указанная первая, вторая и третья нуклеиновые последовательности-штрихкоды расположены в обратном порядке от 5'-конца к 3'-концу по отношению к первой, второй и третьей кодирующим последовательностям.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая содержит множество уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества, представляют собой полипептиды химерного антигенного рецептора (CAR).

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая содержит множество уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества, представляют собой модульные

рецепторные полипептиды.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, кодирующих модульные рецепторные полипептиды, причем указанные модульные рецепторные полипептиды представляют собой химерные полипептиды рецептора Notch.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая содержит множество уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества, представляют собой модульные каркасные белки.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая содержит множество уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества, представляют собой модульные белки протеинкиназ или протеинфосфатаз.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая содержит множество уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества, представляют собой модульные транскрипционные регуляторные белки.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая содержит множество уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества, представляют собой модульные эпигенетические регуляторные белки.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая содержит множество

уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества, представляют собой модульные белки рекомбиназы или нуклеаз.

В настоящем изобретении также предложена библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которая содержит множество уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества, содержат модульный домен, выбранный из группы состоящий из: антигенсвязывающего домена, специфичного связывающего домена или белка специфичного партнера по связыванию, костимулирующего домена, коингибирующего домена, внутриклеточного сигнального домена, трансмембранного домена, домена каркасного белка, домена белка протеинкиназы, домена белка протеинфосфатазы, домена белка рецепторной тирозинкиназы, домена белка липидкиназы, домена белка липидфосфатазы, домена белка убиквитинилазы, домена белка деубиквитинилазы, домена белка SUMOилазы, домена белка ацетилазы, домена белка деацетилазы, домена белка метилазы, домена белка деметилазы, домена белка нуклеазы, домена белка рекомбиназы, домена белка транскрипционного фактора и их комбинаций.

В настоящем изобретении предложена клеточная библиотека, содержащая: множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем каждый уникальный полинуклеотид содержит: i) кодирующую область, кодирующую уникальный синтетический модульный полипептид, содержащую первую кодирующую последовательность, кодирующую первый модуль, соединенную с сохранением открытой рамки считывания со второй кодирующей последовательностью, кодирующей второй модуль, ii) область нуклеиновой последовательности-штрихкода, содержащую первую нуклеиновую последовательность-штрихкод, специфичную для первой кодирующей последовательности, соединенную со второй нуклеиновой последовательностью-штрихкодом, специфичной для второй кодирующей последовательности, причем указанная первая и вторая последовательности-штрихкоды расположены в обратном порядке от 5'-конца к 3'-концу по отношению к первой и второй кодирующим последовательностям; и при этом секвенирование каждой области

нуклеиновой последовательности-штрихкода позволяет идентифицировать каждый уникальный синтетический модульный полипептид каждой клетки библиотеки.

В настоящем изобретении также предложена клеточная библиотека, в которой клетки представляют собой прокариотические клетки или эукариотические клетки. В настоящем изобретении также предложена клеточная библиотека, в которой эукариотические клетки представляют собой клетки человека. В настоящем изобретении также предложена клеточная библиотека, в которой клетки человека представляют собой Т-клетки человека.

В настоящем изобретении также предложена клеточная библиотека, которая содержит множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, которая содержит кодирующую область, кодирующую уникальный синтетический модульный полипептид, содержащую первую кодирующую последовательность, кодирующую первый модуль, соединенную с сохранением открытой рамки считывания со второй кодирующей последовательностью, кодирующей второй модуль, причем указанная первая и вторая кодирующие последовательности непосредственно соединены без каких-либо промежуточных не кодирующих нуклеотидов.

В настоящем изобретении также предложена клеточная библиотека, которая содержит множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, которая содержит кодирующую область, кодирующую уникальный синтетический модульный полипептид, содержащую первую кодирующую последовательность, кодирующую первый модуль, соединенную с сохранением открытой рамки считывания со второй кодирующей последовательностью, кодирующей второй модуль, причем указанная кодирующая область дополнительно содержит последовательность, кодирующую репортер, соединенную с сохранением открытой рамки считывания со второй кодирующей последовательностью.

В настоящем изобретении также предложена клеточная библиотека, которая содержит последовательность, кодирующую репортер, соединенную с сохранением открытой рамки считывания со второй кодирующей последовательностью, причем указанный репортер представляет собой эпигенетическую метку.

В настоящем изобретении также предложена клеточная библиотека, которая содержит последовательность, кодирующую репортер, соединенную с сохранением открытой рамки считывания со второй кодирующей последовательностью, причем указанный репортер представляет собой флуоресцентный белок.

В настоящем изобретении также предложена клеточная библиотека, которая содержит множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанное множество клеток содержит по меньшей мере 1000 уникальных полинуклеотидов.

В настоящем изобретении также предложена клеточная библиотека, которая содержит множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, которая содержит кодирующую область, кодирующую уникальный синтетический модульный полипептид, причем указанная кодирующая область дополнительно содержит третью кодирующую последовательность, кодирующую третий модуль, соединенную с сохранением открытой рамки считывания со второй кодирующей последовательностью, и область нуклеиновой последовательности-штрихкода содержит третью последовательность-штрихкод, специфичную для третьей кодирующей последовательности, соединенную со второй нуклеиновой последовательностью-штрихкодом, причем указанная первая, вторая и третья нуклеиновые последовательности-штрихкоды расположены в обратном порядке от 5'-конца к 3'-концу по отношению к первой, второй и третьей кодирующим последовательностям.

В настоящем изобретении также предложена клеточная библиотека, которая содержит множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества клеток, представляют собой полипептиды химерного антигенного рецептора (CAR).

В настоящем изобретении также предложена клеточная библиотека, которая содержит множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества клеток, представляют собой модульные рецепторные полипептиды.

В настоящем изобретении также предложена клеточная библиотека, которая содержит множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные модульные рецепторные полипептиды представляют собой химерные полипептиды рецептора Notch.

В настоящем изобретении также предложена клеточная библиотека, которая содержит множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества клеток, представляют собой модульные каркасные белки.

В настоящем изобретении также предложена клеточная библиотека, которая содержит множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества клеток, представляют собой модульные белки протеинкиназ или протеинфосфатаз.

В настоящем изобретении также предложена клеточная библиотека, которая содержит множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества клеток, представляют собой модульные транскрипционные регуляторные белки.

В настоящем изобретении также предложена клеточная библиотека, которая содержит множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества клеток, представляют собой модульные эпигенетические регуляторные белки.

В настоящем изобретении также предложена клеточная библиотека, которая содержит множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества клеток, представляют собой модульные белки рекомбиназ или нуклеаз.

В настоящем изобретении также предложена клеточная библиотека, которая содержит множество клеток, каждая из которых содержит уникальный полинуклеотид, содержащий нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем указанные уникальные синтетические модульные полипептиды, кодируемые каждым уникальным полинуклеотидом из множества клеток,

содержат модульный домен, выбранный из группы, состоящей из: антигенсвязывающего домена, специфичного связывающего домена или белка специфичного партнера по связыванию, костимулирующего домена, коингибирующего домена, внутриклеточного сигнального домена, трансмембранного домена, домена каркасного белка, домена белка протеинкиназы, домена белка протеинфосфатазы, домена белка рецепторной тирозинкиназы, домена белка липидкиназы, домена белка липидфосфатазы, домена белка убиквитинилазы, домена белка деубиквитинилазы, домена белка SUMOилазы, домена белка ацетилазы, домена белка деацетилазы, домена белка метилазы, домена белка деметилазы, домена белка нуклеазы, домена белка рекомбиназы, домена белка транскрипционного фактора и их комбинаций.

В настоящем изобретении предложен способ получения библиотеки нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, каждая из которых кодирует уникальный синтетический модульный полипептид, причем указанный способ включает: приведение первого полинуклеотида, содержащего последовательность, кодирующую первый модуль, соединенную с первой нуклеиновой последовательностью-штрихкодом, в контакт со вторым полинуклеотидом, содержащим последовательность, кодирующую второй модуль, соединенную со второй нуклеиновой последовательностью-штрихкодом, в условиях, достаточных для введения первого полинуклеотида во второй полинуклеотид в месте соединения второй кодирующей последовательности со второй нуклеиновой последовательностью-штрихкодом с получением бимодульного полинуклеотида, содержащего нуклеиновые последовательности-штрихкоды, причем указанный бимодульный полинуклеотид, содержащий нуклеиновые последовательности-штрихкоды, содержит вторую модульную кодирующую последовательность, соединенную с сохранением открытой рамки считывания с первой модульной кодирующей последовательностью, соединенной с первой нуклеиновой последовательностью-штрихкодом, соединенной со второй нуклеиновой последовательностью-штрихкодом.

В настоящем изобретении также предложен способ получения библиотеки нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, каждая из которых кодирует уникальный синтетический модульный полипептид, причем указанный способ дополнительно включает приведение первого и второго полинуклеотидов в контакт с третьим полинуклеотидом, содержащим третью модульную кодирующую последовательность, соединенную с третьей нуклеиновой последовательностью-штрихкодом, причем указанный третий полинуклеотид встраивается в первый полинуклеотид в месте соединения первой кодирующей

последовательности и первой нуклеиновой последовательности-штрихкода с получением трехмодульного полинуклеотида, содержащего нуклеиновые последовательности-штрихкоды, при этом трехмодульный полинуклеотид, содержащий нуклеиновые последовательности-штрихкоды, содержит вторую модульную кодирующую последовательность, соединенную с сохранением открытой рамки считывания с первой модульной кодирующей последовательностью, соединенной с сохранением открытой рамки считывания с третьей модульной кодирующей последовательностью, соединенной с третьей нуклеиновой последовательностью-штрихкодом, соединенной с первой нуклеиновой последовательностью-штрихкодом, соединенной со второй нуклеиновой последовательностью-штрихкодом.

В настоящем изобретении также предложен способ получения библиотеки нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, каждая из которых кодирует уникальный синтетический модульный полипептид, причем указанная первая и вторая модульные кодирующие последовательности соединены с сохранением открытой рамки считывания без каких-либо промежуточных некодирующих нуклеотидов.

В настоящем изобретении также предложен способ получения библиотеки нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, каждая из которых кодирует уникальный синтетический модульный полипептид, причем указанная библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, содержит 1000 или более уникальных нуклеиновых кислот, каждая из которых кодирует уникальный синтетический модульный полипептид.

В настоящем изобретении также предложен способ получения библиотеки нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, каждая из которых кодирует уникальный синтетический модульный полипептид, причем указанные нуклеиновые последовательности-штрихкоды расположены в направлении 5'-конца относительно последовательностей, кодирующих модули, или в направлении 3'-конца относительно последовательностей, кодирующих модули.

В настоящем изобретении также предложен способ получения библиотеки нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, каждая из которых кодирует уникальный синтетический модульный полипептид, причем введение первого полинуклеотида во второй полинуклеотид в месте соединения второй кодирующей последовательности и второй нуклеиновой последовательности-штрихкода опосредовано активностью фермента рестрикции, который распознает сайт распознавания фермента рестрикции на втором полинуклеотиде и расщепляет второй полинуклеотид

между второй кодирующей последовательностью и второй нуклеиновой последовательностью-штрихкодом, включая случаи, когда используемый фермент рестрикции представляет собой фермент рестрикции типа II, включая случаи, когда используемый фермент рестрикции представляет собой фермент рестрикции типа IIS.

В настоящем изобретении также предложен способ получения библиотеки нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, каждая из которых кодирует уникальный синтетический модульный полипептид, причем указанный первый полинуклеотид также содержит сайт распознавания фермента рестрикции, который присутствует на втором полинуклеотиде.

В настоящем изобретении также предложен способ получения библиотеки нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, каждая из которых кодирует уникальный синтетический модульный полипептид, причем указанный способ дополнительно включает приведение бимодульных полинуклеотидов, содержащих нуклеиновые последовательности-штрихкоды, в контакт с полинуклеотидом, кодирующим репортер, в условиях, достаточных для введения полинуклеотида, кодирующего репортер, в бимодульный полинуклеотид, содержащий нуклеиновые последовательности-штрихкоды, в месте соединения первой модульной кодирующей последовательности и первой нуклеиновой последовательности-штрихкода с получением связанного с репортером бимодульного полинуклеотида, содержащего нуклеиновые последовательности-штрихкоды, причем указанный связанный с репортером бимодульный полинуклеотид, содержащий нуклеиновые последовательности-штрихкоды, содержит вторую модульную кодирующую последовательность, соединенную с сохранением открытой рамки считывания с первой модульной кодирующей последовательностью, соединенной с сохранением открытой рамки считывания с последовательностью, кодирующей репортер, соединенной с первой нуклеиновой последовательностью-штрихкодом, соединенной со второй нуклеиновой последовательностью-штрихкодом. В настоящем изобретении также предложен способ получения библиотеки нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, в которой последовательность, кодирующая репортер, кодирует эпигенетическую метку или флуоресцентный репортер.

В настоящем изобретении предложен химерный антигенный рецептор (CAR), идентифицированный путем скрининга библиотеки нуклеиновых кислот, кодирующих синтетические модульные полипептиды CAR, причем указанный CAR содержит по меньшей мере один комодулирующий домен, перечисленный в таблице 3 или таблице 4.

В настоящем изобретении также предложен CAR, идентифицированный путем

скрининга библиотеки нуклеиновых кислот, кодирующих синтетические модульные полипептиды CAR, причем указанный CAR содержит CD19-специфичный антигенсвязывающий домен.

В настоящем изобретении также предложен CAR, идентифицированный путем скрининга библиотеки нуклеиновых кислот, кодирующих синтетические модульные полипептиды CAR, причем указанный CAR содержит первичный сигнальный домен CD3-дзета.

В настоящем изобретении также предложен CAR, идентифицированный путем скрининга библиотеки нуклеиновых кислот, кодирующих синтетические модульные полипептиды CAR, причем указанный CAR стимулирует активность Т-клеток и содержит по меньшей мере один костимулирующий домен, перечисленный в таблице 3.

В настоящем изобретении также предложен CAR, идентифицированный путем скрининга библиотеки нуклеиновых кислот, кодирующих синтетические модульные полипептиды CAR, причем указанный CAR ингибирует активность Т-клеток и содержит по меньшей мере один коингибирующий домен, перечисленный в таблице 4.

В настоящем изобретении также предложена нуклеиновая кислота, кодирующая CAR, идентифицированный путем скрининга библиотеки нуклеиновых кислот, кодирующих синтетические модульные полипептиды CAR, включая, например, любой из CAR, описанных в настоящем документе.

Краткое описание чертежей

На Фиг. 1 представлена модульная последовательность, содержащая нуклеиновую последовательность-штрихкод, субклонированная для сборки библиотеки согласно одному варианту реализации настоящего изобретения.

На Фиг. 2 представлена модульная последовательность, содержащая нуклеиновую последовательность-штрихкод, показанная на Фиг. 1, после расщепления ферментом рестрикции типа IIS при подготовке к сборке библиотеки.

На Фиг. 3 представлен примерный фрагмент последовательности в соответствии с Фиг. 1 (перед расщеплением ферментом рестрикции типа IIS (RE), SEQ ID NO: 16) и Фиг. 2 (после расщепления ферментом рестрикции IIS, SEQ ID NO: 17), содержащий последовательность, кодирующую костимулирующий домен CD28 (трансляция фрагмента после расщепления ферментом рестрикции IIS, SEQ ID NO: 18).

На Фиг. 4 представлена таблица 1. Идентификационные номера последовательностей для последовательностей, представленных в таблице 1, были следующими: индекс 1 – SEQ ID NO: 19; индекс 2 – SEQ ID NO: 20; индекс 3 – SEQ ID NO: 21; индекс 4 – SEQ ID NO: 22; индекс 5 – SEQ ID NO: 23; индекс 6 – SEQ ID NO: 24;

индекс 7 – SEQ ID NO: 25; индекс 8 – SEQ ID NO: 26; индекс 9 – SEQ ID NO: 27; индекс 10 – SEQ ID NO: 28; индекс 11 – SEQ ID NO: 29; индекс 12 – SEQ ID NO: 30; индекс 13 – SEQ ID NO: 31; индекс 14 – SEQ ID NO: 32; индекс 15 – SEQ ID NO: 33; индекс 16 – SEQ ID NO: 34; индекс 17 – SEQ ID NO: 35; индекс 18 – SEQ ID NO: 36; индекс 19 – SEQ ID NO: 37; индекс 20 – SEQ ID NO: 38; индекс 21 – SEQ ID NO: 39; индекс 22 – SEQ ID NO: 40; индекс 23 – SEQ ID NO: 41; индекс 24 – SEQ ID NO: 42; индекс 25 – SEQ ID NO: 43; индекс 26 – SEQ ID NO: 44; индекс 27 – SEQ ID NO: 45; индекс 28 – SEQ ID NO: 46; индекс 29 – SEQ ID NO: 47; индекс 30 – SEQ ID NO: 48; индекс 31 – SEQ ID NO: 49; индекс 32 – SEQ ID NO: 50; индекс 33 – SEQ ID NO: 51; индекс 34 – SEQ ID NO: 52; индекс 35 – SEQ ID NO: 53; индекс 36 – SEQ ID NO: 54; индекс 37 – SEQ ID NO: 55; индекс 38 – SEQ ID NO: 56; индекс 39 – SEQ ID NO: 57; индекс 40 – SEQ ID NO: 58; индекс 41 – SEQ ID NO: 59; индекс 42 – SEQ ID NO: 60; индекс 43 – SEQ ID NO: 61; индекс 44 – SEQ ID NO: 62; индекс 45 – SEQ ID NO: 63; индекс 46 – SEQ ID NO: 64; индекс 47 – SEQ ID NO: 65; индекс 48 – SEQ ID NO: 66; индекс 49 – SEQ ID NO: 67; индекс 50 – SEQ ID NO: 68; индекс 51 – SEQ ID NO: 69; индекс 52 – SEQ ID NO: 70; индекс 53 – SEQ ID NO: 71; индекс 54 – SEQ ID NO: 72; индекс 55 – SEQ ID NO: 73; индекс 56 – SEQ ID NO: 74; индекс 57 – SEQ ID NO: 75; индекс 58 – SEQ ID NO: 76; индекс 59 – SEQ ID NO: 77; индекс 60 – SEQ ID NO: 26; индекс 61 – SEQ ID NO: 26; индекс 62 – SEQ ID NO: 26.

На Фиг. 5 представлена общая схема ступенчатой сборки двумерной библиотеки комодулирующих модулей, содержащей нуклеиновые последовательности-штрихкоды, в соответствии с одним вариантом реализации настоящего изобретения.

На Фиг. 6 представлена общая схема получения двумерной библиотеки из 62×62 комодулирующих модулей согласно одному варианту реализации настоящего изобретения.

На Фиг. 7 представлена общая конфигурация элементов одномерной библиотеки химерных антигенных рецепторов (CAR) в соответствии с одним вариантом реализации настоящего изобретения.

На Фиг. 8 представлены результаты функциональной сортировки (т.е. бининга) стимулированных CAR-экспрессирующих Т-клеток на основании конечного уровня активации.

На Фиг. 9 представлены результаты количественной оценки относительного влияния каждого комодулирующего домена библиотеки на активность Т-клеток, которую осуществляли с помощью количественного секвенирования специфичных для модуля последовательностей-штрихкодов.

На Фиг. 10 представлены характеристики зависимости ответа от дозы для шести комодулирующих доменов при трех повышающихся уровнях вносимого антигена, полученные с использованием 62-элементной одномерной библиотеки.

На Фиг. 11 представлен неограничивающий пример библиотеки CAR, который свидетельствует о том, что различные домены CAR могут варьироваться в библиотеке в соответствии с одним вариантом реализации настоящего изобретения.

На Фиг. 12 представлена схема стратегии сборки кодирующих модули последовательностей, содержащих нуклеиновые последовательности-штрихкоды, описанных в настоящем документе.

На Фиг. 13 представлена схема стратегии сборки кодирующих модули последовательностей, содержащих нуклеиновые последовательности-штрихкоды, описанных в настоящем документе.

На Фиг. 14 представлена схема стратегии сборки кодирующих модули последовательностей, содержащих нуклеиновые последовательности-штрихкоды, описанных в настоящем документе.

На Фиг. 15 представлена схема стратегии сборки кодирующих модули последовательностей, содержащих нуклеиновые последовательности-штрихкоды, описанных в настоящем документе.

На Фиг. 16 представлена схема стратегии сборки кодирующих модули последовательностей, содержащих нуклеиновые последовательности-штрихкоды, описанных в настоящем документе.

На Фиг. 17 представлена схема стратегии сборки кодирующих модули последовательностей, содержащих нуклеиновые последовательности-штрихкоды, описанных в настоящем документе.

На Фиг. 18 представлена схема стратегии сборки кодирующих модули последовательностей, содержащих нуклеиновые последовательности-штрихкоды, описанных в настоящем документе.

На Фиг. 19 представлена схема стратегии сборки кодирующих модули последовательностей, содержащих нуклеиновые последовательности-штрихкоды, описанных в настоящем документе.

На Фиг. 20 представлена схема стратегии сборки кодирующих модули последовательностей, содержащих нуклеиновые последовательности-штрихкоды, описанных в настоящем документе.

На Фиг. 21 представлена схема стратегии сборки кодирующих модули последовательностей, содержащих нуклеиновые последовательности-штрихкоды,

описанных в настоящем документе.

На Фиг. 22 представлена схема стратегии комбинаторной вложенной сборки библиотеки модульных химерных антигенных рецепторов (CAR) в соответствии с вариантом реализации настоящего изобретения.

На Фиг. 23 представлена схема получения объединенной клеточной модульной библиотеки CAR в соответствии с вариантом реализации настоящего изобретения.

На Фиг. 24 представлен схематичный пример интегрированного способа детектирования фенотипа и идентификации модульного полипептида для применения в условиях *in vitro* и/или в условиях *in vivo* в соответствии с вариантом реализации настоящего изобретения.

На Фиг. 25 представлена таблица 2.

На Фиг. 26 исчерпывающе представлен каждый элемент двумерной библиотеки из 61×61 элемента, согласно результатам определения путем глубокого секвенирования собранной библиотеки.

На Фиг. 27 представлены результаты количественной оценки элементов предварительно нормированной комбинаторной библиотеки нуклеиновых кислот, описанной в настоящем документе.

На Фиг. 28 представлено линейное уравнение, использованное при расчете корректировок для нормирования, относящихся к элементам библиотеки, количественно определенным, как показано на Фиг. 27.

На Фиг. 29 представлены результаты, подтверждающие предсказуемость корректировок для нормирования, выполненных в соответствии с описанным в настоящем документе способом.

ОПРЕДЕЛЕНИЯ

В настоящей заявке термины «полинуклеотид» и «нуклеиновая кислота» используются взаимозаменяемо и относятся к полимерной форме нуклеотидов любой длины, как рибонуклеотидов, так и дезоксирибонуклеотидов. Следовательно, данный термин включает, но не ограничивается ими, одно-, двух- или многоцепочечную ДНК или РНК, геномную ДНК, кДНК, гибриды ДНК-РНК или полимер, содержащий пуриновые и пиримидиновые основания или другие природные, химически или биохимически модифицированные, неприродные или дериватизированные нуклеотидные основания.

В настоящей заявке термины «полипептид», «пептид» и «белок» используются взаимозаменяемо и относятся к полимерной форме аминокислот любой длины, которая может включать генетически кодируемые и не кодируемые генетически аминокислоты, химически или биохимически модифицированные или дериватизированные

аминокислоты и полипептиды, имеющие модифицированные пептидные остовы. Термин включает гибридные белки, включая, но не ограничиваясь ими, гибридные белки с гетерологичной аминокислотной последовательностью, гибриды с гетерологичными и гомологичными лидерными последовательностями, с N-концевыми остатками метионина или без указанных остатков; иммунологически меченые белки; и т.п.

В настоящей заявке термины «домен» и «мотив» используются взаимозаменяемо и относятся к структурированным доменам, имеющим одну или более конкретных функций, и неструктурированным сегментам полипептида, которые, несмотря на отсутствие структурированности, сохраняют одну или более конкретных функций. Например, структурированный домен может включать, но не ограничивается ими, непрерывное или дискретное множество аминокислот, или их частей, в свернутом полипептиде, который имеет трехмерную структуру, которая вносит вклад в определенную функцию полипептида. В других случаях домен может включать неструктурированный сегмент полипептида, содержащий множество из двух или более аминокислот, или их частей, который поддерживает определенную функцию полипептида, развернутого или неупорядоченного. Данное определение также включает домены, которые могут быть неупорядоченными или неструктурированными, но становятся структурированными или упорядоченными при связывании с мишенью или партнером по связыванию. Неограничивающие примеры исходно неструктурированных доменов и доменов исходно неструктурированных белков описаны, например, в Dyson & Wright. *Nature Reviews Molecular Cell Biology* 6:197-208.

В настоящей заявке термин «модуль» относится к непрерывной полипептидной последовательности или ее фрагменту, который связан с той или иной функцией, в частности с биологической функцией.

В настоящей заявке термины «химерный антигенный рецептор» и «CAR» используются взаимозаменяемо и относятся к искусственным многомодульным молекулам, способным инициировать или ингибировать активацию иммунной клетки, которые обычно, но не исключительно, содержат внеклеточный домен (например, домен, связывающий лиганд/антиген), трансмембранный домен и один или более внутриклеточных сигнальных доменов. Термин CAR не ограничивается конкретными молекулами CAR, а также включает варианты CAR. Варианты CAR включают разделенные CAR, в которых внеклеточная часть (например, участок связывания лиганда) и внутриклеточная часть (например, внутриклеточная сигнальная часть) CAR присутствуют на двух отдельных молекулах. Варианты CAR также включают CAR-включатели и CAR-выключатели, которые представляют собой

активируемые/подавляемые при определенных условиях CAR, например, указанные варианты включают расщепленный CAR, в котором условная гетеродимеризация двух частей расщепленного CAR контролируется фармакологически. Варианты CAR также включают биспецифичные CAR, которые содержат домен связывания дополнительного CAR, который может усиливать или ингибировать активность основного CAR. Варианты CAR также включают ингибиторные химерные антигенные рецепторы (iCAR), которые можно применять, например, в качестве компонента биспецифичной системы CAR, когда связывание домена связывания с дополнительным CAR приводит к ингибированию активации основного CAR. Молекулы CAR и их производные (т.е. варианты CAR) описаны, например, в заявке PCT US2014/016527; Fedorov et al. *Sci Transl Med* (2013) ;5(215):215ra172; Glienke et al. *Front Pharmacol* (2015) 6:21; Kakarla & Gottschalk *Cancer J* (2014) 20(2):151-5; Riddell et al. *Cancer J* (2014) 20(2):141-4; Pegram et al. *Cancer J* (2014) 20(2):127-33; Cheadle et al. *Immunol Rev* (2014) 257(1):91-106; Barrett et al. *Annu Rev Med* (2014) 65:333-47; Sadelain et al. *Cancer Discov* (2013) 3(4):388-98; Cartellieri et al., *J Biomed Biotechnol* (2010) 956304, содержание которых полностью включено в настоящую заявку посредством ссылки.

Термин «ген» относится к определенной единице наследственности, присутствующей в определенном локусе в генетическом компоненте организма. Ген может представлять собой последовательность нуклеиновой кислоты, например, последовательность ДНК или РНК, присутствующую в геноме нуклеиновых кислот, геноме ДНК или РНК, организма, и, в некоторых случаях, может присутствовать на хромосоме. Ген может представлять собой последовательность ДНК, кодирующую мРНК, которая кодирует белок. Ген может состоять из одного экзона и может не иметь интронов или может содержать несколько экзонов и один или более интронов. Одна из двух или более идентичных или альтернативных форм гена, присутствующих в определенном локусе, называется «аллель», и, например, диплоидный организм обычно будет иметь два аллеля определенного гена. Новые аллели определенного гена могут возникнуть природным или искусственным путем посредством природной или индуцированной мутации и могут распространяться путем размножения или клонирования. Ген или аллель может быть выделен из генома организма и реплицирован и/или подвергнут манипуляциям, или ген или аллель может быть модифицирован в условиях *in situ* методами генной терапии. Локус гена или аллель может иметь связанные регуляторные элементы, и генная терапия, в некоторых случаях, может включать модификацию регуляторных элементов гена или аллеля, оставляя кодирующие последовательности гена или аллеля немодифицированными.

Термин «функционально соединенный» относится к сопоставлению, при котором описанные компоненты взаимосвязаны, что позволяет им функционировать предполагаемым способом. Например, промотор функционально соединен с кодирующей последовательностью, если промотор влияет на ее транскрипцию или экспрессию.

«Вектор» или «вектор экспрессии» представляет собой репликон, такой как плазида, фаг, вирус или космида, к которому может быть присоединен другой сегмент ДНК, т.е. «вставка», чтобы вызвать репликацию присоединенного сегмента в клетке.

В настоящей заявке термин «гетерологичный» означает нуклеотидную или полипептидную последовательность, которая не встречается в нативной (например, природной) нуклеиновой кислоте или белке, соответственно.

Термины «антитела» и «иммуноглобулин» включают антитела или иммуноглобулины любого изотипа, фрагменты антител, которые сохраняют специфичное связывание с антигеном, включая, но не ограничиваясь ими, фрагменты Fab, Fv, scFv и Fd, химерные антитела, гуманизированные антитела, одноцепочечные антитела и гибридные белки, содержащие антигенсвязывающую часть антитела и белка, не относящегося к антителам.

«Фрагменты антитела» содержат часть интактного антитела, например, антигенсвязывающую или переменную область интактного антитела. Примеры фрагментов антител включают фрагменты Fab, Fab', F(ab')₂ и Fv; диатела; линейные антитела (Zapata et al., Protein Eng. 8(10): 1057-1062 (1995)); молекулы одноцепочечных антител; и полиспецифичные антитела, образованные из фрагментов антител. Расщепление антител папином позволяет получить два идентичных антигенсвязывающих фрагмента, называемых фрагментами Fab, каждый из которых содержит один антигенсвязывающий сайт, и остаточный фрагмент «Fc», обозначение, которое отражает способность фрагмента легко кристаллизоваться. Обработка пепсином позволяет получить фрагмент F(ab')₂, который содержит два антигенсвязывающих сайта и по-прежнему способен к сшиванию антигена.

«Одноцепочечные Fv» или «sFv» фрагменты антитела содержат домены V_H и V_L антитела, причем указанные домены присутствуют в одной полипептидной цепи. Согласно некоторым вариантам реализации настоящего изобретения полипептид Fv дополнительно содержит полипептидный линкер между доменами V_H и V_L, который позволяет sFv формировать желаемую структуру для связывания антигена. Для обзора sFv см. Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

В настоящей заявке термин «аффинность» относится к константе равновесия для

обратимого связывания двух агентов и выражается как константа диссоциации (K_d). Аффинность может быть по меньшей мере в 1 раз больше, по меньшей мере в 2 раза больше, по меньшей мере в 3 раза больше, по меньшей мере в 4 раза больше, по меньшей мере в 5 раз больше, по меньшей мере в 6 раз больше, по меньшей мере в 7 раз больше, по меньшей мере в 8 раз больше, по меньшей мере в 9 раз больше, по меньшей мере в 10 раз больше, по меньшей мере в 20 раз больше, по меньшей мере в 30 раз больше, по меньшей мере в 40 раз больше, по меньшей мере в 50 раз больше, по меньшей мере в 60 раз больше, по меньшей мере в 70 раз больше, по меньшей мере в 80 раз больше, по меньшей мере в 90 раз больше, по меньшей мере в 100 раз больше или по меньшей мере в 1000 раз больше или более, чем аффинность антитела в отношении нецелевых аминокислотных последовательностей. Аффинность антитела в отношении целевого белка может быть, например, от приблизительно 100 наномоль (нМ) до приблизительно 0,1 нМ, от приблизительно 100 нМ до приблизительно 1 пикомоль (пМ) или от приблизительно 100 нМ до приблизительно 1 фемтомоль (фМ) или более.

Термин «связывание» относится к непосредственной связи двух молекул посредством, например, ковалентных, электростатических, гидрофобных и ионных и/или водородных связей, включая такие взаимодействия как солевые мостики и водные мостики. Неспецифичное связывание относится к связыванию с величиной аффинности менее приблизительно 10^{-7} М, например, к связыванию с аффинностью 10^{-6} М, 10^{-5} М, 10^{-4} М и т.д.

В настоящей заявке термин «иммунные клетки» обычно включает белые клетки крови (лейкоциты), которые получены из гемопоэтических стволовых клеток (HSC), вырабатываемых в костном мозге. «Иммунные клетки» включают, например, лимфоциты (Т-клетки, В-клетки, природные клетки-киллеры (NK)) и клетки миелоидного происхождения (нейтрофилы, эозинофилы, базофилы, моноциты, макрофаги и дендритные клетки).

Термин «Т-клетка» включает все типы иммунных клеток, экспрессирующих CD3, включая хелперные Т-клетки ($CD4^+$ клетки), цитотоксические Т-клетки ($CD8^+$ клетки), регуляторные Т-клетки (Treg) и гамма-дельта Т-клетки.

Термин «цитотоксическая клетка» включает $CD8^+$ Т-клетки, природные клетки-киллеры (NK) и нейтрофилы, которые способны опосредовать цитотоксические ответы.

В настоящей заявке термин «стволовая клетка» обычно включает плюрипотентные или мультипотентные стволовые клетки. «Стволовые клетки» включают, например, эмбриональные стволовые клетки (ES); мезенхимальные стволовые клетки (MSC); индуцированные плюрипотентные стволовые клетки (iPS); и коммитированные клетки-

предшественники (гемопозитические стволовые клетки (HSC), клетки, полученные из костного мозга, и т.д.).

В настоящей заявке термин «лечение», «лечащий» и т.п. относятся к получению желательного фармакологического и/или физиологического действия. Действие может быть профилактическим с точки зрения полного или частичного предотвращения заболевания или его симптома и/или может быть терапевтическим с точки зрения частичного или полного излечения заболевания и/или нежелательного явления, связанного с заболеванием. В настоящей заявке термин «лечение» охватывает любой способ лечения заболевания у млекопитающего, например, у человека, и включает: (а) предотвращение возникновения заболевания у субъекта, который может быть предрасположен к заболеванию, но у которого данное заболевание еще не было диагностировано; (b) ингибирование заболевания, т.е. прекращение его развития; и (с) облегчение заболевания, т.е. индукцию регресса заболевания.

В настоящей заявке термины «индивидуум», «субъект», «хозяин» и «пациент» используются взаимозаменяемо и относятся к млекопитающему, включая, но не ограничиваясь ими, мышевидных грызунов (например, крыс, мышей), зайцеобразных (например, кроликов), приматов, отличных от человека, человека, собачьих, кошачьих, копытных (например, лошадей, крупный рогатый скот, овец, свиней, коз) и т.д.

«Терапевтически эффективное количество» или «эффективное количество» относится к количеству агента, или комбинированным количествам двух агентов, которое, при введении млекопитающему или другому субъекту для лечения заболевания, является достаточным для осуществления указанного лечения заболевания. «Терапевтически эффективное количество» будет варьироваться в зависимости от агента(ов), заболевания и степени его тяжести, а также от возраста, массы и т.д., субъекта, подлежащего лечению.

Термины «контроль», «контрольная реакция», «контрольное количественное исследование» и т.п. относятся к реакции, тесту или другой части экспериментальной или диагностической процедуры или дизайна эксперимента, для которых ожидаемый результат известен с высокой степенью достоверности, например, для того чтобы указать, являются ли результаты, полученные из соответствующих экспериментальных образцов надежными, указать с какой степенью достоверности соответствующие экспериментальные результаты указывают на истинный результат и/или чтобы обеспечить калибровку экспериментальных результатов. Например, в некоторых случаях, контроль может представлять собой «отрицательный контроль» так, что существенный компонент количественного исследования исключается из реакции отрицательного контроля так, что экспериментатор может иметь высокую степень уверенности в том, что

реакция отрицательного контроля не приведет к положительному результату. В некоторых случаях контроль может представлять собой «положительный контроль» так, что все компоненты конкретного количественного исследования охарактеризованы и известны, при их объединении, для получения конкретного результата в количественном исследовании, выполняемом экспериментатором, который может иметь высокую степень уверенности в том, что реакция положительного контроля не приведет к отсутствию положительного результата.

В настоящей заявке термин «праймер» или «олигонуклеотидный праймер» относится к олигонуклеотиду, который инициирует синтез комплементарной цепи нуклеиновой кислоты в условиях, когда индуцируется синтез продукта удлинения праймера, например, в присутствии нуклеотидов и агента, индуцирующего полимеризацию, такого как ДНК- или РНК-полимераза, и при подходящей температуре, рН, концентрации металлов и концентрации соли. Праймеры обычно имеют длину, совместимую с их использованием при синтезе продуктов удлинения праймера, и могут иметь длину, которая находится в диапазоне от 8 до 100 нуклеотидов, например, от 10 до 75, от 15 до 60, от 15 до 40, от 18 до 30, от 20 до 40, от 21 до 50, от 22 до 45, от 25 до 40 и т.д., включая диапазон 18-40, 20-35, 21-30 нуклеотидов и любую длину в пределах указанных диапазонов. В некоторых случаях длина праймеров может находиться в пределах диапазона от 10 до 50 нуклеотидов, например, 15-45, 18-40, 20-30, 21-25 и т.д., и праймер может иметь любую длину в пределах указанных диапазонов. Согласно некоторым вариантам реализации настоящего изобретения праймеры обычно содержат не более 10, 12, 15, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65 или 70 нуклеотидов.

Термины «гибридизоваться» и «гибридизация» относятся к образованию комплексов между нуклеотидными последовательностями, которые достаточно комплементарны, чтобы образовывать комплексы посредством спаривания оснований по Уотсону-Крику. Например, при «гибридизации» праймера с мишенью (матрицей) такие комплексы (или гибриды) достаточно стабильны, чтобы обеспечивать прайминг, необходимый, например, для инициирования синтеза ДНК ДНК-полимеразой.

Термин «биологический образец» включает множество типов образцов, полученных от индивидуума или популяции индивидуумов, и может быть использован в диагностическом, контрольном или скрининговом количественном исследовании. Определение включает кровь и другие жидкие образцы биологического происхождения, образцы плотных тканей, такие как образец биопсии или культур тканей или полученные из них клетки и их потомство. Определение также включает образцы, которые были

обработаны каким-либо образом после их получения, например, путем смешивания или объединения отдельных образцов, обработки реагентами, солубилизации или обогащения определенными компонентами, такие как клетки, полинуклеотиды, полипептиды и т.д. Термин «биологический образец» включает клинический образец, а также включает клетки в культуре, клеточные супернатанты, клеточные лизаты, сыворотку, плазму, биологическую жидкость и образцы тканей. Термин «биологический образец» включает мочу, слюну, спинномозговую жидкость, интерстициальную жидкость, глазную жидкость, синовиальную жидкость, фракции крови, такие как плазма крови и сыворотка крови, и т.п. Термин «биологический образец» также включает образцы плотных тканей, образцы культур тканей и клеточные образцы.

Термин «оценка» включает любую форму измерения и включает определение присутствия или отсутствия элемента. Термины «определение», «измерение», «исследование», «оценка» и «количественное исследование» используются взаимозаменяемо и включают количественные и качественные определения. Оценка может быть относительной или абсолютной. «Оценка присутствия» включает определение количества присутствующего объекта и/или определение его присутствия или отсутствия. В настоящей заявке термины «определение», «измерение» и «оценка» и «количественное исследование» используются взаимозаменяемо и включают количественные и качественные определения.

Настоящее изобретение будет дополнительно описано ниже, однако следует понимать, что настоящее изобретение не ограничено конкретными описанными вариантами реализации, поскольку очевидно, что они могут варьироваться. Также следует понимать, что терминология, используемая в настоящем документе, предназначена исключительно для описания конкретных вариантов реализации и не предназначена для ограничения, поскольку объем настоящего изобретения будет ограничен только прилагаемой формулой изобретения.

В тех случаях, когда представлен диапазон значений, следует понимать, что каждое промежуточное значение, до десятой части единицы нижнего предела, если из контекста явно не следует иное, между верхним и нижним пределом указанного диапазона и любым другим заявленным или промежуточным значением в указанном диапазоне, включено в область настоящего изобретения. Верхний и нижний пределы указанных меньших диапазонов могут быть независимо включены в меньшие диапазоны и также включены в область настоящего изобретения, с учетом любого конкретно исключенного предела в указанном диапазоне. В тех случаях, когда указанный диапазон включает один или оба предела, в область настоящего изобретения также включены диапазоны, исключаящие

один или оба из указанных включенных пределов.

Если не указано иное, в настоящей заявке все технические и научные термины имеют то же значение, которое обычно понятно специалисту в данной области техники, к которой относится настоящее изобретение. Несмотря на то, что любые способы и материалы, аналогичные или эквивалентные тем, которые описаны в настоящем документе, также могут быть использованы при реализации настоящего изобретения или при испытаниях настоящего изобретения, в настоящем документе описаны предпочтительные способы и материалы. Все публикации, упомянутые в настоящем документе, включены в настоящую заявку посредством ссылки для раскрытия и описания способов и/или материалов, в связи с которыми цитируются публикации.

Следует отметить, что в настоящей заявке и в прилагаемой формуле изобретения формы единственного числа «a», «an» и «the» включают множественные формы референтов, если из контекста явно не следует иное. Например, ссылка на «полипептид» включает множество таких полипептидов, и ссылка на «антиген» включает ссылку на один или более антигенов и их эквивалентов, известных специалистам в данной области техники, и т.д. Также следует отметить, что формула настоящего изобретения может быть составлена для исключения любого необязательного элемента. Следовательно, данное заявление предназначено для применения в качестве предшествующей основы для использования такой исключительной терминологии как «исключительно», «только» и т.п. в отношении описания элементов формулы настоящего изобретения или использования «отрицательного» ограничения.

Следует понимать, что некоторые признаки настоящего изобретения, которые для ясности описаны применительно к отдельным вариантам реализации, также могут быть представлены в комбинации в одном варианте реализации. И наоборот, различные признаки изобретения, которые для краткости описаны применительно к одному варианту реализации, также могут быть представлены по отдельности или в любой подходящей дополнительной комбинации. Все комбинации вариантов реализации, относящихся к настоящему изобретению, конкретно охватываются настоящим изобретением и раскрыты в настоящем документе точно так же, как если бы каждая комбинация была индивидуально и явно раскрыта. Помимо этого все дополнительные комбинации различных вариантов реализации и их элементов также конкретно включены в область настоящего изобретения и раскрыты в настоящем документе точно так же, как если бы каждая такая дополнительная комбинация была индивидуально и явно раскрыта в настоящем документе.

Публикации, обсуждаемые в настоящем документе, представлены исключительно

для их раскрытия до даты подачи настоящей заявки. Ничто из содержащегося в настоящем документе не должно быть истолковано как допущение того, что изобретение не имеет права предшествовать такому раскрытию на основании предшествующего изобретения. Помимо этого даты представленных публикаций могут отличаться от фактических дат публикаций, которые, возможно, необходимо будет подтвердить самостоятельно.

Подробное описание изобретения

В настоящем изобретении предложены библиотеки синтетических модульных полипептидов и нуклеиновые кислоты, кодирующие указанные библиотеки синтетических модульных полипептидов. В настоящем изобретении также предложены способы получения библиотек синтетических модульных полипептидов и нуклеиновых кислот, кодирующих библиотеки синтетических модульных полипептидов. В настоящем изобретении также предложены способы скрининга библиотеки синтетических модульных полипептидов для идентификации выбранного фенотипа, связанного с элементом библиотеки синтетических модульных полипептидов, причем указанные способы можно применять в количественных исследованиях в условиях *in vitro* и в условиях *in vivo*.

Библиотеки

Аспекты настоящего изобретения относятся к библиотекам нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которые кодируют множество синтетических модульных полипептидов. Аспекты настоящего изобретения также включают клеточные библиотеки, в которых каждая клетка экспрессирует отдельную нуклеиновую кислоту, содержащую нуклеиновую последовательность-штрихкод, кодирующую уникальный синтетический модульный полипептид, или множество указанных индивидуальных нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды. Как описано более подробно ниже, аспекты настоящего изобретения также включают способы получения указанных библиотек и способы скрининга указанных библиотек для детектирования конкретных фенотипов.

Библиотеки нуклеиновых кислот согласно настоящему изобретению обычно комбинаторно собраны из модульных компонентов и включают многозвенную последовательность-штрихкод, которую можно секвенировать, чтобы определить идентичность и конфигурацию модульных компонентов каждого элемента библиотеки. Следовательно, каждый отдельный элемент библиотеки нуклеиновых кислот будет содержать по меньшей мере кодирующую область и область нуклеиновой последовательности-штрихкода. В настоящей заявке термин «кодирующая область»,

поскольку он относится к отдельным элементам библиотеки нуклеиновых кислот, относится к области каждой нуклеиновой кислоты, которая кодирует синтетический полипептид, представляющий интерес, например, синтетический полипептид, содержащий один или более полипептидных модулей, которые могут быть подвергнуты скринингу для определения их влияния на конкретный желательный или нежелательный фенотип. В некоторых случаях кодирующая область может дополнительно содержать последовательности, кодирующие репортерную молекулу, например, молекулу, которая используется для детектирования и/или измерения экспрессии представляющего интерес синтетического полипептида.

Кодирующая область обычно кодирует один модульный полипептид, более подробно описанный ниже; однако настоящее описание не исключает библиотеки, в которых кодирующая область каждого элемента библиотеки кодирует два или более отдельных модульных полипептидов. Такие элементы библиотеки могут включать одну кодирующую область, которая является полицистронной (например, бицистронной, полицистронной и т.д.), например, путем включения разделимого линкера (например, участка внутренней посадки рибосомы (IRES), рибосомальной шунтирующей последовательности, нуклеиновой кислоты, кодирующей саморасщепляющийся пептид, и т.д.) между последовательностями, кодируемыми разделяемые модульные полипептиды. Кодирующая область, содержащая последовательность, кодирующую два или более отдельных модульных полипептидов, будет непрерывной и будет собрана и подвергнута скринингу в соответствии со способами, описанными в настоящем документе для одномодульных полипептидных конструкций.

Кодирующая область будет содержать переменные модули и невариабельные модули, причем предполагаемый способ применения библиотеки будет определять, является ли конкретная часть синтетического модульного полипептида переменной или невариабельной. Переменные модули кодирующей области обычно представляют собой те модули, которые подвергают скринингу, индивидуально или комбинаторно, для определения функционального свойства и/или влияния на представляющий интерес фенотип. В целом, переменные модули будут связаны с последовательностью-штрихкодом, идентифицирующей модуль, тогда как невариабельные модули не будут связаны с последовательностью-штрихкодом. Любой модуль может служить переменным или невариабельным модулем в зависимости от конструкции конкретной библиотеки и/или способа скрининга, с которым связана библиотека.

В качестве неограничивающего примера в некоторых вариантах реализации, в которых кодирующая область кодирует химерный антигенный рецептор (CAR), любой

модуль CAR может служить в качестве переменного модуля в зависимости от активности CAR, которую подвергают скринингу, включая, помимо прочего, внеклеточный домен, корегуляторный домен или первичный сигнальный домен. Например, на Фиг. 11 представлен один элемент библиотеки CAR, в которой каждый член библиотеки содержит внеклеточный домен, корегуляторный домен и первичный сигнальный домен и предоставляет неограничивающие примеры случаев, когда каждый домен может служить в качестве переменного модуля.

В некоторых случаях внеклеточный домен может быть переменным модулем, например, в тех случаях, когда интерес представляет нацеленное воздействие на антиген, в тех случаях, когда интерес представляет сила взаимодействия между антигеном и внеклеточным доменом и т.д. В некоторых случаях корегуляторный домен может быть переменным модулем, например, в тех случаях, когда индивидуальный корегуляторный домен или комбинации корегуляторных доменов должны быть подвергнуты скринингу для оценки совместной модуляции функциональной активности. В некоторых случаях первичный сигнальный домен может быть переменным модулем, например, в тех случаях, когда различные внутриклеточные пути передачи сигналов должны быть подвергнуты скринингу.

В качестве неограничивающего примера, в некоторых случаях, кодируемый синтетический модульный полипептид может представлять собой химерный полипептид рецептора Notch, содержащий внеклеточный связывающий домен, расщепляемый домен полипептида рецептора Notch (включая индуцируемый связыванием сайт расщепления) и внутриклеточный домен. Подходящие химерные полипептиды рецептора Notch и их домены включают, но не ограничиваются ими, например, те, которые описаны в заявке на патент США PCT US0201/019188; содержание которой полностью включено в настоящую заявку посредством ссылки. В некоторых случаях один или более доменов химерного полипептида рецептора Notch служат в качестве переменного домена, включая, но не ограничиваясь ими, например, внеклеточный связывающий домен, домен полипептида рецептора Notch, внутриклеточный домен, внеклеточный связывающий домен и домен полипептида рецептора Notch, внеклеточный связывающий домен и внутриклеточный домен или домен полипептида рецептора Notch и внутриклеточный домен.

Невариабельные модули, включенные в кодирующую область, будут выбраны, если это необходимо, так, чтобы все элементы библиотеки имели функцию невариабельного модуля, т.е. когда функция, которая обеспечивается невариабельным модулем, поддерживается постоянной во всех элементах библиотеки. В качестве неограничивающих примеров, в некоторых случаях, внеклеточный домен,

корегуляторный домен или первичный сигнальный домен, описанные выше, могут служить в качестве невариабельных модулей, если это необходимо, или если для проведения количественного исследования необходимо, чтобы все элементы библиотеки имели функцию внеклеточного домена, корегуляторного домена или первичного сигнального домена. В качестве неограничивающих примеров, в некоторых случаях, внеклеточный связывающий домен, расщепляемый домен полипептида рецептора Notch или внутриклеточный домен, описанные выше, могут представлять собой невариабельные модули, если это необходимо, или если для проведения количественного исследования необходимо, чтобы все элементы библиотеки имели функцию внеклеточного связывающего домена, расщепляемого домена полипептида рецептора Notch или внутриклеточного домена.

В зависимости от конкретных применяемых модулей и скринингового количественного исследования, для которых используется библиотека согласно настоящему изобретению, кодирующая область каждого элемента библиотеки может кодировать любую комбинацию переменных и невариабельных модулей, при этом простейшая кодирующая область кодирует только переменный модуль. В некоторых случаях каждый отдельный элемент библиотеки, например, каждый вектор, в который вставлен переменный модуль, может содержать, перед вставкой переменного модуля, одну или более кодирующих последовательностей так, что при вставке переменного модуля кодирующая область будет содержать переменный модуль и ранее существовавшие кодирующие последовательности, которые могут содержать невариабельный модуль и/или немодульную кодирующую последовательность или не содержать их. В некоторых случаях такая ранее существовавшая кодирующая последовательность, которая присутствует во всех элементах библиотеки, может упоминаться как «константный домен».

В некоторых случаях кодирующая область может кодировать два или более переменных модулей, включая, но не ограничиваясь ими, например, три или более переменных модулей, четыре или более переменных модулей, пять или более переменных модулей, шесть или более переменных модулей, семь или более переменных модулей, восемь или более переменных модулей и т.д. В некоторых случаях кодирующая область может кодировать два или более невариабельных модулей, включая, но не ограничиваясь ими, например, три или более невариабельных модулей, четыре или более невариабельных модулей, пять или более невариабельных модулей, шесть или более невариабельных модулей, семь или более невариабельных модулей, восемь или более невариабельных модулей и т.д. Количество переменных и

невариабельных модулей в кодирующей области может быть лимитировано практическими ограничениями способа клонирования кодирующей последовательности и конструирования библиотеки.

Количество уникальных элементов библиотеки, кодирующей уникальные синтетические модульные полипептиды, будет зависеть от количества вариабельных модулей, используемых при конструировании библиотеки и общей предполагаемой сложности библиотеки. Согласно некоторым вариантам реализации настоящего изобретения сложность библиотеки может быть описана применительно к размерности библиотеки, причем размерность библиотеки коррелирует с количеством последовательностей, кодирующих вариабельные модули, которые присутствуют в кодирующей области элементов библиотеки. Следовательно, одномерная библиотека содержит одну последовательность, кодирующую вариабельный модуль, в каждой кодирующей области каждого элемента библиотеки, двумерная библиотека содержит две последовательности, кодирующие вариабельные модули, в каждой кодирующей области каждого элемента библиотеки, трехмерная библиотека содержит три последовательности, кодирующие вариабельные модули, в каждой кодирующей области каждого элемента библиотеки, четырехмерная библиотека содержит четыре последовательности, кодирующие вариабельные модули, в каждой кодирующей области каждого элемента библиотеки и так далее.

Размерность библиотеки не обязательно должна быть однородной во всей библиотеке, и, следовательно, в некоторых случаях, библиотека может иметь смешанную размерность. Под «смешанной размерностью» подразумевается, что библиотека содержит элементы библиотеки, отличные от одномерных, описанных выше. Например, библиотека может быть частично одномерной, частично двумерной, частично трехмерной и т.д. Указанные библиотеки со смешанной размерностью могут представлять собой смешанные одно- и двумерные библиотеки, двух- и трехмерные библиотеки, одно-, двух- и трехмерные библиотеки и т.д. Представленные описания библиотек со смешанной размерностью не являются ограничивающими, и специалист в данной области техники легко поймет, что разнообразие библиотек со смешанной размерностью, включенных в область настоящего изобретения, выходит далеко за рамки явно описанных.

Следовательно, количество уникальных элементов библиотеки, кодирующих уникальные синтетические модульные полипептиды, будет различаться и будет являться результатом размерности библиотеки и количества модулей, использованных при конструировании библиотеки. Например, одномерная библиотека, сконструированная из 20 уникальных вариабельных модулей, будет содержать 20 уникальных элементов

библиотеки. Двумерная библиотека, сконструированная из 20 уникальных переменных модулей, будет содержать 20×20 (т.е. 400) уникальных элементов библиотеки. Трехмерная библиотека, сконструированная из 20 уникальных переменных модулей, будет содержать $20 \times 20 \times 20$ (т.е. 8000) уникальных элементов библиотеки. В некоторых случаях уникальный элемент библиотеки из многомерной библиотеки может содержать две или более последовательностей, кодирующих один и тот же переменный модуль.

В некоторых случаях, когда каждый элемент библиотеки содержит два или более переменных модулей, указанные модули могут иметь конкретное положение, это означает, что подмножество переменных модулей может быть помещено только в определенном месте в пределах синтетического модульного полипептида по отношению к другим переменным модулям. Например, синтетический модульный полипептид, содержащий два переменных модуля, может содержать первый набор и второй набор переменных модулей, причем модули первого набора всегда находятся в определенном положении относительно модулей второго набора. Следовательно, двумерная библиотека, сконструированная с учетом специфического положения модулей, состоящая из 30 уникальных переменных модулей, включая первый набор из 10 уникальных переменных модулей и второй набор из 20 уникальных переменных модулей, может содержать 10×20 (т.е. 200) уникальных элементов библиотеки.

Специалист в данной области техники легко поймет значительную изменчивость конфигураций библиотек, учитывая описанную выше переменную размерность и переменные количества модулей, которые могут быть объединены, с использованием специфических положений или без них, при конструировании библиотек согласно настоящему изобретению. Следовательно, общее количество уникальных элементов библиотеки будет сильно варьироваться и может находиться в пределах диапазона от менее чем 20 до 50000 и более, включая, но не ограничиваясь ими, например, 20 или более, 30 или более, 40 или более, 50 или более, 60 или более, 70 или более, 80 или более, 90 или более, 100 или более, 150 или более, 200 или более, 250 или более, 300 или более, 350 или более, 400 или более, 450 или более, 500 или более, 550 или более, 600 или более, 650 или более, 700 или более, 750 или более, 800 или более, 850 или более, 900 или более, 950 или более, 1000 или более, 1500 или более, 2000 или более, 2500 или более, 3000 или более, 3500 или более, 4000 или более, 4500 или более, 5000 или более, 5500 или более, 6000 или более, 6500 или более, 7000 или более, 7500 или более, 8000 или более, 8500 или более, 9000 или более, 9500 или более, 10000 или более, 20000 или более, 30000 и более, 40000 или более, 50000 и более и т.д.

В тех случаях, когда многомерные библиотеки сконструированы путем

объединения, общее количество уникальных элементов библиотеки в пуле может значительно варьироваться и может находиться в пределах диапазона от менее чем 20 до 50000 или более, как описано выше, и может иметь более высокую степень разнообразия, включая, но не ограничиваясь указанными, например, 10^5 или более, 10^6 или более, 10^7 или более, 10^8 или более, или 10^9 или более.

Уникальные элементы библиотек, описанные в настоящем документе, не обязательно должны быть физически упорядочены для целей конструирования библиотеки или скрининга. Как описано более подробно ниже, комбинаторная сборка компонентов нуклеиновых кислот, кодирующих каждый синтетический модульный полипептид, и совместная сборка многозвенных последовательностей-штрихкодов с помощью способа, который характеризует идентичность и ориентацию собранных модулей, обеспечивает одnoreакторный синтез и общий скрининг библиотек, описанных в настоящем документе. Следовательно, библиотека, и множество уникальных элементов библиотеки, может присутствовать в одном подходящем растворе и/или присутствовать в одном подходящем контейнере.

Синтетические модульные полипептиды

В настоящем изобретении предложены библиотеки синтетических модульных полипептидов и нуклеиновые кислоты, кодирующие синтетические модульные полипептиды. Под «модульным полипептидом» подразумевают функциональный белок, содержащий два или более функционально связанных модульных компонентов, так, что модули физически соединены и совместно функционируют в одной молекуле полипептида. Модули модульного полипептида могут иметь связанные или несвязанные функции или виды активности. Во многих вариантах реализации по меньшей мере два из модульных компонентов синтетических модульных полипептидов получены из отдельных белков. Модули, полученные из «отдельных белков», могут быть получены из различных белков, которые функционально не связаны или функционально связаны, и могут быть получены из разных молекул организма, из одной и той же молекулы организма, разных ортологичных белков, различных паралогичных белков и т.д.

Согласно некоторым вариантам реализации настоящего изобретения отдельные элементы библиотеки синтетических модульных полипептидов, описанной в настоящем документе, будут включать, как минимум, модуль, который является членом специфичной связывающейся пары (например, связывающейся пары антиген-антитело, связывающейся пары лиганд-рецептор и т.д.), и функциональный или сигнальный модуль, который участвует в индукции клеточного ответа. В некоторых случаях элемент конкретной связывающейся пары может быть упомянут в настоящем документе как внеклеточный

домен или внеклеточный распознаваемый домен. В некоторых случаях элемент конкретной связывающейся пары может относиться к белку, участвующему в сигнальном взаимодействии белок-белок, или к белку, участвующему во взаимодействии белок-липид.

Во многих вариантах реализации элемент конкретной связывающейся пары, который можно применять в отдельных элементах библиотеки, может включать антигенсвязывающий домен. Антигенсвязывающий домен, подходящий для применения в элементах библиотеки согласно настоящему изобретению, может представлять собой любой антигенсвязывающий полипептид, широкий спектр которых известен в данной области техники. В некоторых случаях антигенсвязывающий домен представляет собой одноцепочечный фрагмент Fv (scFv). Для применения подходят другие домены распознавания на основе антител (сAb V_HH (вариабельные домены антител верблюдовых) и гуманизированные варианты, IgNAR V_H (вариабельные домены антител акул) и гуманизированные варианты, sdAb V_H (вариабельные домены однодоменных антител) и вариабельные домены антител, содержащих фрагменты антител верблюдовых. В некоторых случаях для применения также подходят домены распознавания на основе Т-клеточных рецепторов (TCR), таких как одноцепочечный TCR (scTv, одноцепочечный двухдоменный TCR, содержащий V α V β).

Антигенсвязывающий домен, подходящий для применения в элементах библиотек согласно настоящему изобретению, может иметь множество видов антигенсвязывающей специфичности. В некоторых случаях антигенсвязывающий домен специфичен в отношении эпитопа, присутствующего в антигене, который экспрессируется (синтезируется) раковой клеткой, т.е. в антигене, связанном с раковой клеткой. Антиген, связанный с раковой клеткой, может представлять собой антиген, связанный, например, с клеткой рака молочной железы, В-клеточной лимфомой, клеткой лимфомы Ходжкина, клеткой рака яичников, клеткой рака предстательной железы, мезотелиомой, клеткой рака легких (например, клеткой мелкоклеточного рака легких), клеткой неходжкинской В-клеточной лимфомы (B-NHL), клеткой рака яичников, клеткой рака предстательной железы, клеткой мезотелиомы, клеткой рака легких (например, клеткой мелкоклеточного рака легких), клеткой меланомы, клеткой хронического лимфоцитарного лейкоза, клеткой острого лимфоцитарного лейкоза, клеткой нейробластомы, глиомой, глиобластомой, медуллобластомой, клеткой колоректального рака и т.д. Антиген, связанный с раковой клеткой, также может быть экспрессирован нераковой клеткой.

Неограничивающие примеры антигенов, с которыми может связываться антигенсвязывающий домен библиотеки согласно настоящему изобретению, включают,

например, CD19, CD20, CD38, CD30, Her2/neu, ERBB2, CA125, MUC-1, специфичный для предстательной железы мембранный антиген (PSMA), молекулу адгезии CD44, мезотелин, онкоэмбриональный антиген (CEA), рецептор эпидермального фактора роста (EGFR), EGFRvIII, рецептор фактора роста эндотелия сосудов 2 (VEGFR2), высокомолекулярный антиген, связанный с меланомой (HMW-MAA), MAGE-A1, IL-13R-a2, GD2 и т.п.

В некоторых случаях элемент конкретной связывающейся пары, подходящий для применения в элементах библиотек согласно настоящему изобретению, представляет собой лиганд для рецептора. Лиганды включают, но не ограничиваются ими, цитокины (например, ИЛ-13 и т.д.); факторы роста (например, херегулин, фактор роста эндотелия сосудов (VEGF) и т.п.); пептид, связывающий интегрин (например, пептид, содержащий последовательность Arg-Gly-Asp); лиганды Notch (например, Delta, Serrate, Delta-like, X-Delta, Jagged и т.д., а также их гомологи и ортологи) и т.п.

В случае если элемент конкретной связывающейся пары в элементах библиотек согласно настоящему изобретению является лигандом, то конкретный элемент библиотеки может быть активирован в присутствии второго члена конкретной связывающейся пары, причем второй элемент конкретной связывающейся пары представляет собой рецептор для лиганда. Например, если лиганд представляет собой VEGF, второй элемент специфичной связывающейся пары может представлять собой рецептор VEGF, включая растворимый рецептор VEGF. В другом примере, если лиганд является лигандом Notch, то второй член специфичной связывающейся пары может представлять собой рецептор Notch или его лигандсвязывающий участок. В качестве другого примера, если лиганд представляет собой херегулин, то второй член специфичной связывающейся пары может представлять собой Her2.

В некоторых случаях член специфичной связывающейся пары, который включен в элементы библиотеки согласно настоящему изобретению, представляет собой рецептор, например, рецептор для лиганда, корецептор и т.д. Рецептор может представлять собой лигандсвязывающий фрагмент рецептора. Подходящие рецепторы включают, но не ограничиваются ими, рецептор фактора роста (например, рецептор VEGF); полипептид члена 1 подсемейства лектиноподобных рецепторов клеток-киллеров К (NKG2D) (рецептор для MICA, MICB и ULB6); рецептор цитокинов (например, рецептор ИЛ-13, рецептор ИЛ-2 и т.д.); Her2; CD27; полипептид рецептора естественной цитотоксичности (NCR) (например, NKP30 (NCR3/CD337) (рецептор для HLA-B-ассоциированного транскрипта 3 (BAT3) и B7-H6) и т.д.); рецептор Notch (например, NOTCH1 человека, NOTCH2 человека, NOTCH3 человека, NOTCH4 человека и т.д.) и т.п.

Специфичные партнеры по связыванию также включают димерные пары.

Отдельные члены димерных пар подходят для применения в качестве элемента библиотек согласно настоящему изобретению и включают, но не ограничиваются ими, связывающиеся с димеризатором пары. Связывающиеся с димеризатором пары связываются с различными участками одной и той же молекулы (называемой в настоящем документе «димеризатор»). В присутствии димеризатора оба члена связывающейся с димеризатором пары связываются с различными участками димеризатора и, следовательно, сближаются друг с другом. Согласно некоторым вариантам реализации настоящего изобретения связывание с димеризатором является обратимым. Согласно некоторым вариантам реализации настоящего изобретения связывание с димеризатором является необратимым. Согласно некоторым вариантам реализации настоящего изобретения связывание с димеризатором является нековалентным. Согласно некоторым вариантам реализации настоящего изобретения связывание с димеризатором является ковалентным.

Другие димерные пары, подходящие для применения, включают связывающиеся с димеризатором пары, которые димеризуются при связывании первого члена димерной пары с димеризирующим агентом, причем димеризирующий агент индуцирует конформационное изменение в первом члене димерной пары, и при этом конформационное изменение позволяет первому члену димерной пары связываться (ковалентно или нековалентно) со вторым членом димерной пары. Другие димерные пары, подходящие для применения, включают димерные пары, в которых воздействие света (например, синего света) индуцирует димеризацию димерной пары.

В некоторых случаях член специфической связывающейся пары может включать белок, участвующий в сигнальном взаимодействии белок-белок или сигнальном взаимодействии белок-липид, и в этой связи синтетические модульные полипептиды, описанные в настоящем документе, могут включать один или более доменов взаимодействия белок-белок или доменов взаимодействия белок-липид. Подходящие домены взаимодействия белок-белок или домены взаимодействия белок-липид включают, но не ограничиваются ими, домен 14-3-3 (например, тот, который присутствует в структуре 2B05 PDB (база данных белковых структур RCSB, доступная на веб-сайте www.rcsb.org)), домен фактора деполимеризации актина (ADF) (например, присутствующий в структуре 1CFY PDB), домен ANK (например, присутствующий в структуре 1SW6 PDB), домен ANTH (N-концевой гомологии AP180) (например, присутствующий в структуре 5AHV PDB), домен Armadillo (ARM) (например, присутствующий в структуре 1BK6 PDB), домен BAR (Bin/Amphiphysin/Rvs) (например, присутствующая в структуре 1I4D PDB), домен BEACH («beige and CHS») (например,

присутствующий в структуре 1MI1 PDB), домены ВН (гомологии Vcl-2) (ВН1, ВН2, ВН3 и ВН4) (например, присутствующие в структуре 1BXL PDB), домен повтора бакуловируса IAP (BIR) (например, присутствующий в структуре 1G73 PDB), домен BRCT (BRCA1 C-конец) (например, присутствующий в структуре 1T29 PDB), бромдомен (например, присутствующий в структуре 1E6I PDB), домен ВТВ (BR-C, ttk и bab) (например, присутствующий в структуре 1R2B PDB), домен С1 (например, присутствующий в структуре 1PTQ PDB), домен С2 (например, присутствующий в структуре 1A25 PDB), домены привлечения каспазы (CARD) (например, присутствующие в структуре 1CWW PDB), суперспиральный домен (CC) (например, присутствующий в структуре 1QEY PDB), домен CALM (Clathrin Assembly Lymphoid Myeloid) (например, присутствующий в структуре 1HFA PDB), домен гомологичности калопонина (CH) (например, присутствующий в структуре 1BKR PDB), домен, модифицирующий организацию хроматина (Chromo) (например, присутствующий в структуре 1KNA PDB), домен CUE (например, присутствующий в структуре 1OTR PDB), домены гибели (DD) (например, присутствующие в структуре 1FAD PDB), домен гибели-эффектора (DED) (например, присутствующий в структуре 1A1W PDB), домен Discheveled, EGL-10 и плекстрин (DEP) (например, присутствующий в структуре 1FSH PDB), домен гомологии Dbl (DH) (например, присутствующий в структуре 1FOE PDB), домен EF-руки (EFh) (например, присутствующий в структуре 2PMY PDB), домен Eps15-гомологии (EH) (например, присутствующий в структуре 1EH2 PDB), домен NH2-концевой гомологии эпсина (ENTH) (например, присутствующий в структуре 1EDU PDB), домен гомологии Eps/Vasp 1 (EVH1) (например, присутствующий в структуре 1QC6 PDB), домен F-box (например, присутствующий в структуре 1FS1 PDB), домен FERM (полоса 4.1, эзрин, радиксин, моззин) (например, присутствующий в структуре 1GC6 PDB), домен FF (например, присутствующий в структуре 1UZC PDB), домен гомологии формина-2 (FH2) (например, присутствующий в структуре 1UX4 PDB), домен, ассоциированный с Forkhead (FHA) (например, присутствующий в структуре 1G6G PDB), домен FYVE (Fab-1, YGL023, Vps27 и EEA1) (например, присутствующий в структуре 1Vfy PDB), домен GAT (GGA и Tom1) (например, присутствующий в структуре 1O3X PDB), домен гомологии гелсолина (GEL) (например, присутствующий в структуре 1H1V PDB), домен GLUE (GRAM-подобный убиквитин-связывающий домен в EAP45) (например, присутствующий в структуре 2CAY PDB), домен GRAM (из глюкозилтрансфераз, Rab-подобные активаторы ГТФаз и миотубуларины) (например, присутствующий в структуре 1LW3 PDB), домен GRIP (например, присутствующий в структуре 1UPT PDB), домен глицин-тирозин-фенилаланин (GYF) (например, присутствующий в структуре 1GYF PDB), домен HEAT (хантингтон,

фактор удлинения 3, PR65/A, TOR) (например, присутствующий в структуре 1IBR PDB), домен, гомологичный С-концевому домену E6-AP (HECT) (например, присутствующий в структуре 1C4Z PDB), домен IQ (например, присутствующий в структуре 1N2D PDB), домен LIM (Lin-1, Isl-1 и Mec-3) (например, присутствующий в структуре 1QLI PDB), домен обогащенных лейцином повторов (LRR) (например, присутствующий в структуре 1YRG PDB), домен злокачественной опухоли головного мозга (MBT) (например, присутствующий в структуре 1OYX PDB), домен MH1 (гомологии Mad 1) (например, присутствующий в структуре 1OZJ PDB), домен MH2 (гомологии Mad 2) (например, присутствующий в структуре 1DEV PDB), домен MIU (взаимодействующий с убиквитином мотив) (например, присутствующий в структуре 2C7M PDB), домен NZF (цинковый палец Npl4) (например, присутствующий в структуре 1Q5W PDB), домен PAS (Peg-ARNT-Sim) (например, присутствующий в структуре 1P97 PDB), домен Phox и Bem1 (PB1) (например, присутствующий в структуре 1IPG PDB), домен PDZ (постсинаптическая плотность 95, PSD-85, discs large, Dlg, zonula occludens-1, ZO-1) (например, присутствующей в структуре 1BE9 PDB), домен гомологичности плекстрина (PH) (например, присутствующий в структуре 1MAI PDB), домен Polo-Box (например, присутствующий в структуре 1Q4K PDB), домен связывания фосфотирозина (PTB) (например, присутствующего в структуре 1SHC PDB), домен Pumilio/Puf (PUF) (например, присутствующий в структуре 1M8W PDB), домен PWWP (например, присутствующий в структуре 1KHC PDB), домен гомологии Phox (PX) (например, присутствующий в структуре 1H6H PDB), домен RGS (регулятор опосредованной G-белком передачи сигналов) (например, присутствующий в структуре 1AGR PDB), домен RING (например, присутствующий в структуре 1FBV PDB), домен SAM (стерильный альфа-мотив) (например, присутствующий в структуре 1BOX PDB), домен Shadow Chromo (SC) (например, присутствующий в структуре 1E0B PDB), домен Src-гомологии 2 (SH2) (например, присутствующий в структуре 1SHB PDB), домен Src-гомологии 3 (SH3) (например, присутствующий в структуре 3SEM PDB), домен SOCS (супрессоры передачи сигналов с участием цитокинов) (например, присутствующий в структуре 1VCB PDB), домен SPRY (например, присутствующий в структуре 2AFJ PDB), родственный стероидогенному острому регуляторному белку (StAR) домен переноса липидов (START) (например, присутствующий в структуре 1EM2 PDB), домен SWIRM (например, присутствующий в структуре 2AQF PDB), домен рецептора Toll/Ил-1 (TIR) (например, присутствующий в структуре 1FYV PDB), домен тетрапептидного повтора (TPR) (например, присутствующий в структуре 1ELW PDB), домен TRAF (факторы, связанные с рецептором фактора некроза опухолей (ФНО)) (например, присутствующий в структуре

1F3V PDB), домен tSNARE (домен SNARE (растворимый чувствительный к N-этилмалеимиду рецептор, обеспечивающий прикрепление белков (SNAP)) (например, присутствующий в структуре 1SFC PDB), домен Tubby (например, присутствующий в структуре 1I7E PDB), домен TUDOR (например, присутствующий в структуре 2GFA PDB), связанный с убиквитином (UBA) домен (например, присутствующий в структуре 1IFY PDB), домен UEV (вариант убиквитина E2) (например, присутствующий в структуре 1S1Q PDB), домен мотива, взаимодействующего с убиквитином (UIM) (например, присутствующий в структуре 1Q0W PDB), домен VHL (например, присутствующий в структуре 1LM8 PDB), домен VHS (Vps27p, Hrs и STAM) (например, присутствующий в структуре 1ELK PDB), домен WD40 (например, присутствующий в структуре 1NEX PDB), домен WW (например, присутствующий в структуре 1I6C PDB) и т.п.

Отдельные элементы библиотеки согласно настоящему изобретению, описанной в настоящем документе, могут содержать модулирующий домен. Модулирующие домены включают домены со стимулирующими и ингибирующими функциями и домены, которые модулируют активирующие и/или ингибирующие функции других «сигнальных доменов», реакции с участием которых протекают раньше, чем реакции с участием модулирующих доменов. В некоторых случаях модулирующие домены включают костимулирующие домены. В некоторых случаях модулирующие домены включают коингибирующие домены.

Модулирующий домен, пригодный для включения в элементы библиотеки согласно настоящему изобретению, может представлять собой любую функциональную единицу полипептида, длина которой равна длине аминокислотного мотива из 3 аминокислот, и длина которой равна длине полноразмерного белка, причем размер модулирующего домена ограничен только тем, что домен должен быть достаточно большим, чтобы сохранить свою функцию, и достаточно маленьким, чтобы быть совместимым с желательным способом сборки. Соответственно, модулирующий домен может иметь размер в диапазоне от 3 аминокислот до 1000 аминокислот или более, и, в некоторых случаях, может иметь в длину от приблизительно 30 аминокислот до приблизительно 70 аминокислот (а.к.), например, модулирующий домен может иметь в длину от приблизительно 30 а.к. до приблизительно 35 а.к., от приблизительно 35 а.к. до приблизительно 40 а.к., от приблизительно 40 а.к. до приблизительно 45 а.к., от приблизительно 45 а.к. до приблизительно 50 а.к., от приблизительно 50 а.к. до приблизительно 55 а.к., от приблизительно 55 а.к. до приблизительно 60 а.к., от приблизительно 60 а.к. до приблизительно 65 а.к. или от приблизительно 65 а.к. до приблизительно 70 а.к. В других случаях модулирующий домен может иметь в длину от

приблизительно 70 а.к. до приблизительно 100 а.к., от приблизительно 100 а.к. до приблизительно 200 а.к. или более 200 а.к.

В некоторых случаях «костимулирующие домены» можно применять в отдельных элементах библиотек согласно настоящему изобретению. Совместная стимуляция обычно относится к вторичному неспецифичному механизму активации, посредством которого распространяется первичная специфичная стимуляция. Примеры совместной стимуляции включают неспецифичную для антигена Т-клеточную стимуляцию после антигенспецифичной передачи сигналов через рецептор Т-клеток и неспецифичную для антигена В-клеточную совместную стимуляцию после передачи сигналов через рецептор В-клеток. Совместная стимуляция, например, совместная Т-клеточная стимуляция и связанные с ней факторы были описаны в работе Chen & Flies. *Nat Rev Immunol* (2013) 13(4):227-42, содержание которой полностью включено в настоящую заявку посредством ссылки. Костимулирующие домены обычно представляют собой полипептиды, полученные из рецепторов. Согласно некоторым вариантам реализации настоящего изобретения костимулирующие домены гомодимеризуются. Костимулирующий домен согласно настоящему изобретению может представлять собой внутриклеточную часть трансмембранного белка (т.е., костимулирующий домен может быть получен из трансмембранного белка). Неограничивающие примеры подходящих костимулирующих полипептидов включают, но не ограничиваются ими, 4-1BB (CD137), CD28, ICOS, OX-40, VTLA, CD27, CD30, GITR и HVEM. В некоторых случаях костимулирующий домен, например, используемый в элементе библиотеки согласно настоящему изобретению, может включать костимулирующий домен, перечисленный в таблице 1. В некоторых случаях костимулирующий домен отдельного элемента библиотеки содержит аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична аминокислотной последовательности костимулирующего домена, описанного в настоящем документе.

В некоторых случаях «коингибирующие домены» можно применять в отдельных элементах библиотек согласно настоящему изобретению. Подходящие коингибирующие домены обычно представляют собой полипептиды, полученные из рецепторов. Совместное ингибирование обычно относится к вторичному ингибированию первичных антигенспецифичных механизмов активации, которое предотвращает совместную стимуляцию. Совместное ингибирование, например, совместное ингибирование Т-клеток и связанные с ним факторы были описаны в работах Chen & Flies. *Nat Rev Immunol* (2013)

13(4):227-42 и Thaventhiran et al. J Clin Cell Immunol (2012) S12, содержание которых полностью включено в настоящую заявку посредством ссылки. Согласно некоторым вариантам реализации настоящего изобретения коингибирующие домены гомодимеризуются. Коингибирующий домен согласно настоящему изобретению может быть внутриклеточной частью трансмембранного белка (т.е. коингибирующий домен может быть получен из трансмембранного белка). Неограничивающие примеры подходящих коингибирующих полипептидов включают, но не ограничиваются ими, CTLA-4 и PD-1. В некоторых случаях коингибирующий домен, например, используемый в элементе библиотеки согласно настоящему изобретению, может включать коингибирующий домен, перечисленный в таблице 1. В некоторых случаях костимулирующий домен отдельного элемента библиотеки содержит аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична аминокислотной последовательности костимулирующего домена, описанного в настоящем документе.

В некоторых случаях отдельные элементы библиотеки из библиотеки синтетических модульных полипептидов могут включать модуль внутриклеточного сигнального домена. Внутриклеточные сигнальные домены, подходящие для применения в качестве модулей библиотеки согласно настоящему изобретению, включают любой желательный сигнальный домен, который обеспечивает различимый и детектируемый сигнал (например, увеличение выработки одного или более цитокинов клеткой; изменение транскрипции гена-мишени; изменение активности белка; изменение поведения клеток, например, гибель клеток, пролиферацию клеток, дифференцировку клеток, выживаемость клеток; модуляцию клеточных сигнальных реакций и т.д.) в ответ на активацию отдельного элемента библиотеки. Домен внутриклеточной сигнализации может быть ковалентно присоединен к отдельным элементам библиотеки или может не иметь такой связи, например, в тех случаях, когда все элементы библиотеки используют общий внутриклеточный сигнальный домен, внутриклеточный сигнальный домен может быть не связан с элементами библиотеки, например, может диффундировать в цитоплазме.

В некоторых случаях отдельные элементы библиотеки согласно настоящему изобретению будут включать трансмембранный домен для введения в мембрану эукариотических клеток. Трансмембранный домен может присутствовать в любом удобном месте в пределах элементов библиотек, например, на N-конце для модульных компонентов библиотеки, на C-конце для модульных компонентов библиотеки, может

быть расположен между по меньшей мере двумя модульными компонентами библиотеки (например, между антигенсвязывающим доменом и костимулирующим доменом элементов модульной библиотеки CAR) и т.д.

Любой трансмембранный (TM) домен, который обеспечивает введение полипептида в мембрану эукариотических клеток (например, млекопитающего), подходит для применения. В качестве одного неограничивающего примера, подходящей для использования является TM-последовательность IYIWAPLAGTCGVLLLSLVITLYC (SEQ ID NO: 1). Дополнительные неограничивающие примеры подходящих TM-последовательностей включают:

a) последовательности, полученные из CD8-бета: LGLLVAGVLVLLVSLGVAIHLCC (SEQ ID NO: 2);

b) последовательности, полученные из CD4: ALIVLGGVAGLLLFIGLGIFFCVRC (SEQ ID NO: 3);

c) последовательности, полученные из CD3-дзета: LCYLLDGILFIYGVILTALFLRV (SEQ ID NO: 4);

d) последовательности, полученные из CD28: WVLVVVGGLVACYSLLVTVAFIIFWV (SEQ ID NO: 5);

e) последовательности, полученные из CD134 (OX40): VAAILGLGLVLLGPLAILLALYLL (SEQ ID NO: 6); и

f) последовательности, полученные из CD7: ALPAALAVISFLLGLGLGVACVLA (SEQ ID NO: 7).

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более дополнительных модульных элементов. Например, в тех случаях, когда библиотека представляет собой библиотеку химерных антигенных рецепторов (CAR), отдельные элементы библиотеки могут быть сконфигурированы так, чтобы содержать один или более дополнительных компонентов, как описано, например, в РСТ публикации заявки на патент WO2014/127261, содержание которой полностью включено в настоящую заявку посредством ссылки.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей многодоменного каркасного белка. В настоящей заявке термин «каркасные белки» включает якорные белки и адаптерные белки. Указанные белки содержат несколько связывающих доменов, каждый из которых привлекает или закрепляет определенные элементы сигнального пути, например, закрепляет их в комплексах,

локализует их в клетке или модулирует процесс передачи сигналов (например, контролирует положительную и/или отрицательную обратную связь, стабилизирует активированные сигнальные компоненты от инактивации и т.д.). Следовательно, домены многодоменных каркасных белков, многодоменных якорных белков и многодоменных адаптерных белков обычно включают связывающие домены членов пути передачи сигналов.

Неограничивающие примеры каркасных белков и путей передачи сигналов, в которых они функционируют, включают, например, каркасный белок Ste5 пути с участием митоген-активируемой протеинкиназы (МАРК); якорные белки А-киназы (АКАР) пути передачи сигналов с участием протеинкиназы А (РКА); супрессор киназы Ras 1 (КСР) пути передачи сигналов с участием МАРК; белок В-клеточной лимфомы 10 (BCL-10) пути передачи сигналов с участием N-концевой киназы JUN (JNK) и пути передачи сигналов с участием МАРК; киназу киназы митоген-активируемой протеинкиназы 1 (МЕКК1) пути передачи сигналов с участием JNK и пути передачи сигналов с участием МАРК; АННАК-1 кальций-зависимого пути передачи сигналов; HOMER кальций-зависимого пути передачи сигналов; белки Pellino пути передачи сигналов врожденной иммунной системы; семейство NLR, содержащие пириновые домены (NLRP) белки пути передачи сигналов врожденной иммунной системы; гомолог Disks large 1 (DLG1) пути передачи сигналов с участием рецептора Т-клеток; спинофилин пути передачи сигналов дендритных клеток; и т.п. Каркасные белки также включают, но не ограничиваются ими, например, те, которые описаны в работе Buday & Tompa. *FEBS Journal* (2010) 277:4348–4355; содержание которой полностью включено в настоящую заявку посредством ссылки. В некоторых случаях каркасный белок может представлять собой белок, связанный с терминами генной онтологии (GO) «каркасный белковый комплекс» (GO: 0032947) и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к протеинкиназам, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей протеинкиназ. Протеинкиназы представляют собой белки, которые функционируют путем добавления фосфатных групп к белкам-субстратам для управления активностью белка-субстрата, ассоциацией с другими белками и/или локализацией. Протеинкиназы могут включать белки, которые связаны с терминами GO «фосфорилирование белка» (GO: 0006468), «активность протеинкиназы» (GO: 0004672), «активность сериновой/треониновой протеинкиназы» (GO: 0004674), «активность

киназы» (GO: 0016301) и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к протеинкиназам, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO. Протеинкиназы содержат один или более доменов киназы, которые осуществляют каталитическую функцию протеинкиназы. Домены протеинкиназы связаны с идентификатором Pfam PF00069, который может быть использован для получения информации о протеинкиназных доменах тех белков, которые содержат протеинкиназные домены, включая последовательности и структуры, представленные, например, на веб-сайте pfam.xfam.org.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей протеинфосфатаз. Протеинфосфатазы представляют собой белки, которые функционируют путем удаления фосфатной группы из фосфорилированного остатка аминокислоты своего белка-субстрата, что приводит к дефосфорилированию для управления активностью белка-субстрата, и выполняют функции, обратные таковым протеинкиназ. Протеинфосфатазы сгруппированы в три класса: фосфопротеиновые фосфатазы (например, PPP1CA, PPP1CB, PPP1CC, PPP2CA, PPP2CB, PPP3CA, PPP3CB, PPP3CC, PPP4C, PPP5C, PPP6C и т.д.), протеинтирозинфосфатазы (например, CDC14A, CDC14B, CDC14C, CDKN3, PTEN, SSH1, SSH2, SSH3 и т.д.) и протеинфосфатазы с двойной специфичностью (например, DUSP1, DUSP2, DUSP3, DUSP4, DUSP5, DUSP6, DUSP7, DUSP8, DUSP9, DUSP10, DUSP11, DUSP12, DUSP13, DUSP14, DUSP15, DUSP16, DUSP18, DUSP19, DUSP21, DUSP22, DUSP23, DUSP26, DUSP27, DUSP28 и т.д.); однако некоторые протеинфосфатазы остаются неклассифицированными. Протеинфосфатазы могут включать белки, которые связаны с термином GO «активность фосфатазы» (GO: 0016791) и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к протеинфосфатазам, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO. Протеинфосфатазы содержат один или более фосфатазных доменов, которые осуществляют функцию дефосфорилирования протеинфосфатазы. Домены протеинфосфатаз связаны с идентификатором Pfam PF15698, который может быть использован для получения информации о протеинфосфатазных доменах тех белков, которые содержат протеинфосфатазные домены, включая последовательности, например, представленные на веб-сайте pfam.xfam.org.

В некоторых случаях отдельные элементы библиотеки синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей белка рецепторной тирозинкиназы, также называемых рецепторными

тирозинкиназами (RTK). RTK представляют собой связанные с мембраной рецепторы клеточной поверхности. Существует подкласс тирозинкиназ, RTK, которые функционируют посредством связывания внеклеточного лиганда и последующего фосфорилирования цитоплазматической части белка. RTK можно разделить на семейства, включающие, например, семейство рецепторов эпидермального фактора роста, семейство рецепторов фактора роста фибробластов (FGFR), семейство рецепторов фактора роста эндотелия сосудов (VEGFR), семейство рецепторов RET и семейство рецепторов с дискоидиновым доменом (DDR). RTK могут включать белки, которые связаны с терминами GO «активность трансмембранной рецепторной тирозинкиназы» (GO: 0004714), «путь передачи сигналов с участием трансмембранной рецепторной тирозинкиназы» (GO: 0007169) и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к RTK, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO. RTK содержат один или более тирозинкиназных доменов, которые осуществляют киназную функцию RTK. Киназные домены RTK связаны с идентификатором Pfam PF07714, который может быть использован для получения информации о киназных доменах RTK и белках, содержащих киназные домены RTK, включая последовательности и структуры, например, представленные на веб-сайте pfam.xfam.org.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей белка липидкиназы. Белки липидкиназы фосфорилируют клеточные липиды, что приводит к модуляции реактивности липида, передаче сигнала и/или локализации липидов. Липидкиназы можно разделить на семейства, включая, например, фосфатидилинозитолкиназы и сфингозинкиназы. Липидкиназы могут включать белки, которые связаны с терминами GO «активность липидкиназы» (GO: 0001727), «фосфорилирование липидов» (GO: 0046834) и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к липидкиназам, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей белка липидфосфатазы. Белки липидфосфатаз дефосфорилируют клеточные липиды и их активность противоположна активности липидкиназ, которая приводит к модуляции реактивности липида, передаче сигнала и/или локализации липидов. Липидфосфатазы могут включать белки, которые связаны с терминами GO «активность

липидфосфатазы» (GO: 0042577), «дефосфорилирование фосфолипидов» (GO: 0046839) и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к липидфосфатазам, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей белка убиквитинилазы. Белки убиквитинилазы представляют собой белки, которые опосредуют посттрансляционную модификацию убиквитинирования, присоединение убиквитина к белку-субстрату. Убиквитинирование белка-субстрата может привести к деградации белка-субстрата, перемещению белка-субстрата, модуляции активности белка-субстрата, модуляции белок-белкового взаимодействия белка-субстрата и т.д. Убиквитинилазы могут включать белки, связанные с терминами GO «убиквитинирование белка» (GO: 0016567), «активность убиквитин-протеинтрансферазы» (GO: 0004842) и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к убиквитинилазам, включая последовательности, представленные, например, на веб-сайте www.ebi.ac.uk/QuickGO.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей белка деубиквитинилазы. Белки деубиквитинилазы представляют собой белки, которые опосредуют обращение убиквитинирования, удаление убиквитина из белка-субстрата. Деубиквитинирование белка-субстрата может обратить действие убиквитинилаз и предотвратить деградацию белка-субстрата, обратить связанное с убиквитином перемещение белка-субстрата, обратить связанную с убиквитином модуляцию активности белка-субстрата, обратить связанную с убиквитином модуляцию белок-белкового взаимодействия белка-субстрата и т.д. Деубиквитинилазы могут включать белки, которые связаны с термином GO «деубиквитинирование белка» (GO: 0016579) и синонимичными терминами, которые можно использовать для получения информации, относящейся к деубиквитинилазам, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей белка SUMOилазы. Белки SUMOилазы представляют собой белки, которые опосредуют посттрансляционную модификацию SUMOилирования, добавление небольшого убиквитинподобного модифицирующего белка (SUMO) к белку-субстрату. SUMOилирование может модулировать различные функции белка, включая стабильность

белка, ядерно-цитозольный транспорт, регуляцию транскрипции и т.д. SUMOилазы могут включать белки, связанные с термином GO «SUMOилирование белка» (GO: 0016925), и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к SUMOилазам, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей белка ацетилазы, также называемых ацетилтрансферазами. Ацетилтрансферазы представляют собой ферменты-трансферазы, которые катализируют перенос ацетильной группы на белок-субстрат и участвуют в эпигенетической и транскрипционной модуляции. Ацетилтрансферазы могут быть классифицированы на группы, включающие, например, гистонацетилтрансферазы, холинацетилтрансферазы, хлорамфениколацетилтрансферазы, серотонин-N-ацетилтрансферазы, NatA-ацетилтрансферазы, NatB-ацетилтрансферазы. Ацетилтрансферазы могут включать белки, которые связаны с термином GO «активность ацетилтрансферазы» (GO: 0016407) и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к ацетилтрансферазам, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей белка деацетилазы. Белки деацетилазы обращают действие ацетилтрансфераз и удаляют ацетильные группы, перенесенные на белок-субстрат, и, следовательно, также участвуют в эпигенетической и транскрипционной модуляции. Деацетилазы включают, например, гистондеацетилазы и сиртуины. Деацетилазы могут включать белки, которые связаны с термином GO «активность деацетилазы» (GO: 0019213) и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к деацетилазам, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей белка метилазы, также называемых метилтрансферазами. Метилтрансферазы алкилируют субстраты (включая такие субстраты как белки и нуклеиновые кислоты) путем переноса метильной группы на субстрат и участвуют в эпигенетической и транскрипционной модуляции. Метилтрансферазы можно разделить на классы на основании их структуры, включая, например, класс I, класс II, класс III, и можно

сгруппировать в соответствии с их субстратами или способом метилирования, включая, например, протеинметилтрансферазы, ДНК-метилтрансферазы, метилтрансферазы природного продукта и SAM-независимые метилтрансферазы. Метилтрансферазы могут включать белки, которые связаны с терминами GO «активность ДНК-метилтрансферазы» (GO: 0009008), «активность гистонметилтрансферазы» (GO: 0042054) и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к метилтрансферазам, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей белка деметилазы. Деметилазы обращают действие метилтрансфераз и катализируют удаление метильных групп с субстратов (включая такие субстраты как белки и нуклеиновые кислоты) и, следовательно, также участвуют в эпигенетической и транскрипционной модуляции. Деметилазы могут включать белки, которые связаны с терминами GO «активность деметилазы» (GO: 0032451), «активность ДНК-деметилазы» (GO: 0035514), «активность гистондеметилазы» (GO: 0032452) и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к деметилазам, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей белка нуклеазы. Нуклеазы катализируют расщепление фосфодиэфирных связей между нуклеотидными субъединицами нуклеиновых кислот-субстратов. Нуклеазы можно подразделить на не взаимоисключающие категории, такие как эндонуклеазы и экзонуклеазы. Нуклеазы могут включать белки, которые связаны с терминами GO «активность нуклеазы» (GO: 0004518), «активность дезоксирибонуклеазы I» (GO: 0004530), «активность рибонуклеазы в отношении гибрида РНК-ДНК» (GO: 0004523), «гидролиз фосфодиэфирной связи нуклеиновых кислот» (GO: 0090305), «активность эндонуклеазы» (GO: 0004519), «активность экзонуклеазы» (GO: 0004527) и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к нуклеазам, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей белка рекомбиназы. Рекомбиназы катализируют чувствительные к направлению

реакции обмена нуклеиновых кислот между последовательностями участков-мишеней, специфичных для каждой рекомбиназы, что приводит к вырезанию/вставке, инверсии, транслокации и обмену фрагментов нуклеиновых кислот. Примеры рекомбиназ включают, но не ограничиваются ими, рекомбиназу *Cte*, рекомбиназу *Hip*, рекомбиназу *Tge*, рекомбиназу *FLP* и т.п. Рекомбиназы могут включать белки, которые связаны с терминами GO «активность рекомбиназы» (GO: 0000150), «рекомбинация ДНК» (GO: 0006310) и синонимичными терминами, которые могут быть использованы для получения информации, относящейся к рекомбиназам, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей белка фактора транскрипции. Факторы транскрипции представляют собой белки, которые связываются с определенными последовательностями ДНК и контролируют транскрипцию. Факторы транскрипции могут являться активаторами, которые усиливают транскрипцию, или репрессорами, которые подавляют транскрипцию. Факторы транскрипции могут быть классифицированы на суперклассы на основании структуры их ДНК-связывающих доменов, включая, например, содержащие основной домен факторы транскрипции, факторы транскрипции, содержащие цинк-координирующий ДНК-связывающий домен, факторы транскрипции спираль-поворот-спираль, факторы транскрипции, содержащие бета-каркасные факторы, и те факторы транскрипции, которые не включены ни в один из вышеуказанных суперклассов (см., например, Stegmaier et al. *Genome Inform* (2004) 15(2):276-86). Факторы транскрипции содержат один или более ДНК-связывающих доменов. ДНК-связывающие домены факторов транскрипции могут представлять собой один или более ДНК-связывающих доменов, связанных с идентификаторами Pfam PF00010, PF00170, PF00172, PF00046, PF00319, PF08279, PF00096, PF00105 и т.п., которые могут быть использованы для получения информации о последовательностях и структуре доменов, например, на веб-сайте pfam.xfam.org. Факторы транскрипции также могут содержать один или более трансактивирующих доменов, один или более доменов, воспринимающих сигналы. Факторы транскрипции могут включать белки, связанные с термином GO «комплекс факторов транскрипции» (GO: 0005667), и синонимичными и родственными терминами, которые могут быть использованы для получения информации, относящейся к факторам транскрипции, включая последовательности, например, представленные на веб-сайте www.ebi.ac.uk/QuickGO.

Другие ДНК-связывающие домены, например, ДНК-связывающие домены,

полученные не из факторов транскрипции, или ДНК-связывающие домены, не принадлежащие факторам транскрипции, также можно применять в некоторых случаях в одном или более модулях отдельных элементов библиотек синтетических модульных полипептидов, описанных в настоящем документе. Подходящие ДНК-связывающие домены, полученные не из факторов транскрипции, включают природные и синтетические полипептидные домены, которые неспецифично связываются с ДНК или связываются со специфичными последовательностями ДНК. ДНК-связывающие домены, полученные не из факторов транскрипции, могут включать, но не ограничиваются ими, например, природный или модифицированный ДНК-связывающий домен из полипептида эндонуклеазы, содержащей цинковый палец, природный или модифицированный ДНК-связывающий домен из полипептида эффекторной нуклеазы, подобной активаторам транскрипции (TALEN), природный или модифицированный ДНК-связывающий домен полипептида Cas9 (включая, например, Cas9 с полной инактивацией нуклеазной активности (dCas9) и тому подобное) и т.д.

В некоторых случаях отдельные элементы библиотек синтетических модульных полипептидов могут быть сконфигурированы так, чтобы содержать один или более модулей, описанных в РСТ заявке US2004/019778, содержание которой полностью включено в настоящую заявку посредством ссылки.

Компоновка модулей в синтетические модульные полипептиды, описанные в настоящем документе, позволит получить библиотеку, содержащую множество отдельных модульных функциональных белков, которые могут, но необязательно, обладать общей функцией. Например, отдельные элементы библиотеки могут содержать или могут состоять из синтетических модульных полипептидов, которые представляют собой модульные каркасные белки, синтетических модульных полипептидов, которые представляют собой модульные рецепторные белки, синтетических модульных полипептидов, которые представляют собой модульные белки протеинкиназ или протеинфосфатаз, синтетических модульных полипептидов, которые представляют собой модульные транскрипционные регуляторные белки, синтетических модульных полипептидов, которые представляют собой модульные эпигенетические регуляторные белки, синтетических модульных полипептидов, которые представляют собой модульные белки рекомбиназ или нуклеаз и т.д. Подходящие библиотеки можно подвергать скринингу для определения желательного фенотипа в соответствии со способами, описанными в настоящем документе.

Репортеры

Библиотеки, описанные в настоящем документе, включают детектируемые

сигнальные белки, экспрессия которых кодируется последовательностями нуклеиновых кислот библиотек. Конкретные детектируемые сигнальные белки, используемые в системах библиотек, описанных в настоящем документе, будут варьироваться и частично зависеть от предпочтительного способа детектирования полученного сигнала. Например, если сигнал детектируется оптическими способами, например, с использованием флуоресцентной микроскопии или проточной цитометрии (включая флуоресцентно-активированную сортировку клеток (FACS)), то используется флуоресцентный репортер.

Подходящие детектируемые сигнальные белки включают, например, флуоресцентные белки; ферменты, которые катализируют реакцию, которая приводит к получению детектируемого сигнала в качестве продукта; эпитопные метки, поверхностные маркеры и т.п. Детектируемые сигнальные белки могут быть обнаружены непосредственно или косвенно. Например, если используется флуоресцентный репортер, то флуоресценцию репортера можно детектировать непосредственно. В некоторых случаях, если используется эпитопная метка или поверхностный маркер, то эпитопную метку или поверхностный маркер можно детектировать косвенно, например, с помощью детектируемого связывающего агента, который специфично связывается с эпитопной меткой или поверхностным маркером, например, с помощью флуоресцентномеченого антитела, которое специфично связывается с эпитопной меткой или поверхностным маркером. В некоторых случаях репортер, который обычно детектируют косвенно, например, эпитопную метку или поверхностный маркер, можно детектировать непосредственно, или репортер, который обычно детектируют непосредственно, можно детектировать косвенно, например, посредством использования детектируемого антитела, которое специфично связывается с флуоресцентным репортером.

Подходящие флуоресцентные белки включают, но не ограничиваются ими, зеленый флуоресцентный белок (GFP) или его варианты, синий флуоресцентный вариант GFP (BFP), голубой флуоресцентный вариант GFP (CFP), желтый флуоресцентный вариант GFP (YFP), улучшенный GFP (EGFP), улучшенный CFP (ECFP), улучшенный YFP (EYFP), GFPS65T, Изумруд, Топаз (TYFP), Venus, Citrine, mCitrine, GFPuv, дестабилизированный EGFP (dEGFP), дестабилизированный ECFP (dECFP), дестабилизированный EYFP (dEYFP), mCFPm, Cerulean, T-Sapphire, CyPet, YPet, mKO, HcRed, t-HcRed, DsRed, DsRed2, мономерный DsRed, J-Red, dimer2, t-dimer2(12), mRFP1, поциллопорин, Renilla GFP, Monster GFP, раGFP, белок Kaede и фотоактивируемый белок, фикобилипротеины и конъюгаты фикобилипротеинов, включая В-фикоэритрин, R-фикоэритрин и аллофикоцианин. Другие примеры флуоресцентных белков включают mHoneydew, mBanana, mOrange, dTomato, tdTomato, mTangerine, mStrawberry, mCherry,

mGrape1, mRaspberry, mGrape2, mPlum (Shaner et al. (2005) Nat. Methods 2:905-909) и т.п. Любой из множества флуоресцентных и окрашенных белков из видов Anthozoa, описанных, например, Matz et al. (1999) Nature Biotechnol. 17:969-973, подходит для применения.

Подходящие ферменты включают, но не ограничиваются ими, пероксидазу хрена (HRP), щелочную фосфатазу (AP), бета-галактозидазу (GAL), глюкозо-6-фосфатдегидрогеназу, бета-N-ацетилглюкозаминидазу, β -глюкуронидазу, инвертазу, ксантиноксидазу, люциферазу светлячков, глюкозооксидазу (GO) и т.п.

Линкеры и соединения

Соединения между модульными компонентами, кодирующими один полипептид, внутри отдельных элементов библиотеки обычно сохраняют «рамку считывания», т.е. открытая рамка считывания кодонов многомодульной кодирующей последовательности поддерживается от одного модуля до следующего. Указанные соединения могут быть упомянуты в настоящем документе как соединения с сохранением открытой рамки считывания и/или связи с сохранением открытой рамки считывания. Линкеры между модульными компонентами многомодульных полипептидов обычно будут гибкими и/или не будут содержать остатков аминокислот, которые нарушают функцию модульных доменов.

Подходящие соединения с сохранением открытой рамки считывания могут быть достигнуты, как описано более подробно ниже, с помощью любого удобного способа компоновки кодирующей последовательности и вложенной сборки модульных компонентов для включения линкера с сохранением открытой рамки считывания.

Подходящие линкеры могут быть легко выбраны и могут иметь любую подходящую длину, например, от 1 аминокислоты (например, Gly) до 20 аминокислот, от 2 аминокислот до 15 аминокислот, от 3 аминокислот до 12 аминокислот, включая от 4 аминокислот до 10 аминокислот, от 5 аминокислот до 9 аминокислот, от 6 аминокислот до 8 аминокислот или от 7 аминокислот до 8 аминокислот и могут содержать 1, 2, 3, 4, 5, 6, или 7 аминокислот.

Типичные линкеры включают полимеры глицина (G)_n, полимеры глицин-серин (включая, например, (GS)_n, (GSGGS)_n (SEQ ID NO: 8) и (GGGS)_n (SEQ ID NO: 9), где n представляет собой целое число, по меньшей мере один), полимеры глицин-аланин, полимеры аланин-серин и другие гибкие линкеры, известные в данной области техники. Подходящими для использования являются полимеры глицина и полимеры глицин-серин; Gly и Ser относительно неструктурированы и поэтому могут служить нейтральным линкером между компонентами. Подходящими для использования являются полимеры

глицина; глицин получает значительно больше конформационного пространства ϕ - ψ , даже по сравнению с аланином, и гораздо менее ограничен, чем остатки с более длинными боковыми цепями (см. Scheraga, Rev. Computational Chem. 11173-142 (1992)). Типичные линкеры могут содержать аминокислотные последовательности, включая, но не ограничиваясь ими, GGSG (SEQ ID NO: 10), GSGG (SEQ ID NO: 11), GSGSG (SEQ ID NO: 12), GSGGG (SEQ ID NO: 13), GGGSG (SEQ ID NO: 14), GSSSG (SEQ ID NO: 15) и т.п.

В некоторых случаях линкер содержит последовательность сайта распознавания фермента рестрикции BamHI для целей клонирования и как часть линкера, поскольку сайт BamHI кодирует GS.

В некоторых случаях линкерные последовательности могут быть устранены с помощью одной или более стадий клонирования (например, с использованием рестрикционных эндонуклеаз типа IIS или гомологичной рекомбинации) или путем прямого расщепления (например, расщепления ферментом рестрикции BamHI).

В некоторых случаях соединение с сохранением открытой рамки считывания достигается за счет отсутствия линкера в месте соединения двух модульных компонентов. Следовательно, соединение с сохранением открытой рамки считывания может, в некотором смысле, не содержать промежуточной нуклеиновой кислоты, кодирующей аминокислоты, между модульными кодирующими последовательностями. В некоторых случаях соединение с сохранением открытой рамки считывания может содержать две или менее пар промежуточных нуклеиновых оснований между модульными кодирующими последовательностями. В некоторых случаях соединение с сохранением открытой рамки считывания может содержать одну или более пар промежуточных нуклеиновых оснований между модульными кодирующими последовательностями. В некоторых случаях соединение с сохранением открытой рамки считывания может не содержать пар промежуточных нуклеиновых оснований между модульными кодирующими последовательностями.

Нуклеиновые последовательности-штрихкоды

Нуклеиновые последовательности-штрихкоды представляют собой специфичные уникальные последовательности нуклеиновых кислот, которые могут быть идентифицированы любым подходящим способом идентификации последовательности нуклеиновой кислоты, включая, но не ограничиваясь ими, например, идентификацию на основе гибридизации (т.е. гибридизацию в условиях *in situ*), идентификацию на основе амплификации (т.е., идентификацию на основе ПЦР) и секвенирование нуклеиновой кислоты. Нуклеиновые последовательности-штрихкоды согласно настоящему

изобретению могут представлять собой нуклеиновые последовательности-штрихкоды, специфичные для модуля, это означает, что каждый уникальный модуль многомодульной библиотеки коррелирован с конкретной уникальной нуклеиновой последовательностью-штрихкодом так, что идентификация конкретной последовательности-штрихкода эквивалентна положительной идентификации соответствующей последовательности, кодирующей модуль.

Специфичная для модуля нуклеиновая последовательность-штрихкод, описанная в настоящем документе, будет ограничена используемым способом сборки библиотеки. Например, если вложенную сборку многомодульных конструкций осуществляют с помощью способа на основе фермента рестрикции, нуклеиновые последовательности-штрихкоды будут исключать любую последовательность, которая представляет собой сайт распознавания фермента рестрикции для ферментов рестрикции, используемых при сборке. Следовательно, в некоторых случаях последовательности-штрихкоды не будут содержать последовательностей, распознаваемых рестриктазами. В некоторых случаях последовательности-штрихкоды не будут содержать последовательностей, распознаваемых рестриктазой типа PIS.

В тех случаях, когда идентификацию нуклеиновой последовательности-штрихкода и/или количественную оценку выполняют путем секвенирования, включая, например, методы секвенирования следующего поколения, будут применены стандартные критерии для нуклеиновых последовательностей-штрихкодов, детектируемых путем секвенирования. В некоторых случаях коммерчески доступные нуклеиновые последовательности-штрихкоды и/или наборы, содержащие нуклеиновые последовательности-штрихкоды и/или адаптеры нуклеиновых последовательностей-штрихкодов, могут быть использованы или модифицированы для применения в способах, описанных в настоящем документе, включая, например, нуклеиновые последовательности-штрихкоды и/или наборы адаптеров нуклеиновых последовательностей-штрихкодов коммерчески доступные от таких поставщиков как, но не ограничиваясь ими, например, New England Biolabs (Ипсвич, Массачусетс, США), Illumina, Inc. (Хейвард, Калифорния, США), Life Technologies, Inc. (Гранд-Айленд, Нью-Йорк, США), Bioo Scientific Corporation (Остин, Техас, США) и т.п., или могут быть изготовлены на заказ, например, как доступно, например, в Integrated DNA Technologies, Inc. (Коралвилл, Айова, США).

Длина нуклеиновой последовательности-штрихкода будет варьироваться и будет зависеть от сложности библиотеки и используемого метода детектирования последовательности-штрихкода. Поскольку нуклеиновые последовательности-штрихкоды

(например, последовательности-штрихкоды ДНК) хорошо известны, то дизайн, синтез и использование нуклеиновых последовательностей-штрихкодов находится в пределах компетенции специалиста в данной области техники.

В некоторых случаях длина используемых нуклеиновых последовательностей-штрихкодов также будет зависеть от вероятности того, что отдельная нуклеиновая последовательность-штрихкод случайно появится в каком-либо другом компоненте библиотеки, включая, но не ограничиваясь ими, например, кодирующую модуль последовательность, вектор, компонент вектора и т.д., или на стыке двух звеньев нуклеиновых последовательностей-штрихкодов. Например, в некоторых случаях, если существует значительная вероятность того, что последовательность, не относящаяся к нуклеиновой последовательности-штрихкоду, может быть случайно детектирована в виде последовательности-штрихкода, например, как это имеет место в очень сложных библиотеках, могут быть использованы звенья нуклеиновых последовательностей-штрихкодов большей длины. В некоторых случаях метод детектирования нуклеиновой последовательности-штрихкода может быть принят во внимание при определении необходимой длины последовательности-штрихкода, например, в тех случаях, когда нуклеиновую последовательность-штрихкод детектируют путем гибридизации, могут быть использованы более длинные нуклеиновые последовательности-штрихкоды, по сравнению с тем случаем, когда применяют секвенирование с использованием определенных праймеров для секвенирования.

В настоящей заявке термин «область нуклеиновой последовательности-штрихкода», поскольку он применяется по отношению к отдельным элементам библиотеки нуклеиновых кислот, относится к области каждой нуклеиновой кислоты, которая содержит последовательность нуклеиновой кислоты, специфичную для переменной модульной части синтетического полипептида. В некоторых случаях область нуклеиновой последовательности-штрихкода можно использовать для специфичной идентификации переменной модуля или модулей, присутствующих в кодирующей области конкретного элемента библиотеки. В некоторых случаях область нуклеиновой последовательности-штрихкода можно использовать для специфичной идентификации переменной модуля(ей) и идентификации порядка переменных модулей, присутствующих в кодирующей области конкретного элемента библиотеки (т.е. архитектуры). В некоторых случаях область нуклеиновой последовательности-штрихкода можно использовать для количественной оценки, например, полуколичественной оценки, частоты конкретного элемента библиотеки или частоты конкретного модуля в популяции, содержащей множество элементов библиотеки.

В пределах синтетического модульного полипептида, содержащего нуклеиновую последовательность-штрихкод, библиотеки, описанной в настоящем документе, звенья нуклеиновых последовательностей-штрихкодов в области последовательностей-штрихкодов будут иметь обратную ориентацию по отношению к последовательностям, кодирующим модули. Помимо этого звенья нуклеиновых последовательностей-штрихкодов в области последовательностей-штрихкодов будут расположены в обратном порядке по отношению к соответствующей им последовательности, кодирующей модуль. Однако по отношению к кодируемому полипептиду область нуклеиновых последовательностей-штрихкодов может быть расположена в направлении «N-конца» или «С-конца» относительно последовательности нуклеиновой кислоты, кодирующей синтетический модульный полипептид.

Специфичные для вектора элементы

Под «специфичными для вектора элементами» понимают элементы, которые используют при получении, конструировании, распространении, поддержании и/или количественном исследовании вектора до, во время или после его конструирования. Подходящие специфичные для вектора элементы включают, но не ограничиваются ими, например, элементы вектора, необходимые для распространения, клонирования и выбора вектора во время его использования, и могут включать, но не ограничиваются ими, например, сайт инициации репликации, сайт множественного клонирования, прокариотический промотор, фаговый промотор, селективный маркер (например, ген устойчивости к антибиотикам, кодируемый ферментный белок, кодируемый флуоресцентный или хромогенный белок и т.д.) и т.п. В векторах, описанных в настоящем документе, можно использовать любые подходящие специфичные для вектора элементы, если это необходимо.

Подходящие промоторные и энхансерные элементы, используемые в качестве специфичных для вектора элементов, известны в данной области техники. Промоторы, подходящие для экспрессии в бактериальной клетке включают, но не ограничиваются ими, *lacI*, *lacZ*, T3, T7, *gpt*, *lambda P* и *trc*. Промоторы, подходящие для экспрессии в эукариотической клетке, включают, но не ограничиваются ими, промоторные и энхансерные элементы гена легкой и/или тяжелой цепей иммуноглобулина; немедленный ранний промотор цитомегаловируса; промотор тимидинкиназы вируса простого герпеса; ранние и поздние промоторы SV40; промотор, присутствующий в длинных концевых повторах ретровируса; промотор металлотioneина-1 мыши; и различные тканеспецифичные промоторы, известные в данной области техники.

Подходящие обратимые промоторы, включая обратимые индуцибельные

промоторы, известны в данной области техники. Подходящие обратимые промоторы могут быть выделены и получены из многих организмов, например, эукариотов и прокариотов. Модификация обратимых промоторов, полученных из первого организма для применения во втором организме, например, первый организм является прокариотом, и второй организм является эукариотом, первый организм является эукариотом, и второй организм является прокариотом и т.д., хорошо известна в данной области техники. Подходящие обратимые промоторы и системы, основанные на подходящих обратимых промоторах, при этом также содержащие дополнительные контрольные белки, включают, но не ограничиваются ими, промоторы, регулируемые спиртом (например, промотор гена алкогольдегидрогеназы I (alcA), промоторы, реагирующие на белки, трансактивируемые спиртом (AlcR) и т.д.), регулируемые тетрациклином промоторы (например, промоторные системы, включая TetActivators, TetON, TetOFF и т.д.), регулируемые стероидами промоторы (например, промоторные системы рецептора глюкокортикоидов крысы, промоторные системы рецептора эстрогенов человека, регулируемые ретиноидами промоторные системы, системы промоторов щитовидной железы, регулируемые экдизоном промоторные системы, регулируемые мифепристоном промоторные системы и т.д.), регулируемые металлом промоторы (например, промоторные системы металлотионеина и т.д.), регулируемые патогенезом промоторы (например, промоторы, регулируемые салициловой кислотой, промоторы, регулируемые этиленом, промоторы, регулируемые бензотиадиазолом и т.д.), регулируемые температурой промоторы (например, индуцируемые тепловым шоком промоторы (например, HSP-70, HSP-90, промотор белка теплового шока сои и т.д.), регулируемые светом промоторы, синтетические индуцибельные промоторы и т.п.

В некоторых случаях локус или конструкцию или трансген, содержащий подходящий промотор, необратимо переключают посредством индукции индуцируемой системы. Подходящие системы для индукции необратимого переключения хорошо известны в данной области техники, например, индукция необратимого переключения может быть основана на использовании рекомбинации, опосредованной Cre-lox (см., например, работу Fuhrmann-Benzakein, et al., PNAS (2000) 28:e99, содержание которой полностью включено в настоящую заявку посредством ссылки). Любая подходящая комбинация рекомбиназы, эндонуклеазы, лигазы, сайтов рекомбинации и т.д., известная в данной области техники, может быть использована для создания промотора с необратимым переключением. Способы, механизмы и требования к выполнению сайтспецифичной рекомбинации, описанные в другом месте в настоящем документе, могут быть использованы для получения промоторов с необратимым переключением и

хорошо известны в данной области техники, см., например, Grindley et al. (2006) *Annual Review of Biochemistry*, 567-605 и Tropp (2012) *Molecular Biology* (Jones & Bartlett Publishers, Sudbury, MA), содержание которой включено в настоящую заявку посредством ссылки.

Способы получения библиотек

В настоящем изобретении предложены способы получения библиотек нуклеиновых кислот, кодирующих синтетические модульные полипептиды, причем каждая нуклеиновая кислота библиотеки содержит многозвенную нуклеиновую последовательность-штрихкод, которая идентифицирует вариабельный модуль (модули) модульного полипептида и, если присутствуют несколько вариабельных модулей, ориентацию модулей друг относительно друга. В многочисленных вариантах реализации настоящего изобретения предложены способы поэтапной комбинаторной сборки синтетических модульных полипептидов из кодирующих модули нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, так, что каждая из полученных нуклеиновых кислот, кодирующих синтетические модульные полипептиды, содержит область, кодирующую модули, встроенную с сохранением открытой рамки считывания, и многозвенную нуклеиновую последовательность-штрихкод, при этом расположение звеньев нуклеиновой последовательности-штрихкода соответствует расположению модулей, встроенных с сохранением открытой рамки считывания. В целом, в тех случаях, когда в каждый элемент библиотеки включено несколько вариабельных модулей, совместно собранная многозвенная нуклеиновая последовательность-штрихкод обеспечивает характеристику сборки вариабельных модулей каждого элемента библиотеки.

Не желая быть связанными соответствием какой-либо теории, авторы настоящего изобретения полагают, что поэтапная комбинаторная сборка синтетических модульных полипептидов, представленных в настоящем документе, обеспечивает конструирование более крупных и более сложных библиотек, чем было бы возможно или осуществимо с использованием стандартных способов конструирования на основе индивидуальных (т.е. «по одному») полипептидов. Авторы настоящего изобретения выявили, что стандартные способы конструирования синтетических модульных полипептидов представляют собой значительное техническое препятствие для высокопроизводительного скрининга синтетических модульных полипептидов. Скоординированная сборка каждого многомодульного синтетического полипептида и соответствующей многозвенной нуклеиновой последовательности-штрихкода, предложенная в настоящем документе, позволяет преодолеть указанное препятствие и обеспечивает «сборку в одном реакторе»

из множества уникальных нуклеиновых кислот, кодирующих синтетические модульные полипептиды, подходящие для скрининга. Упомянутая высокопроизводительная сборка в одном реакторе не может быть выполнена с использованием стандартных способов конструирования полипептидов.

Авторы настоящего изобретения также выявили, что стандартные, сгруппированные физически, синтетические полипептидные библиотеки также представляют собой серьезные технические препятствия для высокопроизводительного скрининга. В частности, при скрининге больших библиотек число физически разделенных реакционных камер и изменчивость условий количественного исследования между каждой камерой представляют значительные технические препятствия при попытке скрининга всей сложной большой библиотеки и анализа полученных данных. Авторы настоящего изобретения выявили, что скрининг объединенной библиотеки может решить указанную проблему, однако имеет другие значительные препятствия, такие как трудность или нецелесообразность идентификации и/или количественного определения отдельных полипептидов, обеспечивающих желательный фенотип при скрининге из сложного смешанного пула уникальных модульных полипептидов. Использование короткой многозвенной нуклеиновой последовательности-штрихкода устраняет эту проблему, обеспечивая апостериорную эффективную положительную идентификацию и/или количественную оценку отдельных уникальных синтетических модульных полипептидов, которые обеспечивают желательный фенотип в сложном пуле, путем секвенирования только многозвенной нуклеиновой последовательности-штрихкода.

Стратегии клонирования

В целом, в настоящем изобретении предложен способ получения библиотек, описанных в настоящем документе, путем вложенной сборки кодирующих полипептидные модули нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды. Например, как показано на Фиг. 12, вектор нуклеиновой кислоты (100), содержащий последовательность, кодирующую полипептидный модуль (101), соединенную со специфичной для модуля нуклеиновой последовательностью-штрихкодом (102), линеаризуют (103) путем расщепления связи между последовательностью, кодирующей полипептидный модуль, и специфичной для модуля нуклеиновой последовательностью-штрихкодом. После линеаризации вектора, содержащего первую кодирующую модуль последовательность и первую специфичную для модуля нуклеиновую последовательность-штрихкод, нуклеиновую кислоту, содержащую вторую кодирующую модуль последовательность (104) и вторую нуклеиновую последовательность-штрихкод (105), специфичную для второго модуля,

вставляют (т.е. встраивают) между первой кодирующей модуль последовательностью и первой нуклеиновой последовательностью-штрихкодом (106). Собранный нуклеиновый кислотный остаток содержит кодирующую область, кодирующую синтетический модульный полипептид и область нуклеиновой последовательности-штрихкода, которая содержит многозвенную нуклеиновую последовательность-штрихкод (т.е. «нуклеиновые последовательности-штрихкоды» (BC)). В некоторых случаях последовательности, кодирующие модули, собирают так, чтобы они соединялись с помощью последовательности, кодирующей желатиновый линкер, описанный в настоящем документе. В некоторых случаях последовательности, кодирующие модули, собирают так, чтобы они соединялись без использования линкерной последовательности, например, без линкерной последовательности, кодирующей одну или более линкерных аминокислот, без каких-либо промежуточных некодирующих нуклеотидов между первой и второй последовательностями, кодирующими модули, и т.д. Следовательно, обозначение «желатиновый линкерная последовательность» или «линкер», в частности, использованное в чертежах, включает полное отсутствие любого линкера или линкерной последовательности и прямое соединение полипептидных модулей и последовательностей, кодирующих модули.

Линеаризация векторной последовательности путем расщепления участка между первой кодирующей модуль последовательностью и первой специфичной для модуля нуклеиновой последовательностью-штрихкодом может быть достигнута с помощью любого удобного и подходящего средства. Например, полинуклеотид, содержащий кодирующую модуль последовательность и нуклеиновую последовательность-штрихкод, может быть сконфигурирован так, чтобы содержать сайт расщепления фермента рестрикции (т.е. рестрикционной эндонуклеазы) между кодирующей модуль последовательностью и нуклеиновой последовательностью-штрихкодом. Сайт расщепления может представлять собой сайт расщепления фермента рестрикции типа II и, более конкретно, может представлять собой сайт расщепления, который содержится в последовательности, распознаваемой используемым ферментом рестрикции типа II. Любой удобный фермент рестрикции типа II, который осуществляет расщепление в пределах своей последовательности распознавания, можно применять для описанной цели, включая ферменты рестрикции типа II, которые осуществляют расщепление в пределах своих последовательностей распознавания, которые известны в данной области техники. В некоторых случаях сайт расщепления может представлять собой сайт расщепления BamHI, который имеет последовательность распознавания 5'-GGATCC-3' и расщепляет нити после первого G последовательности распознавания на обеих нитях,

оставляя липкий конец 5'-GATC-3'. Другие ферменты рестрикции, которые могут быть использованы для указанной цели, включают, но не ограничиваются ими, например.

В некоторых случаях вложенную сборку осуществляют путем использования сайтов двух ферментов рестрикции (RE1 и RE2), расположенных между первой кодирующей модуль последовательностью и первой специфичной для модуля нуклеиновой последовательностью-штрихкодом, и фланкирующих вторую кодирующую модуль последовательность, содержащую нуклеиновую последовательность-штрихкод (Фиг. 13). В случае стратегии сборки, основанной на использовании двух ферментов рестрикции, расщепление в двух сайтах RE1 и RE2 приводит к линеаризации вектора и высвобождению фрагмента, содержащего вторую кодирующую модуль последовательность, содержащую нуклеиновую последовательность-штрихкод. После лигирования линеаризованного вектора и высвобожденного фрагмента желательная линкерная последовательность присутствует между первой и второй последовательностями, кодирующими модули, и первая и вторая нуклеиновые последовательности-штрихкоды имеют обратную ориентацию по отношению к первой и второй кодирующим модули последовательностям.

В некоторых случаях вложенную сборку осуществляют путем использования сайтов четырех ферментов рестрикции (RE1, RE2, RE3 и RE4), причем RE1 и RE2 расположены между обеими кодирующими модули последовательностями и соответствующими им нуклеиновыми последовательностями-штрихкодами, тогда как RE3 и RE4 фланкируют вторую кодирующую модуль последовательность, содержащую нуклеиновую последовательность-штрихкод (Фиг. 14). В случае стратегии сборки с использованием четырех ферментов рестрикции отдельное расщепление первого вектора в сайтах RE1 и RE2 и второго вектора в сайтах RE3 и RE4 приводит к линеаризации вектора и высвобождению фрагмента, содержащего вторую кодирующую модуль последовательность, содержащую нуклеиновую последовательность-штрихкод. После лигирования линеаризованного вектора и высвобожденного фрагмента желательная линкерная последовательность присутствует между первой и второй последовательностями, кодирующими модули, тогда как первая и вторая нуклеиновые последовательности-штрихкоды имеют обратную ориентацию по отношению к первой и второй кодирующим модули последовательностям.

В некоторых случаях ферменты рестрикции намеренно выбирают так, что после лигирования соединения между первой и второй кодирующими модули последовательностями и первой и второй нуклеиновыми последовательностями-штрихкодами не содержат последовательностей распознавания ферментов рестрикции

RE1, RE2, RE3 или RE4. В целом, в соответствии с указанной стратегией использованные сайты RE1, RE2, RE3 и RE4 инактивированы после лигирования так, что полученный в результате вектор содержит только активные сайты RE1 и RE2 между второй кодирующей модуль последовательностью и соответствующей ей нуклеиновой последовательностью-штрихкодом. Следовательно, данная стратегия обеспечивает последовательную вложенную сборку не только двумерной конструкции путем повторной линейаризации полученного вектора с помощью расщепления ферментами рестрикции RE1 и RE2 кодирующей модуль последовательности, содержащей нуклеиновую последовательность-штрихкод, которая была вставлена в последнюю очередь.

В случае способа сборки с использованием четырех ферментов рестрикции RE1 выбирают так, чтобы после расщепления полученный конец линейаризованного вектора был совместимым (т.е. был способен к лигированию) с концом высвобожденного фрагмента, полученного с помощью расщепления ферментом рестрикции RE3. RE2 выбирают так, чтобы после расщепления полученный конец линейаризованного вектора был совместимым (т.е. мог быть лигирован и/или был способен к полной или по меньшей мере частичной гибридизации) с концом высвобожденного фрагмента, полученного при расщеплении с помощью фермента рестрикции RE4. В некоторых случаях один или более концов, полученных в результате расщепления с помощью RE1, RE2, RE3 или RE4, могут быть модифицированы, например, путем добавления или удаления одного или более нуклеотидов или другой химической модификации для получения концов, совместимых между RE1 и RE3 или RE2 и RE4. Любой удобный способ концевой модификации можно применять для получения совместимых концов, включая способы концевой модификации, которые хорошо известны в данной области техники, включая, но не ограничиваясь ими, например, создание тупых концов, фосфорилирование, дефосфорилирование и т.д.

В некоторых случаях вложенную сборку осуществляют путем использования последовательности распознавания фермента рестрикции типа IIS (RE1), причем два сайта RE1 присутствуют между первой кодирующей модуль последовательностью и соответствующей ей нуклеиновой последовательностью-штрихкодом, и два сайта RE1 фланкируют вторую кодирующую модуль последовательность, содержащую нуклеиновую последовательность-штрихкод (Фиг. 15). В случае указанной стратегии, опосредованной ферментом рестрикции типа IIS, расщепление в сайтах, смежных с сайтами распознавания RE1, приводит к линейаризации вектора и высвобождению фрагмента, содержащего вторую кодирующую модуль последовательность, содержащую нуклеиновую последовательность-штрихкод. Сайты расщепления, смежные с сайтами распознавания RE1, сконфигурированы так, что после расщепления 3'-конец вектора

совместим с 5'-концом фрагмента, и 5'-конец вектора совместим с 3'-концом фрагмента. В некоторых случаях указанные совместимые концы могут называться совместимыми «липкими концами». Следовательно, при лигировании линейаризованного вектора и высвобожденного фрагмента желательная линкерная последовательность присутствует между первой и второй кодирующими модули последовательностями, тогда как первая и вторая нуклеиновые последовательности-штрихкоды имеют обратную ориентацию по отношению к первой и второй кодирующим модули последовательностям.

Любой удобный фермент рестрикции типа II^S можно применять в стратегиях сборки с использованием указанных ферментов, описанных в настоящем документе, включая, но не ограничиваясь ими, например, AceIII, AcuI, AlwI, AarI, BbsI, BbvI, BbvII, BccI, Vce83I, VceAI, VcefI, VciVI, BfuAI, BmrI, BmuI, BpmI, BpuEI, BsaI, BsbI, BseRI, BsgI, BslFI, BsmAI, BsmFI, BsoMAI, BspCNI, BspGI, BspMI, BspNCI, BspQI, BsrDI, Bst7II, BtgZI, BtsCI, BtsI, BveI, DrdII, EarI, EciI, FagI, FinI, FokI, HgaI, Hin4II, HphI, HpyAV, LguI, MboII, MmeI, MnlI, NmeAIII, PleI, SapI, SfaNI, SgeI и т.п.

В некоторых случаях фермент рестрикции, использованный, например, при линейаризации вектора и/или высвобождении фрагмента, содержащего второй модуль, представляет собой фермент рестрикции, который расщепляет цепи по обе стороны от своего сайта распознавания. Любой удобный фермент рестрикции, обладающий указанной активностью, можно применять в способах сборки, описанных в настоящем документе, включая, но не ограничиваясь им, например, VcgI.

В некоторых случаях вложенную сборку осуществляют путем использования нескольких последовательностей распознавания ферментов рестрикции типа II^S, например, двух последовательностей распознавания ферментов рестрикции типа II^S (RE1 и RE2), причем два сайта RE1 присутствуют между кодирующими модули последовательностями и соответствующими им нуклеиновыми последовательностями-штрихкодами, и два сайта RE2 фланкируют вторую кодирующую модуль последовательность, содержащую нуклеиновую последовательность-штрихкод (Фиг. 16). В случае стратегии сборки с использованием двух ферментов рестрикции типа II^S раздельное расщепление первого вектора в сайтах RE1 и второго вектора в сайтах RE2 приводит к линейаризации вектора и высвобождению фрагмента, содержащего вторую кодирующую модуль последовательность, содержащую нуклеиновую последовательность-штрихкод. После лигирования линейаризованного вектора и высвобожденного фрагмента желательная линкерная последовательность присутствует между первой и второй кодирующими модули последовательностями, тогда как первая и вторая нуклеиновые последовательности-штрихкоды имеют обратную ориентацию по

отношению к первой и второй кодирующим модули последовательностям.

В целом, в соответствии с данной стратегией сайты RE1 и RE2, использованные в первом раунде линеаризации и высвобождения фрагментов, не сохраняются в векторе и встраиваются так, что полученный вектор содержит только активные сайты RE1 между второй кодирующей модуль последовательностью и соответствующей ей нуклеиновой последовательностью-штрихкодом. Следовательно, данная стратегия обеспечивает последовательную вложенную сборку не только двумерной конструкции путем повторной линеаризации полученного вектора с помощью расщепления ферментами рестрикции сайтов распознавания RE1 в кодирующей модуль последовательности, содержащей нуклеиновую последовательность-штрихкод, которая была вставлена в последнюю очередь.

В данном способе сборки с использованием фермента рестрикции типа PS сайты расщепления, смежные с последовательностями распознавания RE1 и RE2, сконфигурированы так, что после расщепления полученные концы линеаризованного вектора являются совместимыми (т.е. способны к лигированию) с концами высвобожденного фрагмента. В некоторых случаях один или более концов, полученных при расщеплении RE1 и/или RE2, могут быть модифицированы, например, путем добавления или удаления одного или более нуклеотидов или другой химической модификации для получения совместимых концов. Любой удобный способ модификации концов можно применять для получения совместимых концов, включая способы модификации концов, которые хорошо известны в данной области техники, включая, но не ограничиваясь ими, например, получение тупых концов, фосфорилирование, дефосфорилирование и т.д.

В некоторых случаях совместимые концы могут быть получены с помощью фермента с экзонуклеазной активностью, например, экзонуклеазы. Например, в некоторых вариантах реализации первый вектор и второй вектор или фрагмент нуклеиновой кислоты сконфигурированы так, что при расщеплении первого вектора первым ферментом рестрикции (RE1) и второго вектора или нуклеиновой кислоты вторым ферментом рестрикции (RE2) полученные новые концы совместимы для лигирования после использования фермента с экзонуклеазной активностью. Способы получения совместимых концов с использованием фермента с экзонуклеазной активностью хорошо известны в данной области техники и включают, но не ограничиваются ими, например, клонирование с помощью системы In-Fusion[®], сборку по Гибсону и т.п.

В некоторых случаях сборку первого вектора, расщепленного ферментом рестрикции, и второй нуклеиновой кислоты, например, второго вектора, содержащего

фрагмент нуклеиновой кислоты, осуществляют с помощью реакции In-Fusion[®], и, в таких случаях, совместимые концы, полученные в результате использования фермента с экзонуклеазной активностью, могут быть упомянуты как липкие концы In-Fusion[®] (липкие концы IF) (Фиг. 17). Соединение двух липких концов IF может быть сконфигурировано так, чтобы получить желательные линкеры и/или желательные линкерные последовательности между двумя соединенными концами нуклеиновой кислоты. Например, соединенные липкие концы IF между первой кодирующей модуль последовательностью и второй кодирующей модуль последовательностью могут быть сконфигурированы так, что первый и второй модули расположены с сохранением открытой рамки считывания. Как описано более подробно в настоящем документе, кодирующие модули последовательности, которые расположены с сохранением открытой рамки считывания, могут быть разделены последовательностью, кодирующей одну или более линкерных аминокислот, или могут не иметь такой последовательности. Например, как показано на Фиг. 17, расщепление первого вектора в сайте распознавания (RE1) первого фермента рестрикции приводит к получению концов, которые могут быть совмещены посредством реакции In-Fusion[®] с концами, полученными при расщеплении второго вектора или фрагмента нуклеиновой кислоты ферментом рестрикции типа PIS в сайтах расщепления, определенных сайтами распознавания (RE2), присутствующими на втором векторе или фрагменте нуклеиновой кислоты. Совместимые концы лигируют посредством реакции In-Fusion[®], что приводит к получению желательной линкерной последовательности между первой и второй кодирующими модули последовательностями и многозвенной нуклеиновой последовательностью-штрихкодом, содержащей первое звено и второе звено нуклеиновых последовательностей-штрихкодов, которые имеют обратную ориентацию по отношению к первой и второй кодирующим модули последовательностям. В некоторых случаях после сборки In-Fusion[®] первой и второй кодирующих модули последовательностей и многозвенной нуклеиновой последовательности-штрихкода сайт фермента рестрикции присутствует между кодирующей областью и областью последовательности-штрихкода, чтобы обеспечить возможность введения дополнительных последовательностей, кодирующих модули, и звеньев нуклеиновых последовательностей-штрихкодов.

В некоторых случаях после лигирования нуклеиновой кислоты, содержащей первую кодирующую модуль последовательность, со второй нуклеиновой кислотой, содержащей вторую кодирующую модуль последовательность, первая и вторая кодирующие модули последовательности соединены так, что между первой и второй кодирующими модули последовательностями отсутствует линкер и промежуточные

некодирующие нуклеотиды. В одном неограничивающем примере первый вектор, содержащий первую кодирующую модуль последовательность и соответствующую ей нуклеиновую последовательность-штрихкод, соединен без линкера или промежуточных некодирующих нуклеотидов со вторым вектором или фрагментом нуклеиновой кислоты, содержащим вторую кодирующую модуль последовательность и соответствующую ей нуклеиновую последовательность-штрихкод, с помощью расщепления ферментом рестрикции (Фиг. 18). Первый вектор и второй вектор или фрагмент нуклеиновой кислоты расщепляют с использованием двух разных ферментов рестрикции типа II, имеющих два разных сайта распознавания (RE1 и RE2), причем первый фермент рестрикции расщепляет обе нити нуклеиновой кислоты в одном и том же положении на некотором расстоянии от своего сайта распознавания (RE1), оставляя «тупой конец», тогда как второй фермент рестрикции расщепляет обе нити нуклеиновой кислоты в разных положениях на некотором расстоянии от своего сайта распознавания (RE2), оставляя выступающий конец или «липкий конец». Исходные нуклеиновые кислоты сконфигурированы так, что после расщепления первого вектора и второго вектора вторым ферментом рестрикции образующиеся липкие концы являются совместимыми для лигирования. Следовательно, после лигирования полученных тупых концов и липких концов полученный вектор содержит первую кодирующую модуль последовательность, гибридизованную непосредственно, без линкера или промежуточных некодирующих нуклеиновых кислот, со второй кодирующей модуль последовательностью и многозвенной нуклеиновой последовательностью-штрихкодом, содержащей нуклеиновую последовательность-штрихкод первой кодирующей модуль последовательности и нуклеиновую последовательность-штрихкод второй кодирующей модуль последовательности в обратной ориентации по отношению к первой и второй кодирующим модули последовательностям.

Любой удобный фермент рестрикции типа II, который приводит к получению тупых концов после расщепления, можно применять в способах, описанных в настоящем документе, с использованием лигирования тупых концов, включая ферменты, которые хорошо известны в данной области техники, включая, но не ограничиваются ими, например, S₁I, M₁I и т.д. Помимо этого при необходимости можно использовать способы получения тупых концов после расщепления с помощью рестрикционной эндонуклеазы, которая не образует тупые концы, т.е. может быть использовано «затупление», при необходимости, включая, но не ограничиваясь перечисленным, «заполнение конца» ДНК-полимеразой, такой как, например, большой фрагмент ДНК-полимеразы I (например, фрагмент Кленова), ДНК-полимеразой T4, нуклеазой золотистой

фасоли и т.д., или концевые неспаренные нуклеотиды могут быть удалены ферментом с экзонуклеазной активностью.

В некоторых случаях, когда последовательность, совместимая с ферментом рестрикции, не относящимся к типу *PS*, присутствует на конце кодирующей модуль последовательности, расщепление может быть выполнено с помощью фермента рестрикции, не относящегося к типу *PS*, например, фермента рестрикции типа *P*, который расщепляет нити внутри своей последовательности распознавания. В таких случаях может быть использован фермент рестрикции, который образует тупой конец на конце кодирующей модуль последовательности, если кодирующая модуль последовательность содержит всю или часть последовательности распознавания фермента рестрикции на своем 3'-конце или 5'-конце. В некоторых случаях, когда кодирующая модуль последовательность содержит всю или часть последовательности распознавания фермента рестрикции, который разрезает нити внутри своей последовательности распознавания и не образует тупой конец, может быть использован фермент рестрикции, не образующий тупых концов, и полученные липкие концы могут быть затуплены, например, любым удобным способом, включая, но не ограничиваясь ими, например, способы, описанные в настоящем документе.

Поскольку последовательности на концах кодирующих модули последовательностей будут ограничены концевыми аминокислотами модуля, то случаи, когда подходящие последовательности распознавания фермента рестрикции присутствуют в удобном положении, или могут быть модифицированы для этого, на одном или нескольких концах кодирующей модуль последовательности, чтобы обеспечить надлежащее расщепление и/или образование тупого конца, могут быть редкими в зависимости от конкретного используемого модуля. Следовательно, во многих случаях эффективное получение большой или многомерной библиотеки синтетических модульных полипептидов будет зависеть от сайтов распознавания ферментов, которые присутствуют за пределами кодирующей модуль последовательности. Следовательно, в некоторых случаях одна или более, включая все, из последовательностей распознавания ферментов рестрикции, используемых при сборке библиотеки, будут присутствовать за пределами кодирующей модуль последовательности. Во многих случаях одна или более, включая все, из последовательностей распознавания ферментов рестрикции, используемых при сборке библиотеки, будут присутствовать за пределами нуклеиновой последовательности-штрихкода.

В некоторых случаях получают синтетический модульный полипептид, в котором кодирующие модули последовательности легко соединяются с концами линкерного

домена так, что между кодирующими модули последовательностями и соответствующими им соединениями с линкерным доменом отсутствует промежуточная последовательность. В одном неограничивающем примере указанное непрерывное соединение кодирующих модули последовательностей с линкерным доменом облегчается вектором или фрагментом нуклеиновой кислоты, сконфигурированным так, чтобы содержать часть желательной линкерной последовательности, легко присоединяемой к любому концу кодирующей модуль последовательности. Например, первый вектор, сконфигурированный так, чтобы содержать первую кодирующую модуль последовательность, легко присоединенную к первой части линкерного домена, лигируют со вторым вектором или фрагментом нуклеиновой кислоты, сконфигурированным так, чтобы содержать вторую часть линкерного домена, последовательно присоединенную ко второй кодирующей модуль последовательности (Фиг. 19). Первый вектор может быть сконфигурирован так, чтобы содержать два сайта распознавания первого фермента рестрикции типа IIS (RE1) между кодирующей модуль последовательностью и соответствующей ей нуклеиновой последовательности-штрихкодом. Второй вектор, или фрагмент нуклеиновой кислоты, может быть сконфигурирован так, чтобы содержать два сайта распознавания второго фермента рестрикции типа IIS (RE2), фланкирующие вторую кодирующую модуль последовательность, содержащую нуклеиновую последовательность-штрихкод. Может быть использован третий вектор или фрагмент нуклеиновой кислоты, содержащий оставшуюся, например, «среднюю» часть линкерного домена, которая фланкирована двумя сайтами распознавания второго фермента рестрикции типа IIS (RE2). Последовательность предварительно расщепленных векторов и/или фрагментов нуклеиновых кислот конфигурируют так, чтобы после расщепления получить совместимые концы между первой частью линкерного домена и средней частью линкерного домена, второй частью линкерного домена и средней частью линкерного домена и двумя нуклеиновыми последовательностями-штрихкодами. После расщепления и лигирования полученный вектор содержит первую кодирующую модуль последовательность, последовательно соединенную с линкерным доменом, последовательно соединенным со второй кодирующей модуль последовательностью, соединенной с областью нуклеиновой последовательности-штрихкода, содержащей первую и вторую нуклеиновые последовательности-штрихкоды в обратной ориентации по отношению к первой и второй кодирующим модули последовательностям (Фиг. 19). Последовательная сборка, например, между модулями и линкерными доменами, может быть достигнута с использованием опосредованной экзонуклеазами сборки (например, клонирование с помощью системы In-Fusion[®], сборка по Гибсону и т.д.) или без нее.

Вышеописанные стратегии сборки на основе расщепления, в некоторых случаях, можно комбинировать, полностью или частично, любым удобным и подходящим путем, чтобы обеспечить способ, пригодный для получения библиотеки синтетических модульных полипептидов, описанной в настоящем документе. Помимо этого, в тех случаях, когда альтернативные способы сборки на основе расщепления известны в данной области техники и будут совместимы с описанными в настоящем документе способами, подходящие альтернативные способы можно применять при сборке библиотеки синтетических модульных полипептидов, описанной в настоящем документе. В некоторых случаях описанные стратегии на основе расщепления можно комбинировать, полностью или частично, со способами, не связанными с расщеплением, для сборки нуклеиновых кислот.

Сборка нуклеиновых кислот, кодирующих библиотеку синтетических модульных полипептидов, описанную в настоящем документе, не ограничена стратегиями сборки на основе расщепления, т.е. сборки на основе ферментов рестрикции. В некоторых случаях способы, не связанные с расщеплением, можно применять при сборке библиотеки, описанной в настоящем документе, включая, но не ограничиваясь ими, например, стратегии на основе амплификации, стратегии на основе рекомбинации и т.д. Подходящими для применения являются способы, не связанные с расщеплением, вместо стратегий на основе расщепления, т.е. вся стратегия сборки в целом не связана с расщеплением ферментами рестрикции или может быть использована в комбинации со стратегиями на основе расщепления, т.е., так, что стратегия сборки в целом включает как способы, основанные на использовании ферментов рестрикции, так и способы, не связанные с расщеплением ферментами рестрикции.

В некоторых случаях библиотека синтетических модульных полипептидов, описанная в настоящем документе, может быть собрана полностью или частично с использованием способа сборки на основе амплификации, включая, но не ограничиваясь ими, например, клонирование ПЦР, клонирование ТА, удлинение концов методом ПЦР и т.п. Подходящие стратегии на основе амплификации будут варьироваться, но обычно будут основаны на использовании множества сайтов связывания праймеров в исходных векторах и/или фрагментах нуклеиновых кислот. Подходящие сайты связывания праймеров, в зависимости от желательного конечного продукта, могут быть специально добавлены к векторам и/или фрагментам нуклеиновых кислот, или ранее существовавшая последовательность, присутствующая в векторе или фрагменте нуклеиновой кислоты, может быть использована в качестве сайта связывания праймера в соответствии с различными стратегиями сборки, основанными на амплификации. Сайты связывания

праймеров могут быть расположены в любой удобной конфигурации и/или ориентации, достаточной для получения желательного клонированного продукта, включая, но не ограничиваясь ими: расположение между кодирующей модуль последовательностью и соответствующей ей нуклеиновой последовательностью-штрихкодом в направлении от 5'-конца к 3'-концу относительно последовательности, кодирующей модуль; расположение между кодирующей модуль последовательностью и соответствующей ей нуклеиновой последовательностью-штрихкодом в направлении от 5'-конца к 3'-концу относительно последовательности-штрихкода; расположение перед (т.е. в направлении 5'-конца) кодирующей модуль последовательностью в направлении от 5'-конца к 3'-концу относительно кодирующей модуль последовательности; расположение после (т.е. в направлении 3'-конца) кодирующей модуль последовательности в направлении от 5'-конца к 3'-концу относительно кодирующей модуль последовательности, содержащей нуклеиновую последовательность-штрихкод; и т.п. Последовательности сайтов связывания праймеров могут быть сконфигурированы так, что после клонирования на основе амплификации, в том числе после одного раунда клонирования на основе амплификации, одна или более желательных линкерных последовательностей присутствуют между собранными элементами, включая, например, собранные последовательности, кодирующие модули, собранные нуклеиновые последовательности-штрихкоды и т.д.

В качестве неограничивающего примера стратегий сборки на основе амплификации может быть использована стратегия амплификации липких концов методом ПЦР (Фиг. 20), при которой первый вектор содержит первый сайт связывания праймера (PBS1) и второй сайт связывания праймера (PBS2), расположенный, в ориентации, противоположной ориентации от 5'-конца к 3'-концу, между кодирующей модуль последовательностью и соответствующей ей нуклеиновой последовательностью-штрихкодом. Может быть использован второй вектор или фрагмент нуклеиновой кислоты, содержащий вторую кодирующую модуль последовательность, содержащую нуклеиновую последовательность-штрихкод, фланкированную первым и вторым сайтами связывания праймеров (PBS1 и PBS2). После удлинения и амплификации с помощью удлинения липких концов методом ПЦР с использованием праймеров, которые специфично гибридизуются с сайтами PBS1 и PBS2, синтезируется собранный продукт, в котором первая кодирующая модуль последовательность соединена с помощью желательного линкера со второй кодирующей модуль последовательностью, которая соединена с многозвенной нуклеиновой последовательностью-штрихкодом, содержащей первую и вторую нуклеиновые последовательности-штрихкоды в обратной ориентации по

отношению к соответствующим им кодирующим модули последовательностям (Фиг. 20).

Ввиду вышеописанных стратегий сборки специалист в данной области техники легко поймет, как любая из вышеописанных стратегий может быть комбинирована, полностью или частично, чтобы привести к желаемому результату и/или обеспечить наибольшие преимущества и/или свести к минимуму недостатки конкретных методик клонирования в соответствии со сборкой с использованием желательной библиотеки и/или компонента библиотеки. В качестве неограничивающего примера, стратегии на основе амплификации можно комбинировать со стратегиями на основе расщепления, когда комбинирование указанных стратегий приводит к сборке желательной библиотеки и/или компонентов библиотеки. Например, как показано на Фиг. 21, расщепление ферментом рестрикции первого вектора в соответствии с сайтом распознавания фермента рестрикции (RE1), расположенным между первой кодирующей модуль последовательностью и соответствующей ей нуклеиновой последовательностью-штрихкодом, можно комбинировать с амплификацией на основе ПЦР с использованием сайтов связывания праймеров, фланкирующих вторую кодирующую модуль последовательность, содержащую нуклеиновую последовательность-штрихкод, содержащуюся в пределах второго вектора или фрагмента нуклеиновой кислоты, чтобы обеспечить вложенную сборку. После сборки с применением описанной гибридной стратегии сборки может быть получен собранный продукт, в котором первая кодирующая модуль последовательность соединена с помощью желательного линкера со второй кодирующей модуль последовательностью, которая соединена с многозвенной нуклеиновой последовательностью-штрихкодом, содержащей первую и вторую нуклеиновые последовательности-штрихкоды в обратной ориентации по отношению к соответствующим им кодирующим модули последовательностям (Фиг. 21).

Гибридные стратегии не ограничиваются комбинированием стратегий, основанных на расщеплении и амплификации, и могут включать другие способы клонирования и/или синтетической биологии, полностью или частично, включая, но не ограничиваясь ими, например, стратегии клонирования на основе рекомбинации (включая, но не ограничиваясь ими, например, стратегии клонирования на основе технологии Gateway, клонирования на основе рекомбиназы Cre/Flp (включая, стратегии, при которых сайт инактивирован при рекомбинации) и т.д.), сборку последовательности *de novo*, синтез нуклеиновой кислоты *de novo* и т.п. В некоторых случаях стратегии клонирования на основе рекомбинации, сборки последовательности *de novo*, синтеза нуклеиновой кислоты *de novo* и т.п. могут быть использованы независимо, т.е. по отдельности в качестве независимой стратегии клонирования, а не как часть гибридной стратегии клонирования.

В некоторых случаях, когда используемая конкретная стратегия клонирования приводит к присутствию нежелательной промежуточной последовательности между двумя клонированными элементами, также известной как шрамы или швы клонирования, такие шрамы клонирования могут быть уменьшены и/или удалены путем мономолекулярного расщепления и повторного лигирования вектора, содержащего шрам. Мономолекулярное расщепление и повторное лигирование можно осуществлять любым удобным способом, включая, но не ограничиваясь ими, например, мономолекулярное расщепление, опосредованное ферментами рестрикции, и повторное лигирование, такое как, например, мономолекулярное расщепление ферментами рестрикции типа II^S и повторное лигирование. Например, в некоторых случаях, шов, который содержит полный сайт рекомбинации или его часть, в результате сборки на основе рекомбинации, может быть частично или полностью удален путем мономолекулярного расщепления и повторного лигирования. В некоторых случаях шрам после расщепления, который содержит весь сайт распознавания фермента рестрикции или его часть в результате стратегии на основе расщепления, может быть частично или полностью удален путем мономолекулярного расщепления и повторного лигирования. В некоторых случаях шов, который содержит весь сайт связывания праймера или его часть в результате сборки на основе амплификации, может быть частично или полностью удален путем мономолекулярного расщепления и повторного лигирования.

В одном неограничивающем варианте реализации комбинаторная библиотека получена путем итеративного клонирования компонентов библиотеки каждого типа размерности посредством клонирования на основе расщепления и клонирования с помощью системы In-Fusion[®]. Например, как показано на Фиг. 22, на первом этапе исходную нуклеиновую кислоту (например, вектор), содержащую кодирующую scFv последовательность и смежный линкер Gly/Ser, расщепляют в сайте BamHI в пределах или рядом с линкером Gly/Ser. На втором этапе фрагмент нуклеиновой кислоты, содержащий первую кодирующую модуль последовательность, соединенную со вторым линкером Gly/Ser, гибридным с первой специфичной для модуля нуклеиновой последовательностью-штрихкодом, клонируют в расщепленный сайт BamHI путем клонирования с помощью системы In-Fusion[®]. После лигирования путем клонирования с помощью системы In-Fusion[®] линкер Gly/Ser между scFv и первой кодирующей модуль последовательностью сохраняется. На третьем этапе нуклеиновую кислоту, собранную на этапе 2, расщепляют в сайте BamHI, импортированном в пределах или рядом с линкером Gly/Ser фрагмента первой кодирующей модуль нуклеиновой кислоты, и повторяют клонирование с помощью системы In-Fusion[®] со вторым фрагментом нуклеиновой

кислоты, содержащим вторую кодирующую модуль последовательность, соединенную с третьим линкером Gly/Ser, гибридованным со второй нуклеиновой последовательностью-штрихкодом, специфичной для модуля. Как и на этапе 2, после лигирования путем клонирования с помощью системы In-Fusion[®] линкер Gly/Ser между первой и второй кодирующими модули последовательностями сохраняется. Если необходимы элементы библиотеки большей размерности, этап 3 можно повторять итеративно. На четвертом этапе расщепление в сайте BamHI, импортированном в пределах или рядом с линкером Gly/Ser, кодирующей модуль последовательности, содержащей нуклеиновую кислоту, которая была добавлена в последнюю очередь, позволяет осуществлять клонирование In-Fusion[®] концевой репортерной последовательности (например, последовательности, кодирующей GFP, которая также содержит последовательность сигнала «стоп» (например, стоп-кодон)) между готовой кодирующей модуль последовательностью и нуклеиновой последовательностью-штрихкодом, расположенной после нее. В этом варианте реализации каждый готовый элемент библиотеки содержит комбинаторный CAR, встроенный с сохранением открытой рамки считывания, и комбинаторную нуклеиновую последовательность-штрихкод, описывающую архитектуру комбинаторного CAR, встроенного с сохранением открытой рамки считывания.

Объединенные библиотеки

В настоящем изобретении предложены способы получения объединенных библиотек синтетических модульных полипептидов, содержащих нуклеиновые последовательности-штрихкоды. Подразумевают, что в объединенной библиотеке элементы библиотеки присутствуют в общем контейнере и/или общем растворе, и отдельные элементы библиотеки не должны быть физически разделены в пространстве, например, отдельные элементы библиотеки могут быть объединены во время конструирования библиотеки (например, как при «сборке в одном реакторе») или после создания отдельных элементов библиотеки. Например, в случае объединенной библиотеки синтетических модульных полипептидов, сконструированной путем комбинаторной вложенной сборки, компоненты библиотеки могут быть объединены во время сборки элементов библиотеки, до завершения сборки элементов библиотеки и/или после завершения сборки элементов библиотеки и т.д. В настоящей заявке объединенные библиотеки не ограничены библиотеками только синтетических модульных полипептидов, но также включают объединенные библиотеки клеток, экспрессирующих синтетические модульные полипептиды, объединенные библиотеки нуклеиновых кислот, кодирующих синтетические модульные полипептиды, и т.п.

Соответственно, в некоторых случаях, отдельные компоненты нуклеиновых кислот, используемые для сборки элементов библиотеки, можно комбинировать до сборки и они могут оставаться объединенными во время сборки элементов библиотеки. В других случаях собранные элементы библиотеки могут быть смешаны после сборки, чтобы получить объединенную библиотеку.

Поскольку библиотеки и способы сборки библиотек, описанные в настоящем документе, включают получение элементов библиотеки, которые могут быть идентифицированы путем секвенирования соответствующей области нуклеиновой последовательности-штрихкода, отдельные нуклеиновые кислоты могут быть объединены в любой момент до, во время или после сборки, при этом сохраняется возможность последующей идентификации и количественного определения отдельных элементов библиотеки в количественных исследованиях. В некоторых случаях конечную размерность библиотеки может контролировать путем объединения компонентов библиотеки в определенные моменты во время сборки. Например, библиотека смешанной размерности может быть получена путем раздельной сборки частичных библиотек разной размерности и последующего объединения частичных библиотек (например, с использованием сборки «разделить-и-объединить»). В некоторых случаях, например, после объединения частичных библиотек различной размерности, может быть выполнена последующая сборка, включая добавление дополнительных переменных доменов.

В одном неограничивающем варианте реализации, как показано на Фиг. 23, объединенную библиотеку нуклеиновых кислот, кодирующих синтетические модульные полипептиды CAR, трансформируют в общую популяцию первичных Т-клеток человека для получения первичных Т-клеток человека, экспрессирующих синтетический модульный полипептид CAR. Необязательно, если кодируемые элементы библиотеки синтетических модульных полипептидов CAR содержат один или более репортерных модулей, трансформированные Т-клетки могут быть отсортированы на основании экспрессии ими репортерного модуля (который указывает на экспрессию синтетического модульного полипептида CAR), чтобы выделить трансформированные клетки с одинаковыми уровнями экспрессии элементов библиотеки синтетических модульных полипептидов CAR. Указанные отсортированные трансформированные Т-клетки с одинаковыми уровнями экспрессии представляют собой объединенную библиотеку Т-клеток, экспрессирующих синтетические модульные полипептиды CAR.

В некоторых случаях модулируют трансформационную эффективность нуклеиновых кислот и/или первичных Т-клеток человека, экспрессирующих синтетические модульные полипептиды CAR. Такая модуляция может быть выполнена по

различным практическим причинам, например, для контроля вероятности экспрессии каждого элемента библиотеки и/или для контроля вероятности экспрессии каждой клеткой не более чем одного элемента библиотеки. Указанная модуляция может быть достигнута любым удобным способом, включая, но не ограничиваясь ими, например, модулирование исходного количества нуклеиновой кислоты, кодирующей синтетический модульный полипептид, присутствующей во время трансформации, контроль исходного количества первичных Т-клеток, присутствующих во время трансформации, контроль соотношения кодирующей нуклеиновой кислоты и первичных Т-клеток во время трансформации и т.п. Следовательно, в некоторых случаях, эффективность трансформации модулируют так, что по существу каждая трансдуцированная Т-клетка экспрессирует один уникальный синтетический модульный полипептид CAR. В некоторых случаях готовые объединенные библиотеки клеток, экспрессирующих уникальные синтетические модульные полипептиды CAR, можно культивировать, размножать, сохранять и т.д. в соответствии с требованиями последующих количественных исследований.

Подходящие объединенные библиотеки, независимо от того, содержат они объединенные нуклеиновые кислоты, объединенные модульные полипептиды, объединенные трансдуцированные клетки и т.д., не ограничены нуклеиновыми кислотами, кодирующими синтетические модульные полипептиды CAR, синтетическими модульными полипептидами CAR или клетками (например, Т-клетками), экспрессирующими синтетические модульные полипептиды CAR. Любая библиотека нуклеиновых кислот, кодирующих модульные полипептиды, модульные полипептиды и/или соответствующие трансдуцированные клетки, экспрессирующие модульные полипептиды, полученные в соответствии со способами, описанными в настоящем документе, могут быть объединены аналогичным способом, чтобы получить объединенную библиотеку, которую можно применять, например, при скрининговых исследованиях.

Отдельные элементы объединенных библиотек, описанных в настоящем документе, могут быть положительно идентифицированы благодаря специфичной многозвенной нуклеиновой последовательности-штрихкоду, связанной с каждым членом библиотеки. Благодаря специфичной комбинаторной вложенной сборке, которая приводит к предсказуемой позиционной взаимосвязи между каждой кодирующей модуль последовательностью и соответствующей ей нуклеиновой последовательностью-штрихкодом, идентичность и архитектура каждого синтетического модульного полипептида может быть установлена путем простого секвенирования соответствующей

многозвенной нуклеиновой последовательности-штрихкода. Следовательно, в некоторых случаях, идентичность и/или архитектура отдельных элементов библиотеки синтетических модульных полипептидов может быть определена на основании данных о последовательности, связанной с областями нуклеиновых последовательностей-штрихкодов элементов библиотеки. Например, в некоторых случаях, все элементы и/или каждый отдельный элемент объединенной библиотеки, описанной в настоящем документе, могут быть определены, например, после конструирования библиотеки, после использования библиотеки в конкретном количественном исследовании и т.д.

Компартментализованные компоненты библиотеки

Отдельные нуклеиновые кислоты-элементы, содержащие кодирующую область и область нуклеиновой последовательности-штрихкода, в некоторых случаях могут быть компартментализованы. Объединенные библиотеки и библиотеки с компартментализованными компонентами не обязательно должны быть взаимоисключающими, и, например, в некоторых случаях библиотека может быть объединена и может содержать компартментализованные компоненты в разные моменты времени, например, библиотека может быть сконструирована, например, элементы библиотеки могут быть собраны, в виде пула, например, как при сборке в одном реакторе, и затем может быть впоследствии компартментализована, например, путем трансфекции элементов библиотеки нуклеиновых кислот в отдельные клетки или неклеточные компартменты. В других случаях элементы библиотеки могут быть компартментализованы в ходе их сборки и затем объединены для последующих способов, включая, но не ограничиваясь ими, например, дальнейшую сборку или скрининг. В целом, нуклеиновые кислоты, кодирующие синтетические полипептиды, содержащие нуклеиновые последовательности-штрихкоды, собранные в соответствии со способами вложенной сборки и стратегиями клонирования, описанными в настоящем документе, собирают в виде пула и впоследствии могут подвергаться компартментализации или могут оставить без изменений.

Компартментализация собранных нуклеиновых кислот, кодирующих синтетические полипептиды, содержащих нуклеиновые последовательности-штрихкоды, может быть достигнута любым подходящим способом, обеспечивающим транскрипцию и трансляцию отдельных элементов библиотек нуклеиновых кислот так, что каждый элемент библиотеки нуклеиновых кислот остается связанным с продуктом, который он кодирует. В некоторых случаях, как описано более подробно ниже, компартментализация может быть достигнута путем создания клеточных библиотек, в которых отдельные клетки служат в качестве «компартмента» и обеспечивают трансляцию и транскрипцию

элементов библиотек нуклеиновых кислот.

В некоторых случаях компартиментализация может быть достигнута посредством создания неклеточных библиотек, например, библиотек на основе инкапсуляции. Бесклеточные библиотеки на основе инкапсуляции обычно содержат эмульсию двух или более несмешивающихся жидкостей, причем элементы библиотеки нуклеиновых кислот растворимы в первой жидкости, например, в водной жидкости, такой как вода или водный буфер, и нерастворимы во второй жидкости, например, в масле или другом органическом растворителе, так, что первая жидкость образует компартменты, например, капли, содержащие отдельные элементы библиотеки. Эмульсия, содержащая библиотеку, может быть сконфигурирована так, что каждый компартмент содержит любое количество отдельных элементов библиотеки, включая, но не ограничиваясь этим, не более одного элемента в компартменте. Нуклеиновые кислоты-элементы из инкапсулированных библиотек могут быть транскрибированы (например, в условиях *in vitro*) и транслированы (например, в условиях *in vitro*) в условиях, в которых продукты транскрипции и трансляции остаются связанными, т.е. остаются в пределах компартмента, с отдельным элементом библиотеки нуклеиновых кислот, кодирующим их. Любой удобный и подходящий способ получения инкапсулированной библиотеки полипептидов, кодируемых нуклеиновыми кислотами, может быть использован в способах, описанных в настоящем документе, включая, но не ограничиваясь ими, например, способы, описанные в Bernath et al. *Anal Biochem.* (2004) 325(1):151-7; содержание которой полностью включено в настоящую заявку посредством ссылки.

После получения кодируемого продукта элемента библиотеки внутри компартмента, элемент библиотеки нуклеиновых кислот и кодируемый синтетический модульный полипептид могут быть, но необязательно, физически соединены. Например, в некоторых случаях, элемент библиотеки нуклеиновых кислот и кодируемый продукт могут быть соединены, например, с помощью любого удобного и подходящего способа, включая химическую связь или конъюгацию (т.е., путем получения ковалентной связи между элементом библиотеки нуклеиновых кислот и кодируемым синтетическим модульным полипептидом) или посредством молекулярного связывания (например, путем прямого связывания элемента библиотеки с кодируемым синтетическим модульным полипептидом или посредством непрямого связывания элемента библиотеки с кодируемым продуктом, которое опосредовано одним или более связывающими промежуточными продуктами (например, партнерами по связыванию, субстратами и т.д.)). Любой удобный и подходящий способ соединения компартиментализованных кодируемых полипептидов с элементами библиотеки нуклеиновых кислот можно применять в

способах, описанных в настоящем документе, включая, но не ограничиваясь ими, использование, например, субстрата, содержащего присоединенный связывающий агент для эпитопной метки, который специфично связывается с эпитопной меткой, кодируемой элементом библиотеки нуклеиновых кислот, например, как описано в Griffiths & Tawfik. EMBO J (2003) 22(1):24-35, содержание которой полностью включено в настоящую заявку посредством ссылки. В некоторых случаях после связывания элементы библиотек нуклеиновых кислот могут быть компартиментализованы, например, объединены, и затем количественно исследованы как объединенная библиотека.

В некоторых случаях элемент библиотеки нуклеиновых кислот и кодируемый продукт остаются связанными в достаточной степени, не будучи физически соединенными, например, путем компартиментализации внутри компартмента, включая клеточные и неклеточные компартменты.

Компартиментализация, как клеточная, так и неклеточная, обычно позволяет идентифицировать синтетический модульный полипептид или его часть, который коррелирует с детектированным фенотипом посредством секвенирования области нуклеиновой последовательности-штрихкода отдельного элемента библиотеки нуклеиновых кислот, кодирующего синтетический модульный полипептид, который остается связанным с синтетическим модульным полипептидом на основании их компартиментализации или в результате физической связи, образующейся во время их компартиментализации.

Клеточные библиотеки

Как описано более подробно выше, в настоящей заявке библиотеки включают клеточные библиотеки, в которых клетки библиотеки экспрессируют синтетические модульные полипептиды. Трансформация нуклеиновых кислот, кодирующих синтетические модульные полипептиды, может быть осуществлена любым удобным способом, включая, но не ограничиваясь ими, например, вирусную трансфекцию, электропорацию, липофекцию, бомбардировку, химическую трансформацию, использование подходящего для трансдукции носителя (например, подходящего для трансдукции белка-носителя) и т.п. В некоторых случаях клетка, в которую трансформируют нуклеиновую кислоту, кодирующую синтетический модульный полипептид, упоминается в настоящем документе как клетка-хозяин.

Клетки-хозяева могут экспрессировать одну отдельную нуклеиновую кислоту, содержащую нуклеиновую последовательность-штрихкод, кодирующую уникальный синтетический модульный полипептид, или могут экспрессировать множество, включая две или более, отдельных нуклеиновых кислот, содержащих нуклеиновую

последовательность-штрихкод, кодирующих уникальные синтетические модульные полипептиды. Следует понимать, что количество отдельных нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которое экспрессирует клетка-хозяин, можно контролировать, например, путем контроля частоты или вероятности доставки рассматриваемых нуклеиновых кислот в клетку-хозяина, например, путем модуляции параметров способа доставки. В некоторых случаях конечное количество отдельных нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, которые кодируют уникальные синтетические модульные полипептиды, присутствующих в клетке-хозяине, может быть упомянуто как множественность инфекции (МОИ) и может быть определено как соотношение нуклеиновых кислот и клеток-хозяев до или после доставки. Стандартные способы модуляции МОИ, например, путем увеличения или уменьшения соотношения нуклеиновых кислот и клеток-хозяев до доставки, могут быть использованы для получения желательного конечного количества отдельных нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, кодирующих уникальные синтетические модульные полипептиды, на клетку-хозяина после доставки.

Нуклеиновые кислоты, кодирующие синтетические модульные полипептиды, могут быть трансформированы в любую подходящую клетку-хозяина или клеточную линию, включая, например, прокариотические и эукариотические клетки. Выбор типа клетки-хозяина будет зависеть от ряда факторов, включая тип библиотеки синтетических модульных полипептидов, подлежащей скринингу, и конкретное скрининговое количественное исследование. В некоторых случаях клетка-хозяин может представлять собой прокариотическую клетку, включая, но не ограничиваясь ими, *Acidobacteria*, *Actinobacteria*, *Aquificae*, *Bacteroidetes*, *Caldiserica*, *Chlamydiae*, *Chlorobi*, *Chloroflexi*, *Chrysiogenetes*, *Cyanobacteria*, *Deferribacteres*, *Deinococcus-Thermus*, *Dictyoglomi*, *Elusimicrobia*, *Fibrobacteres*, *Firmicutes*, *Fusobacteria*, *Gemmatimonadetes*, *Lentisphaerae*, *Nitrospirae*, *Planctomycetes*, *Proteobacteria*, *Spirochaetes*, *Synergistetes*, *Tenericutes*, *Thermodesulfobacteria*, *Thermotogae* и *Verrucomicrobia*. Согласно некоторым вариантам реализации настоящего изобретения клетка-хозяин может представлять собой бактериальную клетку, например, *E. coli*. В некоторых случаях может быть использован стандартный бактериальный штамм, включая, но не ограничиваясь ими, например, коммерчески доступные штаммы от таких поставщиков как Американская коллекция типовых культур (АТСС) (Манассас, Виргиния, США), Life Technologies, Inc. (Гранд-Айленд, Нью-Йорк, США) и т.п.

Подходящие эукариотические клетки включают первичные клетки и

культивируемые клетки, первоначально полученные из животного-хозяина, включая, но не ограничиваясь ими, например, млекопитающих (включая, например, человека, приматов, человекообразных обезьян, копытных, собачьих, кошачьих, кроликов, грызунов и т.д.), рептилий, земноводных (например, шпорцевых лягушек, саламандру, тритона и т.д.), рыб (например, данио и т.д.), птиц (например, курицу и т.д.), беспозвоночных (например, насекомых (например, плодovou муху и т.д.)), червей (например, нематод и т.д.), морских беспозвоночных (например, морского ежа и т.д.) и т.д.), дрожжи и т.п. Согласно некоторым вариантам реализации настоящего изобретения клетки могут представлять собой первичные клетки грызунов или культивированные клетки грызунов, полученные от мыши или крысы. В других вариантах реализации клетки могут представлять собой первичные клетки человека или культивированные клетки человека. Любая подходящая эукариотическая клетка может быть использована в качестве клетки-хозяина в зависимости от конкретной библиотеки, подлежащей скринингу, и конкретного скринингового количественного исследования, однако в некоторых случаях могут быть использованы подходящие линии эукариотических клеток, включая, но не ограничиваясь ими, например, коммерчески доступные от таких поставщиков как Американская коллекция типовых культур (АТСС) (Манассас, Виргиния, США), Life Technologies, Inc. (Гранд-Айленд, Нью-Йорк, США) и т.п.

В некоторых случаях клетки клеточной библиотеки представляют собой первичные клетки (например, первичные моноциты, первичные лимфоциты (например, первичные Т-клетки, первичные В-клетки, первичные НК-клетки и т.д.), первичные дендритные клетки и т.д.), первичные эндотелиальные клетки, первичные эпителиальные клетки, первичные фибробласты, первичные гемопоэтические стволовые клетки, первичные кератиноциты, первичные меланоциты, первичные мезенхимальные стволовые клетки, первичные преадипоциты, первичные мышечные клетки (например, первичные клетки гладкой мускулатуры, первичные клетки скелетной мускулатуры и т.д.) и т.д. В некоторых случаях клетки клеточной библиотеки представляют собой известные клеточные линии (например, клетки Jurkat и т.д.). В некоторых случаях клетки клеточной библиотеки представляют собой специфичные для пациента клетки (специфичные для пациента иммунные клетки (например, первичные Т-клетки и т.д.), специфичные для пациента стволовые клетки (например, гемопоэтические стволовые клетки, мезенхимальные стволовые клетки, стволовые клетки, полученные из жировой ткани и т.д.), специфичные для пациента раковые клетки (например, опухолевые клетки, клетки рака крови и т.д.).

Подходящие клетки млекопитающих включают первичные клетки и иммортализованные клеточные линии. Подходящие линии клеток млекопитающих

включают линии клеток человека, линии клеток приматов, отличных от человека, линии клеток грызунов (например, мыши, крысы) и т.п. Подходящие линии клеток млекопитающих включают, но не ограничиваются ими, клетки HeLa (например, Американская коллекция типовых культур (ATCC) № CCL-2), клетки CHO (например, ATCC №№CRL9618, CCL61, CRL9096), клетки HEK293 (например, ATCC №CRL-1573), клетки Vero, клетки NIH 3T3 (например, ATCC №CRL-1658), клетки Huh-7, клетки ВНК (например, ATCC №CCL10), клетки PC12 (ATCC №CRL1721), клетки COS, клетки COS-7 (ATCC №CRL1651), клетки RAT1, L-клетки мыши (ATCC №CCLI.3), клетки мезонефроса человека (HEK) (ATCC №CRL1573), линии клеток HLHepG2, Hut-78, Jurkat, HL-60, NK (например, NKL, NK92 и YTS) и т.п.

В некоторых случаях клетка не является клеткой иммортализованной клеточной линии, а, напротив, представляет собой клетку (например, первичную клетку), полученную от индивидуума. Например, в некоторых случаях, клетка представляет собой иммунную клетку, полученную от индивидуума. В качестве примера клетка представляет собой Т-лимфоцит, полученный от индивидуума. В качестве другого примера клетка представляет собой цитотоксическую клетку, полученную от индивидуума. В качестве другого примера клетка представляет собой стволовую клетку или клетку-предшественника, полученную от индивидуума.

После трансформации множества нуклеиновых кислот, кодирующих синтетические модульные полипептиды библиотеки, трансформированные клетки-хозяева могут быть отсортированы на основании экспрессии ими синтетических модульных полипептидов, например, для удаления клеток, которые не экспрессируют синтетические модульные полипептиды, из библиотеки, чтобы выделить только те клетки, которые экспрессируют синтетические модульные полипептиды на уровнях, превышающих определенный порог экспрессии, чтобы выделить только те клетки, которые экспрессируют синтетические модульные полипептиды на уровнях, ниже определенного порога экспрессии, чтобы выделить только те клетки, которые экспрессируют синтетические модульные полипептиды в определенном диапазоне уровней экспрессии, и т.д. В некоторых случаях сортировка трансформированных клеток на основании уровня экспрессии может быть выполнена, чтобы выделить только те клетки в библиотеке, которые имеют одинаковый уровень экспрессии, причем одинаковый уровень экспрессии может варьироваться в зависимости от конкретных способов применения и в некоторых случаях может быть определен как уровень экспрессии в пределах определенного диапазона, который выше и ниже среднего уровня экспрессии популяции.

Согласно некоторым вариантам реализации настоящего изобретения сортировка

трансформированных клеток на основании уровня экспрессии ими элементов библиотеки, например, сортировка клеток, экспрессирующих или имеющих приблизительно равный уровень экспрессии элементов библиотеки, позволяет улучшить оценку влияния элемента библиотеки и/или модулей, входящих в ее состав. Например, при выделении только тех клеток, которые экспрессируют элементы библиотеки в определенном диапазоне уровней, идентификация элементов библиотеки, влияющих на конкретный фенотип, будет основана на фактической функции элементов библиотеки и тех модулей, которые входят в ее состав, а не на эффективности экспрессии конкретного элемента библиотеки. Согласно некоторым вариантам реализации настоящего изобретения, например, при выполнении полуколичественного исследования частоты экспрессии отдельных элементов библиотеки и/или ее модулей, сортировка клеток, экспрессирующих элементы библиотеки в пределах заранее определенного диапазона уровней, обеспечивает более точное количественное исследование и количественное определение элементов библиотеки и/или модулей на основании их влияния на определенный фенотип, а не на основании характерного для них относительного уровня экспрессии.

Помимо этого сортировка позволяет идентифицировать элементы библиотеки и ее модули, функционирование которых создает или влияет на конкретный фенотип при экспрессии в пределах определенного диапазона уровней экспрессии или выше или ниже определенного порога, включая, например, низкий уровень экспрессии или высокий уровень экспрессии.

Нормирование библиотеки

Библиотеки согласно настоящему изобретению могут быть нормированы или могут не быть нормированы в зависимости от ситуации, в которой создается библиотека, и/или предполагаемого окончательного способа применения библиотеки. В настоящей заявке под термином «нормированный», применительно к описанным библиотекам, подразумевают, что относительные количества каждого элемента библиотеки корректируют, чтобы по меньшей мере уравнивать их количества, по сравнению с относительными количествами каждого элемента библиотеки до корректировки. В некоторых случаях нормирование библиотеки приводит к тому, что библиотека имеет меньший диапазон между количеством наиболее представленных элементов библиотеки и количеством наименее представленных элементов библиотеки. В некоторых случаях нормирование приводит к увеличению количества наименее представленного элемента(ов) библиотеки. В некоторых случаях нормирование приводит к уменьшению количества наиболее представленного элемента(ов) библиотеки.

В некоторых случаях нормирование библиотеки может включать количественную

оценку всех или большинства или репрезентативной выборки (или всей или большей части репрезентативной выборки) элементов библиотеки для определения относительного количества каждого элемента библиотеки в библиотеке. После количественной оценки корректировку проводят на основании количественной оценки, чтобы уравнивать относительное присутствие каждого элемента библиотеки в библиотеке. В зависимости от ситуации указанная корректировка может быть выполнена непосредственно в уже полученной библиотеке так, что библиотека подвергается непосредственному нормированию. В другом варианте корректировка может быть выполнена как часть способа получения библиотеки так, что следующая полученная библиотека будет нормирована.

Любая библиотека согласно настоящему изобретению может быть нормирована, включая, но не ограничиваясь ими, например, библиотеки нуклеиновых кислот, библиотеки полипептидов, неклеточные инкапсулированные библиотеки, клеточные библиотеки и т.д., и в зависимости от библиотеки, которую нормируют, могут быть использованы различные способы. Например, в зависимости от типа библиотеки, которую нормируют, могут быть использованы различные способы количественной оценки относительных количеств элементов библиотеки. Различные способы количественного определения элементов библиотеки нуклеиновых кислот, которые могут быть использованы, включают, но не ограничиваются ими, например, количественное секвенирование (например, секвенирование следующего поколения), количественную ПЦР, количественную гибридизацию (например, микрочип) и т.п. Могут быть использованы различные способы количественного определения элементов полипептидной библиотеки, включая, но не ограничиваясь ими, например, количественную масс-спектрометрию, ИФА и т.п. Могут быть использованы различные способы количественного определения элементов клеточной библиотеки, включая, но не ограничиваясь ими, проточную цитометрию, иммуногистохимическое исследование, количественное секвенирование (например, секвенирование следующего поколения), количественную ПЦР, количественную гибридизацию (например, микрочип) и т.п.

В некоторых случаях, когда элементы библиотеки количественно определены, может быть рассчитана корректировка(и), необходимая для нормирования. Любой удобный и подходящий способ расчета нормирования может быть использован в зависимости от типа библиотеки и/или размера библиотеки. В некоторых случаях может быть использовано линейное уравнение, включая, но не ограничиваясь им, линейное уравнение, представленное на Фиг. 28.

В некоторых случаях, когда нормирование рассчитано для каждого элемента

библиотеки, библиотека может быть скорректирована. Могут быть использованы различные способы непосредственной корректировки библиотеки. Например, в некоторых случаях, клеточная библиотека может быть нормирована с использованием FACS для сортировки равного количества клеток, представляющих каждый элемент библиотеки, в пул или в индивидуально доступные компартменты. В некоторых случаях, когда библиотека уже компартиментализована, библиотека может быть нормирована путем корректировки объема каждого компартмента, включая, например, те случаи, когда различные концентрации элементов библиотеки в каждом компартменте нормируют путем добавления определенного объема жидкости в каждый компартмент, достаточного для выравнивания концентраций.

В некоторых случаях может быть выполнено нормирование объединенной библиотеки нуклеиновых кислот. Объединенные библиотеки нуклеиновых кислот могут быть нормированы по разным причинам. В одном варианте реализации объединенная библиотека нуклеиновых кислот может быть нормирована для компенсации избыточной или недостаточной представленности отдельных элементов библиотеки в библиотеке, например, вследствие более или менее эффективного включения отдельных модулей нуклеиновых кислот в элементы библиотеки во время комбинаторной сборки.

В некоторых случаях элементы библиотеки нуклеиновых кислот и/или модули нуклеиновых кислот, составляющие элементы библиотеки, могут быть количественно определены (например, путем количественного секвенирования). После такой количественной оценки рассчитывается корректировка каждого элемента, необходимая для нормирования. В одном варианте реализации рассчитанная корректировка может быть применена к следующей комбинаторной сборке библиотеки, например, количество каждого модуля нуклеиновой кислоты, используемое для сборки библиотеки нуклеиновых кислот, может быть скорректировано на основании относительной представленности этого модуля в количественно оцененной библиотеке. Следовательно, готовая комбинаторная библиотека будет нормирована путем корректировки начального количества модулей нуклеиновых кислот перед следующей сборкой библиотеки. Соответственно, в некоторых случаях нормирование библиотеки, описанной в настоящем документе, может включать сборку библиотеки с последующей количественной оценкой собранной библиотеки и повторной сборкой нормированного варианта библиотеки, который основан на количественной оценке.

Способы скрининга

В настоящем изобретении предложены способы скрининга библиотек синтетических модульных полипептидов, включая, но не ограничиваясь ими, например,

способы скрининга в условиях *in vitro* и способы скрининга в условиях *in vivo*. Под «скринингом *in vivo*» обычно подразумевают, что библиотеку, содержащую множество уникальных синтетических модульных полипептидов, количественно исследуют в биологической среде живого организма. Живые организмы, которые можно количественно исследовать в условиях *in vivo* в соответствии со способами, описанными в настоящем документе, включают одноклеточные и многоклеточные организмы.

Скрининг одноклеточного организма в условиях *in vivo* обычно включает приведение одноклеточного организма в контакт с библиотекой синтетических модульных полипептидов, причем указанная библиотека синтетических модульных полипептидов может представлять собой библиотеку полипептидов или библиотеку клеток, экспрессирующих синтетические модульные полипептиды, и детектирование фенотипа в одноклеточном организме. В других случаях скрининг одноклеточного организма в условиях *in vivo* может включать приведение одноклеточного организма или множества одноклеточных организмов в контакт с библиотекой нуклеиновых кислот, кодирующих синтетические модульные полипептиды, в условиях, достаточных для экспрессии кодируемых синтетических модульных полипептидов одноклеточными организмами.

Скрининг многоклеточного организма в условиях *in vivo* обычно включает приведение многоклеточного организма в контакт с библиотекой синтетических модульных полипептидов, причем указанная библиотека синтетических модульных полипептидов может представлять собой библиотеку полипептидов или библиотеку клеток, экспрессирующих синтетические модульные полипептиды, и детектирование фенотипа в многоклеточном организме. В других случаях скрининг многоклеточного организма в условиях *in vivo* может включать приведение многоклеточного организма или множества многоклеточных организмов в контакт с библиотекой нуклеиновых кислот, кодирующих синтетические модульные полипептиды, в условиях, достаточных для экспрессии кодируемых синтетических модульных полипептидов многоклеточным организм(ами). Любой подходящий многоклеточный организм может быть использован при скрининге в условиях *in vivo* библиотеки, описанной в настоящем документе, в зависимости от конкретной библиотеки, подлежащей скринингу, и конкретного используемого количественного исследования в условиях *in vivo*, причем конкретные многоклеточные организмы включают, но не ограничиваются ими, например, млекопитающих (например, мышей, крыс и т.д.).

Под термином «скрининг в условиях *in vitro*» обычно подразумевают, что библиотеку, содержащую множество уникальных синтетических модульных

полипептидов, количественно исследуют вне нормальной биологической среды, например, полипептидных модулей библиотеки, биологического материала, используемого при скрининге, или фенотипа, который подвергают скринингу. Например, в некоторых случаях, скрининг в условиях *in vitro* может быть выполнен с использованием искусственной или синтетической экспериментальной среды, включая, но не ограничиваясь ими, например, выделенный образец, выделенную клетку, культуру клеток, выделенную или иссеченную ткань, определенный образец, определенную среду, искусственную ткань, искусственный орган, клеточный экстракт, тканевой экстракт, множество образцов и т.д. Скрининг в условиях *in vitro* может быть выполнен в любом удобном и подходящем сосуде, включая, но не ограничиваясь ими, например, реакционный сосуд, реакционную камеру, пробирку, флакон, планшет, колбу, чашку, предметное стекло и т.п.

Скрининг образца в условиях *in vitro*, включая клеточный образец или неклеточный образец, обычно включает приведение образца в контакт с библиотекой синтетических модульных полипептидов, причем указанная библиотека синтетических модульных полипептидов может представлять собой библиотеку полипептидов или библиотеку клеток, экспрессирующих синтетические модульные полипептиды, и детектирование клеточного фенотипа или другой реакции или молекулярного фенотипа. Любой подходящий образец может быть использован при скрининге библиотеки, описанной в настоящем документе, в условиях *in vitro* в зависимости от конкретной библиотеки, подлежащей скринингу, и конкретного используемого количественного исследования в условиях *in vitro*, причем конкретные образцы включают, но не ограничиваются ими, например, биологические образцы, клеточные образцы, образцы полипептидов, образцы нуклеиновых кислот, химические образцы и т.п.

В некоторых случаях бесклеточные библиотеки синтетических модульных полипептидов могут быть подвергнуты скринингу в условиях *in vitro*. Например, компартментализованную библиотеку синтетических модульных полипептидов можно подвергнуть скринингу для определения фенотипа путем приведения компартментализованной библиотеки синтетических модульных полипептидов в контакт с одним или более агентами, с которыми, согласно прогнозам, реагируют отдельные элементы библиотеки. Любые подходящие способы скрининга бесклеточных библиотек полипептидов, как инкапсулированных, так и объединенных, можно применять в способах, описанных в настоящем документе, включая, но не ограничиваясь ими, например, детектирование методом проточной цитометрии или детектирование фенотипа методом флуоресцентно-активированной сортировки клеток в бесклеточных

количественных исследованиях, основанных на инкапсуляции, например, описанных в Griffiths & Tawfik. EMBO J (2003) 22(1):24-35 и Bernath et al. Anal Biochem (2004) 325(1):151-7; содержание которых полностью включено в настоящую заявку посредством ссылки.

Фенотипы и способы идентификации

Способы скрининга, в условиях *in vivo* или в условиях *in vitro*, обычно включают детектирование фенотипа и идентификацию одного или более элементов библиотеки, связанных с фенотипом. В настоящей заявке термин «фенотип» обычно относится к характеристике молекулы, клетки, ткани, органа или организма, которую детектируют в конкретном количественном исследовании, и, следовательно, может включать, но не ограничивается ими, например, молекулярные фенотипы, клеточные фенотипы, фенотипы организма, фенотипы тканей, фенотипы органов, фенотипы организма и т.д. Фенотип, детектируемый в конкретном количественном исследовании, может представлять собой заранее определенный фенотип, например, известный или ожидаемый фенотип (например, включая известный или ожидаемый уровень определенной характеристики, присутствие или отсутствие известной или ожидаемой характеристики уровня и т.д.) или может быть идентифицирован во время количественного исследования, например, впервые детектируемый или ранее неопределенный фенотип (например, включая впервые детектируемый или ранее неопределенный уровень конкретной характеристики, присутствие или отсутствие впервые детектируемой или ранее неопределенной характеристики и т.д.). Любой подходящий количественный способ исследования для детектирования фенотипа, относящегося к библиотеке синтетических модульных полипептидов, описанной в настоящем документе, можно применять при скрининге указанных библиотек.

Скрининг библиотеки синтетических модульных полипептидов или нуклеиновых кислот, кодирующих библиотеку синтетических модульных полипептидов, позволяет идентифицировать полипептиды и/или их модули, которые эффективно обеспечивают желательный фенотип. Соответственно, в область настоящего изобретения в целом включены полипептиды, идентифицированные путем скрининга описанных в настоящем документе библиотек.

В некоторых случаях клеточный фенотип детектируют после приведения популяции клеток в контакт с библиотекой, описанной в настоящем документе. Клеточные фенотипы могут включать, но не ограничиваются ими, например, поведение клеток (включая, но не ограничиваясь ими, жизнеспособность клеток, пролиферацию клеток, активацию клеток, морфологию клеток, миграцию клеток, адгезию клеток,

дифференцировку клеток, плюрипотентность клеток и т.д.), клеточную экспрессию (включая, но не ограничиваясь ими, экспрессию гена, экспрессию белка, экспрессию некодирующей РНК, активацию гена, репрессию гена и т.д.), экспрессию репортера (включая, но не ограничиваясь ими, например, экспрессию трансгенного репортера, экспрессию маркера) и т.п.

В некоторых случаях фенотип ткани, органа или организма детектируют после приведения ткани, органа или организма в контакт с библиотекой, описанной в настоящем документе. Фенотипы ткани включают, но не ограничиваются ими, например, жизнеспособность ткани, морфологию ткани, физические характеристики ткани (включая, но не ограничиваясь ими, защитную функцию, механическую прочность, эластичность и т.д.), экспрессию в ткани (включая, но не ограничиваясь ими, например, экспрессию тканевого гена, экспрессию тканевого белка и т.д.), экспрессию репортера в ткани (включая, но не ограничиваясь ими, например, экспрессию трансгенного репортера, экспрессию маркера) и т.п. Фенотипы органа включают, но не ограничиваются ими, например, внешний вид органа, жизнеспособность органа, морфологию органа, функцию органа (включая, но не ограничиваясь ими, производство биомолекулы (например, фермента, метаболита, белка и т.д.), фильтрацию, механическую функцию и т.д.). Фенотипы организма включают, но не ограничиваются ими, например, внешний вид организма, жизнеспособность организма (например, продолжительность жизни), физиологию организма, фертильность организма/плодовитость, поведение организма и т.д.

В некоторых случаях фенотип может быть количественно исследован в отношении патологического состояния, причем указанное патологическое состояние может представлять собой моделируемое патологическое состояние (например, клетку, ткань или организм, который был изменен или обработан, чтобы проявлять характеристики конкретного заболевания) или может представлять собой клиническое патологическое состояние (например, организм, проявляющий характеристики заболевания или у которого диагностировано заболевание, или клетки или ткани, полученные из него). Фенотипы, связанные с заболеванием, могут быть количественно исследованы на любом удобном уровне, включая, но не ограничиваясь ими, например, клеточный уровень, тканевой уровень, органый уровень, уровень организма. В некоторых случаях количественно исследованные фенотипы заболевания могут представлять собой фенотипы самого возбудителя заболевания, включая, но не ограничиваясь ими, например, фенотипы опухолей, фенотипы раковых клеток, фенотипы аутоиммунных клеток, фенотипы инфекционных агентов (бактериальных, вирусных и т.д.) и т.д. В других

случаях количественно исследованные фенотипы заболевания могут представлять собой фенотипы клетки, ткани или организма, пораженного указанным заболеванием, или связанные с моделируемым заболеванием, которые обеспечивают информацию о присутствии и/или прогрессировании заболевания, включая, но не ограничиваясь ими, например, активацию клеток (например, активацию иммунных клеток), ответ на заболевание (например, иммунный ответ), биомаркеры, количество клеток, физиологические реакции организма, клинические исходы и т.д.

В некоторых случаях оценка фенотипа может быть проведена на уровне популяции, например, может быть проведена оценка популяции клеток, оценка популяции организмов и т.д. В некоторых случаях при популяционной оценке фенотипа действие конкретного элемента библиотеки на присутствие или отсутствие популяционного фенотипа может быть измерено. Например, можно оценить действие конкретного элемента библиотеки на клеточный фенотип популяции клеток. В других случаях можно оценить действие конкретного элемента библиотеки на фенотип организмов в популяции организмов.

Согласно некоторым вариантам реализации настоящего изобретения фенотип оценивают в ответ на применяемый стимул, причем применение стимула включает, но не ограничивается перечисленными, например, приведение клеток в контакт со стимулом, приведение ткани в контакт со стимулом, приведение органа в контакт со стимулом, приведение организма в контакт со стимулом и т.д. По существу исследуемый образец или исследуемый субъект может быть приведен в контакт со стимулом в условиях *in vitro* или *in vivo* в зависимости от используемого количественного способа исследования, в зависимости от стимула и в зависимости от конкретной библиотеки, которую подвергают скринингу. Различные стимулы могут быть использованы по отдельности или в комбинации. Стимул может представлять собой свободный стимул (например, растворимый стимул, свободный лиганд и т.д.), связанный стимул (например, связанный с твердым носителем), экспрессируемый клетками стимул (например, экспрессированная костимулирующая молекула, экспрессированный антиген, экспрессированный клеточный лиганд и т.д.) и т.п.

Согласно некоторым вариантам реализации настоящего изобретения популяцию Т-клеток, экспрессирующую библиотеку синтетических модульных CAR, приводят в контакт со стимулом в условиях *in vitro* и детектируют полученный в результате фенотип. Стимулы в условиях *in vitro*, пригодные для скрининга клеточной библиотеки, экспрессирующей синтетические модульные CAR, обычно будут включать антигены, включая, например, свободный антиген, связанный антиген, экспрессируемый клетками

антиген (например, экспрессируемый на антигенпрезентирующей клетке, экспрессируемый на клетке-мишени и т.д.) и т.д. Подходящие антигены будут варьироваться в зависимости от конкретной библиотеки CAR, подлежащей скринингу, и желательного результата скрининга. Неограничивающие примеры антигенов включают, но не ограничиваются ими, например, растворимый антиген, антиген, связанный с твердой подложкой (например, связанный с планшетами антиген, связанный с гранулами антиген, связанный с предметным стеклом антиген и т.д.), экспрессируемый антиген (например, трансгенная клетка, экспрессирующая антиген, клетка, экспрессирующая антиген в природных условиях (например, клетка, экспрессирующая нативный антиген, раковая клетка, экспрессирующая раковый антиген и т.д.). Клетки, экспрессирующие нативный антиген, которые можно применять для скрининга библиотеки в условиях *in vitro*, будут варьироваться и могут включать, но не ограничиваются ими, например, наивные опухолевые клетки (например, полученные при биопсии опухоли).

Согласно некоторым вариантам реализации настоящего изобретения популяцию Т-клеток, экспрессирующую библиотеку синтетических модульных CAR, приводят в контакт со стимулом в условиях *in vivo* и детектируют полученный в результате фенотип. Ситуации скрининга библиотеки синтетических модульных CAR в условиях *in vivo* будут сильно различаться и могут включать, но не ограничиваются ими, животные модели. В некоторых случаях скрининг в условиях *in vivo* может быть выполнен у небольших модельных животных, таких как, например, моделирование на грызунах, включая, но не ограничиваясь ими, мышинные модели, крысиные модели и т.д. В некоторых случаях скрининг в условиях *in vivo* выполняют на мышинных моделях опухолей, включая трансгенные и нетрансгенные мышинные модели опухолей.

В некоторых случаях используемая модель может представлять собой ксенотрансплантатную модель. Например, используемая модель может представлять собой «гуманизованную» модель, причем указанные гуманизованные модели определяют как имеющие один или более компонентов человеческого происхождения, например, гуманизованная иммунная система, гуманизованные Т-клетки, экспрессирующие белок человека, несущие раковые клетки человека и т.д. В этой связи гуманизованные модели могут быть полностью или частично гуманизованы. В других случаях модель может быть не полностью или частично гуманизованной, но вместо этого может быть просто введена с клетками человека или тканью человека посредством инъекции или трансплантации. Например, в некоторых случаях, раковые клетки человека или клетки из клеточных линий рака человека вводят в животную модель. Любые удобные опухолевые клетки человека или линия опухолевых клеток

человека может быть использована в указанных моделях, включая, но не ограничиваясь ими, например, клетки K562, клетки лимфомы Дауди и т.д.

Животные модели и/или клетки или ткани, введенные в животные модели, могут быть трансгенными или нетрансгенными, например, могут быть модифицированы для экспрессии одного или более трансгенов. Например, в некоторых случаях, животная модель может быть трансгенно модифицирована для экспрессии гетерологичного гена, например, репортерного гена (например, для идентификации клеток животного-хозяина), гена-мишени (например, гена, кодирующего генный продукт, который должен быть мишенью при скрининге в условиях *in vivo*). В некоторых случаях клетка, введенная в животную модель, может быть трансгенно модифицирована для экспрессии гетерологичного гена, например, репортерного гена (например, для идентификации введенной клетки), гена-мишени (например, гена, кодирующего генный продукт, который должен быть мишенью при скрининге в условиях *in vivo*). В качестве неограничивающего примера мышьяная модель опухоли может быть подвергнута скринингу в соответствии со способами, описанными в настоящем документе, при этом мыши вводят опухолевые клетки человека, экспрессирующие раковый трансген-мишень (например, CD19, мезотелин и т.д.).

Элементы библиотеки, внедренные в системы *in vivo*, могут быть подвергнуты скринингу для определения любого подходящего фенотипа, причем указанный фенотип может зависеть от конкретной библиотеки, которую подвергают скринингу, конкретной ситуации в условиях *in vivo* (например, животной модели) и т.д. В некоторых случаях, например, если система в условиях *in vivo* представляет собой животную модель опухоли, библиотека может быть подвергнута скринингу для определения фенотипов, связанных с селективностью элементов библиотеки в отношении опухолей, например, путем введения библиотеки в животную модель, содержащую две разные опухоли, путем введения библиотеки в животную модель или несколько животных моделей, содержащих опухоли, экспрессирующие различные уровни опухолевого антигена и т.д. Любой удобный способ количественного исследования фенотипа, включая описанные в настоящем документе клеточные и биохимические/молекулярные способы, можно применять при оценке систем *in vivo*, причем такие оценки обычно включают получение биологического образца из животной модели. В некоторых случаях биологический образец, пригодный для оценки модели в условиях *in vivo*, может включать образец ткани (например, крови, опухоли и т.д.) или образец органа (например, селезенки).

В некоторых случаях библиотека может быть подвергнута скринингу в условиях *in vitro* или в условиях *in vivo* в соответствии с фенотипом Т-клеток. Фенотипы Т-клеток

будут варьироваться и будут включать стимулированные фенотипы Т-клеток, т.е. ответ на антиген. Неограничивающие примеры фенотипов Т-клеток включают, но не ограничиваются ими, например, пролиферацию Т-клеток, выработку цитокинов (например, ИЛ-2, ИФН- γ , ФНО, ЛТ- α , ИЛ-4, ИЛ-5, ИЛ-6, ИЛ-13, ИЛ-9, ИЛ-10, ИЛ-17А, ИЛ-17F, ИЛ-21, ИЛ-22, ИЛ-26, ССL20, ИЛ-21, ФРО- β и т.д.), экспрессию поверхностных маркеров Т-клеток (например, CD3, CD4, CD8 и т.д.), маркеры активации Т-клеток (например, CD69 и т.д.), маркеры внутриклеточной сигнализации (например, фосфорилированную форму ERK1/2, фосфорилированную форму p38MAPK и т.д.) и т.п.

Фенотипы Т-клеток могут быть количественно исследованы в условиях *in vitro* и в условиях *in vivo* и могут быть детектированы любым удобным способом. В некоторых случаях устройство для подсчета клеток или проточный цитометр можно использовать для количественного исследования фенотипа Т-клеток, включая, например, пролиферацию Т-клеток и/или количественную оценку Т-клеток. Например, пролиферация Т-клеток может быть количественно исследована с помощью метода проточной цитометрии путем разбавления красителя для мечения клеток. В некоторых случаях экспрессия маркеров клеточной поверхности также может быть количественно исследована с помощью проточной цитометрии. Экспрессия внутриклеточных маркеров может быть определена клеточными способами (например, с помощью проточной цитометрии, проточной цитометрии с использованием фосфо-специфичных антител, проточной цитометрии внутриклеточных маркеров, иммунофлуориметрии, гибридизации *in situ*, флуоресцентной гибридизации *in situ* и т.д.) или может быть количественно исследована с помощью молекулярных и/или биохимических способов (например, ИФА, захвата цитокинов, способов на основе амплификации (например, количественной ПЦР), способов, основанных на секвенировании (например, с помощью количественного секвенирования), количественной масс-спектрометрии и т.д.).

В некоторых случаях Т-клетки могут быть количественно исследованы для определения фенотипа активации «природных киллеров». Любой подходящий способ оценки активации природных киллеров можно применять в указанных количественных исследованиях. Например, Т-клетки могут быть количественно исследованы для определения экспрессии CD107a/b, например, методом проточной цитометрии.

В некоторых случаях Т-клетки могут быть количественно исследованы для определения одного или более фенотипов дифференцировки. Любой подходящий способ оценки дифференцировки Т-клеток можно применять в указанных количественных исследованиях. Например, дифференцировку в Т-клетку памяти можно определить, например, путем количественного исследования маркеров Т-клеток памяти (например,

Th1, Th2, Th17, Treg и т.д.) с использованием любого подходящего клеточного или молекулярно-биохимического способа. В некоторых случаях может быть проведена оценка экспрессии одного или более факторов внутриклеточной транскрипции, указывающих на дифференцировку в Т-клетки памяти (например, Gata3, Tbet, RORyt, FoxP3, Vcl-6, CCR7, CD45RO, CD45RA, CD69 и т.д.).

Скрининг библиотеки синтетических модульных полипептидов CAR или нуклеиновых кислот, кодирующих библиотеку синтетических модульных полипептидов CAR, позволяет идентифицировать CAR и/или их части (например, антигенсвязывающие домены, первичные сигнальные домены, комодулирующие домены и т.д.), которые эффективно приводят к возникновению желательного фенотипа Т-клеток. Соответственно, в область настоящего изобретения включены CAR, идентифицированные путем скрининга библиотек, описанных в настоящем документе, а также нуклеиновых кислот, кодирующих указанные CAR. В область настоящего изобретения также включены CAR, содержащие подходящие модули CAR (например, антигенсвязывающие домены, первичные сигнальные домены, комодулирующие домены и т.д.), идентифицированные путем скрининга библиотек, описанных в настоящем документе, а также нуклеиновых кислот, кодирующих указанные CAR.

В некоторых случаях CAR согласно настоящему изобретению может содержать один или более комодулирующих доменов, идентифицированных как стимулирующие Т-клетки или ингибирующие Т-клетки домены при скрининге библиотеки синтетических модульных полипептидов CAR или нуклеиновых кислот, кодирующих синтетические модульные полипептиды CAR, описанных в настоящем документе. Соответственно, общий Т-клеточный фенотип CAR может заключаться в имитации активности Т-клеток или ингибировании активности Т-клеток. Виды активности Т-клеток, которые могут быть подвергнуты стимулированию или ингибированию, включают, но не ограничиваются ими, например, виды активности Т-клеток, описанные в настоящем документе.

В некоторых случаях CAR, идентифицированный путем скрининга библиотеки, может содержать по меньшей мере один комодулирующий домен, перечисленный в таблице 3 или таблице 4, включая, но не ограничиваясь ими, например, комодулирующие домены, содержащие аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% идентична перечисленной последовательности домена. В некоторых случаях CAR, идентифицированный путем скрининга библиотеки, может содержать два или более

комодулирующих доменов, перечисленных в таблице 3 и таблице 4, включая, но не ограничиваясь ими, комодулирующие домены, содержащие аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% идентична перечисленной последовательности домена. В некоторых случаях CAR согласно настоящему изобретению может содержать комодулирующий домен, идентифицированный при скрининге, описанном в настоящем документе, в качестве костимулирующего домена, включая, но не ограничиваясь ими, например, те, которые перечислены в таблице 3, включая, но не ограничиваясь ими, комодулирующие домены, содержащие аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% идентична перечисленной последовательности домена. В некоторых случаях CAR согласно настоящему изобретению может содержать комодулирующий домен, идентифицированный при скрининге, описанном в настоящем документе, в качестве коингибирующего домена, включая, но не ограничиваясь ими, например, те, которые перечислены в таблице 4, включая, но не ограничиваясь ими, комодулирующие домены, содержащие аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% идентична перечисленной последовательности домена. В некоторых случаях CAR, который содержит один или более коингибирующих доменов, может представлять собой iCAR.

В некоторых случаях CAR, содержащий два или более комодулирующих доменов, может содержать два костимулирующих домена, включая, но не ограничиваясь ими, например, два или более костимулирующих доменов, перечисленных в таблице 3. В некоторых случаях CAR, содержащий два или более комодулирующих доменов, может содержать два коингибирующих домена, включая, но не ограничиваясь ими, например, два или более коингибирующих доменов, перечисленных в таблице 4. В некоторых случаях CAR, содержащий два или более комодулирующих доменов, может включать смесь костимулирующих и коингибирующих доменов, включая, но не ограничиваясь ими, по меньшей мере один костимулирующий домен, перечисленный в таблице 3, и по меньшей мере один коингибирующий домен, перечисленный в таблице 4.

Таблица 3: Комодулирующие домены, обладающие стимулирующей функцией (костимулирующие домены)

Домен	Последовательность	SEQ ID NO:
4-1BB	KRGRKLLYIFKQPFMRPVQTTQEEDGCSCRFPEEEEGGCEL	26
CD7	ARTQIKKLCRWDKNSAACVVYEDMSHSRCNTLSSPNQYQ	25
2B4	WRRKRKEKQSETSPKEFLTIYEDVKDLKTRRNHEQEQTFFPGGGSTIYSMIQS QSSAPTSQEPAYTLYSLIQPSRKSGSRKRNHSPSFNSTIYEVIGKSQPKAQN ARLSRKELENFDVYS	41
HVEM	MEESVVRPSVFFVVDGQTDIPFTRLGRSHRRQSCSV	23
CRTAM	KLRKAHVIVKKENEVSEHTLESYRSRNNNEETSSEEKNGQSSHPMRCMNYI TKLYSEAKTKRKENVQHSKLEEKHIQVPESIV	35
CD30	RRACRKRIRQKLHLCPVQTSQPKLELVDSRPRRSSTQLRSGASVTEPVAEE RGLMSQPLMETCHSVGAAYLESPLQDASPAGGPSSPRDLPEPRVSTEHTN NKIEKIYIMKADTVIVGTVKAELPEGRGLAGPAEPELEEELEADHTPHYPEQ ETEPPLGSCSDVMLSVEEEGKEDPLPTAASGK	42
TLT2	KKRHMASYSMCDPSTRDPPGRPEPYVEVYLI	21
CD27	HQRRKYRSNKGESPVPEAEPCHYSCPREEEGSTIPIQEDYRKPEPACSP	28
CTLA4	SLSKMLKRSPLTTGVYVKMPTEPECEKQFQPYFIPIN	24

Таблица 4: Комодулирующие домены, обладающие ингибиторной функцией (коингибирующие домены)

Домен	Последовательность	SEQ ID NO:
DNAM-1	NRRRRRERRDLFTESWDTQKAPNNYRSPISTSQPTNQSMDDTREDIYVNYPTFSRR PKTRV	31
CD80	RCRERRRNERLRRESVRPV	19
PD-1	ICSRAARGTIGARRTGQPLKEDPSAVPVFSVDYGELDFQWREKTPEPPVPCVPEQT EYATIVFSPGMGTSSPARRGSADGPRSAQPLRPEDGHCSWPL	36
TIM-3	KWYSHSKEKIQNLSLISLANLPPSGLANAVAEGIRSEENIYTIENVYEVVEEPNEY CYVSSRQQPSQPLGCRFAMP	33
BTLA	CCLRRHQGKQNELSDTAGREINLVDAHLKSEQTEASTRQNSQVLLSETGIYDNDPD LCFRMQEGSEVYSNPCLEENKPGIVYASLNHSGVGNPNSRLARNVKEAPTEYASICV RS	39
CD40	KKVAKKPTNKAPHPKQEPQEIFPDDLPGSNTAAPVQETLHGCQPVTQEDGKESRI SVQERQ	32
ICOS	TKKKYSSSVHDPNGEYMFMRVAVNTAKKSRLTDVTL	22
LAG3	RRQWRPRRFSALEQGIHPPQAQSKIEELEQEPEPEPEPEPEPEPEPEPEQL	29
GITR	HIWQLRSQCMWPRETQLLLEVPSTEDARSCQFPEEERGERSAEEKGRLGDLWV	30
TIGIT	RKKKALRIHSVEGDLRKRKASAGQEEWSPSAPSPPGSCVQAEAAPAGLCGEQRGEDC AELHDYFNVLSYRSLGNCFFTETG	34
PD-L1	RKGRMMDVKKCGIQDTNSKKQSDTHLEET	20
CD28	WVRSKRSLRHSDYMNMTPRRPGPTRKHYPYAPPRDFAAYRS	27
LAIR1	HRQNQIKQGPPRSKDEEQKPPQRPDLAVDVLERTADKATVNGLPKDRDTSAL AAGSSQEVTYAQLDHWALTQRTARAVSQSTKPMASITYAAVARH	37
CAR	CCRKKRREEKYEKEVHHDIREDVPPPKSRTSTARSYIGSNHSSLGSMSPSNMEGYS KTQYNQVPSEDFERTPQSPTLPPAKVAAPNLSRMGAIPVMIPAQSKDGSIV	38
CD2	TKRKKQRSRRNDEELETRAHRVATEERGRKPHQIPASTPQNPATSQHPPPPGHRS QAPSHRPPPPGHRVQHQPQRPPAPSGTQVHQKGPPLPRPRVQPKPPHGAENSL SPSSN	40

CAR, идентифицированный путем скрининга библиотеки синтетических модульных полипептидов CAR или библиотеки нуклеиновых кислот, кодирующих синтетические полипептиды CAR, может содержать любой подходящий антигенсвязывающий домен, включая, но не ограничиваясь ими, те, которые используются в клинических условиях в различных конструкциях CAR, включая, например, ВСМА-специфичный антигенсвязывающий домен, CD123-специфичный

антигенсвязывающий домен, CD138-специфичный антигенсвязывающий домен, CD171-специфичный антигенсвязывающий домен, CD19-специфичный антигенсвязывающий домен, CD22-специфичный антигенсвязывающий домен, CD30-специфичный антигенсвязывающий домен, CD33-специфичный антигенсвязывающий домен, CD7-специфичный антигенсвязывающий домен, CD70-специфичный антигенсвязывающий домен, SEA-специфичный антигенсвязывающий домен, EGFRvIII-специфичный антигенсвязывающий домен, EPCAM-специфичный антигенсвязывающий домен, EphA2-специфичный антигенсвязывающий домен, ErbB-специфичный антигенсвязывающий домен, FAP-специфичный антигенсвязывающий домен, GD2-специфичный антигенсвязывающий домен, GPC3-специфичный антигенсвязывающий домен, HER2-специфичный антигенсвязывающий домен, IL1RAP-специфичный антигенсвязывающий домен, каппа-специфичный антигенсвязывающий домен, LeY-специфичный антигенсвязывающий домен, Meso-специфичный антигенсвязывающий домен, MG7-специфичный антигенсвязывающий домен, MUC1-специфичный антигенсвязывающий домен, NKG2D-специфичный антигенсвязывающий домен, PSCA-специфичный антигенсвязывающий домен, ROR1-специфичный антигенсвязывающий домен и т.п.

CAR, идентифицированный путем скрининга библиотеки синтетических модульных полипептидов CAR или библиотеки нуклеиновых кислот, кодирующих синтетические модульные полипептиды CAR, может содержать любой подходящий первичный сигнальный домен (также называемый в настоящем документе внутриклеточным сигнальным доменом), включая, но не ограничиваясь ими, например, те, которые содержат один или более иммунорецепторных тирозиновых активирующих мотивов (ITAM).

Подходящий внутриклеточный сигнальный домен может быть частью, содержащей мотив ITAM, которая получена из полипептида, который содержит мотив ITAM. Например, подходящий внутриклеточный сигнальный домен может представлять собой домен, содержащий мотив ITAM, из любого содержащего мотив ITAM белка. Следовательно, подходящий внутриклеточный сигнальный домен необязательно содержит полную последовательность всего белка, из которого он получен. Примеры подходящих полипептидов, содержащих мотив ITAM, включают, но не ограничиваются ими: DAP12; FCER1G (гамма-цепь рецептора Fcε I); CD3D (CD3-дельта); CD3E (CD3-эпсилон); CD3G (CD3-гамма); CD3Z (CD3-дзета); и CD79A (альфа-цепь белка, связанного с антигенраспознающим рецепторным комплексом).

В некоторых случаях внутриклеточный сигнальный домен получен из DAP12 (также известного как TYROBP; белок, связывающий тирозинкиназу TYRO, KARAP;

PLOSL; DNAX-активирующий белок 12; связанный с KAR белок; связывающий тирозинкиназу TYRO белок; активирующий киллеры ассоциированный с рецептором белок; активирующий киллеры рецептор-ассоциированный белок и т.д.). Например, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична любой из следующих аминокислотных последовательностей (4 изоформы):

MGGLEPCSRLLLLPLLLAVSGLRPVQAQAQSDCSCSTVSPGVLAGIVMGDLVLTVLIALAVYFLGRLVPRGRGAAEAATRKQRITETESPYOELQGQRSDVYSDLNTQRPYYK (SEQ ID NO: 107);

MGGLEPCSRLLLLPLLLAVSGLRPVQAQAQSDCSCSTVSPGVLAGIVMGDLVLTVLIALAVYFLGRLVPRGRGAAEAATRKQRITETESPYOELQGQRSDVYSDLNTQRPYYK (SEQ ID NO: 108);

MGGLEPCSRLLLLPLLLAVSDCSCSTVSPGVLAGIVMGDLVLTVLIALAVYFLGLVPRGRGAAEAATRKQRITETESPYOELQGQRSDVYSDLNTQRPYYK (SEQ ID NO: 109); или

MGGLEPCSRLLLLPLLLAVSDCSCSTVSPGVLAGIVMGDLVLTVLIALAVYFLGLVPRGRGAAEAATRKQRITETESPYOELQGQRSDVYSDLNTQRPYYK (SEQ ID NO: 110),

в которых мотивы ITAM выделены жирным шрифтом и подчеркнуты.

Аналогичным образом, подходящий полипептид внутриклеточного сигнального домена может содержать часть, содержащую мотив ITAM, из полноразмерной аминокислотной последовательности DAP12. Следовательно, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична следующей аминокислотной последовательности:

ESPYOELQGQRSDVYSDLNTQ (SEQ ID NO:111),

в которой мотивы ITAM выделены жирным шрифтом и подчеркнуты.

В некоторых случаях внутриклеточный сигнальный домен получен из FCER1G (также известного как FCRG; гамма-цепь рецептора Fcε I; гамма-цепь рецептора Fc; Fc-эпсилон RI-гамма; FcR-гамма; FcεRI гамма; высокоаффинная гамма-субъединица

эпсилон-рецептора иммуноглобулина; рецептор иммуноглобулина Е; высокоаффинная гамма-цепь и т.д.). Например, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична следующей аминокислотной последовательности:

**MIPAVVLLLLLLVEQAAALGEPQLCYILDAILFLYGIVLTLTYCRLKIQVRKAAITSYEKS
DGVYTGLSTRNQETYETLKHEKPPQ (SEQ ID NO: 112),**

в которой мотивы ITAM выделены жирным шрифтом и подчеркнуты.

Аналогичным образом, подходящий полипептид внутриклеточного сигнального домена может содержать часть, содержащую мотив ITAM, из полноразмерной аминокислотной последовательности FCER1G. Следовательно, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична следующей аминокислотной последовательности:

DGVYTGLSTRNQETYETLKHE (SEQ ID NO: 113),

в которой мотивы ITAM выделены жирным шрифтом и подчеркнуты.

В некоторых случаях внутриклеточный сигнальный домен получен из дельта-цепи поверхностного гликопротеина CD3 Т-клеток (также известного как CD3D; CD3-DELTA; T3D; антиген CD3; дельта-субъединица; CD3-дельта; CD3d-антиген; дельта-полипептид (комплекс TiT3); OKT3; дельта-цепь; дельта-цепь Т-клеточного рецептора T3; дельта-цепь поверхностного гликопротеина CD3 Т-клеток и т.д.). Например, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична непрерывному участку аминокислотной последовательности от приблизительно 100 аминокислот до приблизительно 110 аминокислот (а.к.), от приблизительно 110 а.к. до приблизительно 115 а.к., от приблизительно 115 а.к. до приблизительно 120 а.к., от приблизительно 120 а.к. до

приблизительно 130 а.к., от приблизительно 130 а.к. до приблизительно 140 а.к., от приблизительно 140 а.к. до приблизительно 150 а.к. или от приблизительно 150 а.к. до приблизительно 170 а.к. любой из следующих аминокислотных последовательностей (2 изоформы):

MEHSTFLSGLVLATLLSQVSPFKIPIEELEDRVFVNCNTSITWVEGTVGTLLSDITRLDLG
KRILDPRGIYRCNGTDIYKDKESTVQVHYRMCQSCVELDPATVAGIIVTDVIATLLLALG
VFCFAGHETGRLSGAADTQALLRNDQVYOPLRDRDDAQYSHLGGNWARNK (SEQ ID
NO: 114) или

MEHSTFLSGLVLATLLSQVSPFKIPIEELEDRVFVNCNTSITWVEGTVGTLLSDITRLDLG
KRILDPRGIYRCNGTDIYKDKESTVQVHYRTADTQALLRNDQVYOPLRDRDDAQYSHL
GGNWARNK (SEQ ID NO: 115),

в которой мотивы ITAM выделены жирным шрифтом и подчеркнуты.

Аналогичным образом, подходящий полипептид внутриклеточного сигнального домена может содержать часть, содержащую мотив ITAM, из полноразмерной аминокислотной последовательности CD3-дельта. Следовательно, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична следующей аминокислотной последовательности:

DQVYOPLRDRDDAQYSHLGGN (SEQ ID NO: 116),

в которой мотивы ITAM выделены жирным шрифтом и подчеркнуты.

В некоторых случаях внутриклеточный сигнальный домен получен из эpsilon-цепи поверхностного гликопротеина CD3 T-клеток (также известной как CD3ε; эpsilon-цепь поверхностного антигена T3/Leu-4 T-клеток; эpsilon-цепь поверхностного гликопротеина CD3 T-клеток; AI504783; CD3; CD3-эpsilon; T3ε и т.д.). Например, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична непрерывному участку аминокислотной последовательности от приблизительно 100 аминокислот до приблизительно 110 аминокислот (а.к.), от приблизительно 110 а.к. до приблизительно

115 а.к., от приблизительно 115 а.к. до приблизительно 120 а.к., от приблизительно 120 а.к. до приблизительно 130 а.к., от приблизительно 130 а.к. до приблизительно 140 а.к., от приблизительно 140 а.к. до приблизительно 150 а.к. или от приблизительно 150 а.к. до приблизительно 205 а.к. следующей аминокислотной последовательности:

MQSGTHWRVGLCLLSVGVWGQDGNEEMGGITQTPYKVSISGTTVILTCPQYPGSEILW
QHNDKNIGGDEDDKNIGSDEDHLSLKEFSELEQSGYYVCYPRGSKPEDANFYLYLRARV
CENCMEMDMVMSVATIVIVDICITGGLLLL VYYWSKNRKA KAKPVTRGAGAGGRQRGQ
NKERPPVPNPDYEPIRGQRDLYSGLNQRR (SEQ ID NO: 117),

в которой мотивы ITAM выделены жирным шрифтом и подчеркнуты.

Аналогичным образом, подходящий полипептид внутриклеточного сигнального домена может содержать часть, содержащую мотив ITAM, из полноразмерной аминокислотной последовательности CD3 эпсилон. Следовательно, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична следующей аминокислотной последовательности:

NPDYEPIRGQRDLYSGLNQR (SEQ ID NO: 118),

в которой мотивы ITAM выделены жирным шрифтом и подчеркнуты.

В некоторых случаях внутриклеточный сигнальный домен получен из гамма-цепи поверхностного гликопротеина CD3 Т-клеток (также известной как CD3G; гамма-цепь Т-клеточного рецептора T3; CD3-гамма; T3G; гамма-полипептид (комплекс T1T3) и т.д.). Например, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична непрерывному участку аминокислотной последовательности от приблизительно 100 аминокислот до приблизительно 110 аминокислот (а.к.), от приблизительно 110 а.к. до приблизительно 115 а.к., от приблизительно 115 а.к. до приблизительно 120 а.к., от приблизительно 120 а.к. до приблизительно 130 а.к., от приблизительно 130 а.к. до приблизительно 140 а.к., от приблизительно 140 а.к. до приблизительно 150 а.к. или от приблизительно 150 а.к. до приблизительно 180 а.к. следующей аминокислотной последовательности:

MEQGKGLAVLILAILLLQGTLAQSIKGNHLVKVYDYQEDGSVLLTCDAEAKNITWFKDG
 KMIGFLTEDKKKWNLGSNAKDPRGMVYQCKGSQNKSKPLQVYYRMCQNCIELNAATIS
 GFLFAEIVSIFVLAVGVYFIAGQDGVQRASDKQTLLPNDQLYQPLKDREDDQYSHLQ
 GNQLRRN (SEQ ID NO: 119),

в которой ITAM мотивы выделены жирным шрифтом и подчеркнуты.

Аналогичным образом, подходящий полипептид внутриклеточного сигнального домена может содержать часть, содержащую мотив ITAM, из полноразмерной аминокислотной последовательности CD3-гамма. Следовательно, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична следующей аминокислотной последовательности:

DQLYQPLKDREDDQYSHLQGN (SEQ ID NO: 120),

в которой мотивы ITAM выделены жирным шрифтом и подчеркнуты.

В некоторых случаях внутриклеточный сигнальный домен получен из дзета-цепи поверхностного гликопротеина CD3 Т-клеток (также известной как CD3Z; дзета-цепь Т-клеточного рецептора Т3; CD247; CD3-дзета; CD3H; CD3Q; Т3Z; TCRZ и т.д.). Например, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична непрерывному участку аминокислотной последовательности от приблизительно 100 аминокислот до приблизительно 110 аминокислот (а.к.), от приблизительно 110 а.к. до приблизительно 115 а.к., от приблизительно 115 а.к. до приблизительно 120 а.к., от приблизительно 120 а.к. до приблизительно 130 а.к., от приблизительно 130 а.к. до приблизительно 140 а.к., от приблизительно 140 а.к. до приблизительно 150 а.к. или от приблизительно 150 а.к. до приблизительно 160 а.к. любой из следующих аминокислотных последовательностей (2 изоформы):

MKWKALFTAAILQAQLPITEAQSFGLLDPKLCYLLDGILFIYGVILTALFLRVKFSRSADA
 PAYQQGQNQLYNELNLGRREEYDVLDKRRGRDPEMGGKPRRKNPQEGLYNELQKDK
 MAEAYSEIGMKGERRRGKGHDGLYOGLSTATKDTYDALHMQALPPR (SEQ ID NO:
 121) или

MKWKALFTAAILQAQLPITEAQSFGLLDPKLCYLLDGILFIYGVILTALFLRVKFSRSADA
 PAYQQGQNQLYNELNLGRREEYDVLDKRRGRDPEMGGKPRRKNPQEGLYNELQKD
 KMAEAYSEIGMKGERRRGKGHDGLYOGLSTATKDTYDALHMQALPPR (SEQ ID NO:
 в которой мотивы ITAM выделены жирным шрифтом и подчеркнуты.

Аналогичным образом, подходящий полипептид внутриклеточного сигнального домена может содержать часть, содержащую мотив ITAM, из полноразмерной аминокислотной последовательности CD3-дзета. Следовательно, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична любой из следующих аминокислотных последовательностей:

RVKFSRSADAPAYQQGQNQLYNELNLGRREEYDVLDKRRGRDPEMGGKPRRKNPQEG
YNELQKDKMAEAYSEIGMKGERRRGKGHDGLYOGLSTATKDTYDALHMQALPPR
 (SEQ ID NO: 123);

NQLYNELNLGRREEYDVLDKR (SEQ ID NO: 124); EGLYNELQKDKMAEAYSEIGMK
 (SEQ ID NO: 125); или DGLYOGLSTATKDTYDALHMQ (SEQ ID NO: 126),

в которой мотивы ITAM выделены жирным шрифтом и подчеркнуты.

В некоторых случаях внутриклеточный сигнальный домен получен из CD79A (также известной как альфа-цепь белка, связанного с антигенным рецептором В-клеток; антиген CD79a (ассоциированная с иммуноглобулином альфа-цепь); мембранный гликопротеин MB-1; ig-альфа; связанный с мембраной ассоциированный с иммуноглобулином белок; поверхностный IgM-ассоциированный белок и т.д.). Например, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична непрерывному участку аминокислотной последовательности от приблизительно 100 аминокислот до приблизительно 110 аминокислот (а.к.), от приблизительно 110 а.к. до приблизительно

115 а.к., от приблизительно 115 а.к. до приблизительно 120 а.к., от приблизительно 120 а.к. до приблизительно 130 а.к., от приблизительно 130 а.к. до приблизительно 150 а.к., от приблизительно 150 а.к. до приблизительно 200 а.к. или от приблизительно 200 а.к. до приблизительно 220 а.к. любой из следующих аминокислотных последовательностей (2 изоформы):

MPGGPGVLQALPATIFLLFLLSAVYLGPGCQALWMHKVPASLMVSLGEDAHFQCPHNS
SNNANVTWWRVHLHGNYTWPPEFLGPGEDPNGTLIIQNVNKS HGGIYVCRVQEGNESYQ
QSCGTYLRVRQPPRPFLDMGEGTKNRIITAEGIILLFCAVVPGTLLLFRKRWQNEKLGL
DAGDEYEDENLYEGLNLDDCSMYEDISRGLQGTYQDVGSLNIGDVQLEKP (SEQ ID
NO: 127); или

MPGGPGVLQALPATIFLLFLLSAVYLGPGCQALWMHKVPASLMVSLGEDAHFQCPHNS
SNNANVTWWRVHLHGNYTWPPEFLGPGEDPNEPPRPFLDMGEGTKNRIITAEGIILLFCA
VVPGTLLLFRKRWQNEKLGLDAGDEYEDENLYEGLNLDDCSMYEDISRGLQGTYQDV
GSLNIGDVQLEKP (SEQ ID NO: 128),

в которой мотивы ITAM выделены жирным шрифтом и подчеркнуты.

Аналогичным образом, подходящий полипептид внутриклеточного сигнального домена может содержать часть, содержащую мотив ITAM, из полноразмерной аминокислотной последовательности CD79A. Следовательно, подходящий полипептид внутриклеточного сигнального домена может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична следующей аминокислотной последовательности: ENLYEGLNLDDCSMYEDISRGL (SEQ ID NO: 129),

в которой мотивы ITAM выделены жирным шрифтом и подчеркнуты.

Внутриклеточные сигнальные домены, подходящие для применения в CAR согласно настоящему изобретению, содержат сигнальную цепь типа DAP10/CD28.

Примером сигнальной цепи DAP10 является аминокислотная последовательность:

RPRRSPAQDGKVIYINMPGRG (SEQ ID NO: 130).

Согласно некоторым вариантам реализации настоящего изобретения подходящий внутриклеточный сигнальный домен содержит аминокислотную последовательность, которая по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или по меньшей мере приблизительно на 99% идентична по всей длине аминокислотной

последовательности RPRRSPAQDGKVVYNMPGRG (SEQ ID NO: 130).

Примером сигнальной цепи CD28 является аминокислотная последовательность: FWLVVVVGGVLACYSLLVTVAFIIFWVRSKRSRLLHSDYMNMTTPRRPGPTRKHYPYA PPRDFAAYRS (SEQ ID NO: 131).

Согласно некоторым вариантам реализации настоящего изобретения подходящий внутриклеточный сигнальный домен содержит аминокислотную последовательность, которая по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или по меньшей мере приблизительно на 99% идентична по всей длине аминокислотной последовательности

FWLVVVVGGVLACYSLLVTVAFIIFWVRSKRSRLLHSDYMNMTTPRRPGPTRKHYPYA PPRDFAAYRS (SEQ ID NO: 131).

Внутриклеточные сигнальные домены, пригодные для применения в CAR согласно настоящему изобретению, включают полипептид ZAP70, например, полипептид, содержащий аминокислотную последовательность, которая по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98%, по меньшей мере приблизительно на 99% или 100% идентична непрерывному участку аминокислотной последовательности от приблизительно 300 аминокислот до приблизительно 400 аминокислот, от приблизительно 400 аминокислот до приблизительно 500 аминокислот или от приблизительно 500 аминокислот до 619 аминокислот следующей аминокислотной последовательности:

MPDPA AHL PFFYGSISR AEAEHLKLAGMADGLFLLRQCLRSLGGYVLSLVHDV
RFHNFPIERQLNGTYAIAGGKAHCGPAELCEFYSRDPDGLPCNLRKPCNRP SGLEPQPGV
FDCLRDAMVRDYVRQ TWKLEGEALEQAIISQAPQVEKLIATTAHERMPWYHSSLTREE
AERKLYSGAQT DGKFLLRPRKEQGTYALSLIYGKTVYHYLISQDKAGKYCIPEGTKFDT
LWQLVEYLK LKADGLIYCLKEACPNSSASNASGAAAPTLP AHPSTLTHPQRRIDTLNSD
GYTPEPARITSPDKPRPMPMDTSVYESPYSDPEELKDKKFLKRDNLLIADIELGCGNFGS
VRQGVYRMRKKQIDVAIKVLKQGTEKADTEEMMREAQIMHQLDNPYIVRLIGVCQAE
ALMLVMEMAGGGPLHKFLVGKREEIPVSNVAELLHQVSMGMKYLEEKNFVHRDLAAR
NLLVNRHYAKISDFGLSKALGADDSYYTARSAGKWPLK WYAPECINFRKFSSRSVDVW
SYGVTMWEALSYGQKPYKKMKGPEVMAFIEQGKRMECPPECPELYALMSDCWIYKW
EDRPDFLTVEQRM RACYYS LASKVEGPPGSTQKAE AACA (SEQ ID NO: 132).

В некоторых случаях CAR, идентифицированный путем скрининга библиотеки синтетических модульных полипептидов CAR или библиотеки нуклеиновых кислот,

кодирующих синтетические модульные полипептиды CAR, включая CAR, содержащий по меньшей мере один или два или более комодулирующих доменов, перечисленных в таблице 3 и таблице 4, может быть разделен на две полипептидные цепи, способные соединяться, в присутствии димеризатора, посредством домена димеризации, присутствующего в каждой цепи. Указанные расщепленные CAR являются условно активными и фармакологически индуцируемыми/подавляемыми, например, как те, которые описаны в РСТ публикации патентной заявки WO2014/127261, содержание которой полностью включено в настоящую заявку посредством ссылки.

Соответственно, в некоторых случаях, каждый полипептид расщепленного варианта CAR из CAR, идентифицированного путем скрининга библиотеки, описанной в настоящем документе, может содержать одну половину пары димеризации (также называемой парой, связывающей димеризатор). Неограничивающие примеры подходящих димеров (например, пар, связывающих димеризатор) включают, но не ограничиваются ими: а) FK506-связывающий белок (FKBP) и FKBP; b) FKBP и каталитическая субъединица кальцинейрина А (CnA); c) FKBP и циклофилин; d) FKBP и FKBP-рапамицин-ассоциированный белок (FRB); e) гираза В (GyrB) и GyrB; f) дигидрофолатредуктаза (DHFR) и DHFR; g) DmrB и DmrB; h) PYL и ABI; i) Cgy2 и CIB1; и j) GAI и GID1.

В некоторых случаях элемент димера (например, пары, связывающей димеризатор) рассматриваемого CAR получен из FKBP. Например, подходящий элемент пары, связывающий димеризатор, может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична следующей аминокислотной последовательности: MG VQVETISPGDGRTFPKRGQTCVVHYTGMLEDGKKFDSSRDRNPKFKFMLGKQEVIRGWEEGVAQM SVGQRAKLTISPDYAYGATGHPGIPPHATLVFDVELLKLE (SEQ ID NO: 78).

В некоторых случаях элемент пары, связывающей димеризатор, рассматриваемого CAR получен из каталитической субъединицы А кальцинейрина (также известной как PPP3CA; CALN; CALNA; CALNA1; CCN1; CNA1; PPP2B; каталитическая субъединица SAM-PRP; кальцинейрин А альфа; кальмодулин-зависимая изоформа альфа-субъединицы кальцинейрина А; протеинфосфатаза 2В; каталитическая субъединица, альфа-изоформа и т.д.). Например, подходящий элемент пары, связывающий димеризатор, может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%,

по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична следующей аминокислотной последовательности (домен PP2Ac):

LEESVALRIITEGASILRQEKNLLDIDAPVTVCGDHGGQFFDLMKLFVGGSPANT
RYLFLGDYVDRGYFSIECVLYLWALKILYPKTLFLLRGNHECRHLTEYFTFKQECKIKYS
ERVYDACMDAFDCLPLAALMNQQFLCVHGGLSPEINTLDDIRKLDRFKEPPAYGPMCDI
LWSDPLEDFGNEKTQEHFTHNTVRGCSYFYSPAVCEFLQHNNLLSILRAHEAQDAGYR
MYRKSQTTGFPSLITIFSAPNYLDVYNNKAAVLKYENNVMNIRQFNCSHPYWLPNFM
(SEQ ID NO: 79).

В некоторых случаях элемент димера (например, пары, связывающей димеризатор) получен из циклофилина (также известного как циклофилин А; PPIA; CYPA; CYPH; PPIаза А и т.д.). Например, подходящий элемент пары, связывающий димеризатор, может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична следующей аминокислотной последовательности:

MVNPTVFFDIAVDGEPLGRVVSFELFADKVPKTAENFRALSTGEKGFYKYGSCFHR
IPGFMCQGGDFTRHNGTGGKSIYGEKFEDENFILKHTGPGILSMANAGPNTNGSQFFICT
AKTEWLDGKHVVFVKVKEGMNIVEAMERFGSRNGKTSKKITIADCGQLE (SEQ ID NO:
80).

В некоторых случаях элемент димера (например, пары, связывающей димеризатор) получен из MTOR (также известного как с FKBP-рапамицин-ассоциированный белок; FK506-связывающий белок 12-рапамицин-ассоциированный белок 1; FK506-связывающий белок 12-рапамицин-ассоциированный белок 2; FK506-связывающий белок 12-ассоциированный с рапамициновым комплексом белок 1; FRAP; FRAP1; FRAP2; RAFT1; и RAPT1). Например, подходящий элемент пары, связывающий димеризатор, может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична следующей аминокислотной последовательности (также известной как «Frb»: Fbbp-рапамицин-связывающий домен):

MILWHEMWHEGLEEASRLYFGERNVKGMFEVLEPLHAMMERGPQTLKETSFNQ

AYGRDLMEAQEWCRKYMKSGNVKDLLQAWDLYYHVFRRISK (SEQ ID NO: 81).

В некоторых случаях элемент димера (например, пары, связывающей димеризатор) получен из GyrB (также известного как субъединица В ДНК-гиразы). Например, подходящий элемент пары, связывающий димеризатор, может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична непрерывному участку аминокислотной последовательности от приблизительно 100 аминокислот до приблизительно 200 аминокислот (а.к.), от приблизительно 200 а.к. до приблизительно 300 а.к., от приблизительно 300 а.к. до приблизительно 400 а.к., от приблизительно 400 а.к. до приблизительно 500 а.к., от приблизительно 500 а.к. до приблизительно 600 а.к., от приблизительно 600 а.к. до приблизительно 700 а.к. или от приблизительно 700 а.к. до приблизительно 800 а.к. следующей аминокислотной последовательности GyrB из *Escherichia coli* (или последовательности субъединицы В ДНК-гиразы из любого организма):

MSNSYDSSSIKVLKGLDAVRKRPGMYIGDIDDGTGLHHMVFEVVDNAIDEALA
 GHCKEIIVTIHADNSVSVQDDGRGIPTGIHPPEEGVSAAEVIMTVLHAGGKFDNSYKVS
 GLHGVGVSVVNALSQKLELVIQREGKIHQRQIYEHGVPQAPLAVTGETEKTGTMVRFWPS
 LETFTNVTEFEYEILAKRLRELSFLNSGVSIRLRDKRDGKEDHFHYEGGIKAFVEYLNKN
 КТPIHPNIFYFSTEKDGIGVEVALQWNDGFQENIYCFTNNIPQRDGGTHLAGFRAAMTRT
 LNAYMDKEGYSKKAKVVSATGDDAREGLIAVSVKVPDPKPFSSQTKDKLVSSEVKS
 AVEQQMNELLAEYLLNPTDAKIVVGKIIDAARAREAAARRAREMTRRKGALDLAGLPGKLA
 DCQERDPALSELYLVEGDSAGGSAKQGRNRKNQAILPLK GKILNVEKARFDKMLSSQE
 VATLITALGCGIGRDEYNPDKLRYHSIIIMTDADV DGS HIR TLLL TFFYRQMPEIVERGHV
 YIAQPPLYKVKKGKQEYIKDDEAMDQYQISIALD GATLHTNASAPALAGEALEKLVSE
 YNATQKMINRMERRYPKAMKELIYQPTL TEADLSDEQTVTRWVNALVSELNDKEQH
 GSQWKFDVHTNAEQNLFEPIVRVRTHGVD TDYPLDHEFITGGEYRRIC TLGEKLRGLLE
 EDAFIERGERRQPVASFEQALDWLVKESRRGLSIQRYKGLGEMNPEQLWETTMDPESRR
 MLRVTVKDAIAADQLFTTLMGDAVEPRRAFIEENALKAANIDI (SEQ ID NO: 82). В
 некоторых случаях элемент пары, связывающей димеризатор, содержит аминокислотную
 последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере
 приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере
 приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере
 приблизительно на 98% или 100% идентична аминокислотам 1-220 из вышеуказанной

аминокислотной последовательности GyrB из *Escherichia coli*.

В некоторых случаях элемент димера (например, пары, связывающей димеризатор) получен из DHFR (также известного как дигидрофолатредуктаза, DHFRP1 и DYR). Например, подходящий элемент пары, связывающий димеризатор, может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична следующей аминокислотной последовательности:

MVGS LN CIVAVSQNMGIGKNGDLPWPPLRNEFRYFQRMTTSSVEGKQNLVIMGKKT
WFSIPEKNRPLKGRINLVLSRELKEPPQGAHFLSRSLDDALKLTEQPELANKVDMVWIV
GGSSVYKEAMNHPGHLKLFVTRIMQDFESDTFFPEIDLEKYKLLPEYPGVLSDVQEEKGI
KYKFEVYEKND (SEQ ID NO: 83).

В некоторых случаях элемент димера (например, пары, связывающей димеризатор) получен из DmrB-связывающего домена (т.е. домена гомодимеризации DmrB). Например, подходящий элемент пары, связывающий димеризатор, может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична следующей аминокислотной последовательности:

MASRGVQVETISPGDGRTPKRGQTCVVHYTGMLEDGKKVDSSRDRNKPFKFM LGKQ
EVIRGWEEGVAQMSVGQRAKLTISPDYAYGATGHPGIIPPHATLVFDVELLKLE (SEQ ID
NO: 84).

В некоторых случаях элемент димера (например, пары, связывающей димеризатор) получен из белка PYL (также известного как рецептор абсцизовой кислоты и RCAR). Например, элемент рассматриваемой пары, связывающей димеризатор, может быть получен из белков, например, белков *Arabidopsis thaliana*: PYR1, RCAR1(PYL9), PYL1, PYL2, PYL3, PYL4, PYL5, PYL6, PYL7, PYL8 (RCAR3), PYL10, PYL11, PYL12, PYL13. Например, подходящий элемент пары, связывающий димеризатор, может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична любой из следующих аминокислотных последовательностей:

PYL10:

MNGDETKKVESEYIKKHRRHEL VESQCSSTLVKHIKAPLHLVWSIVRRFDEPQKYKPFIS
RCVVQGGKLEVGSVREVDLKSGLPATKSTEVLEILDDNEHILGIRIVGGDHRLKNYSSTI
SLHSETIDGKTGTLAIESFVVDVPEGNTKEETCFVVEALIQCENLNSLADVTERLQAESME
KKI (SEQ ID NO: 85).

PYL11:

METSQKYHTCGSTLVQTIDAPLSLVWSILRRFDNPQAYKQFVKTCNLSSGDGGEGSVRE
VTVVSGLPAEFSRERLDELDDESHVMMISIIGGDHRLVNYRSKTMAFVAADTEEKTVVV
ESYVVDVPEGNSEEETTSFADTIVGFNLKSLAKLSERVAHLKL (SEQ ID NO: 86).

PYL12:

MKTSQEQHVCGSTVVQTINAPLPLVWSILRRFDNPKTFKHFVKTCCKLRSGDGGE
GSVREVTVVS DLPASFSLERLDELDDESHVMVISIIGGDHRLVNYQSKTTVFVAAEEEEKT
VVVESYVVDVPEGNTEEETTLFADTIVGCNLRSLAKLSEKMMELT (SEQ ID NO: 87).

PYL13:

MESSKQKRCRSSVETIEAPLPLVWSILRSFDKPQAYQRFVKSC TMRS GGGGGK
GGEGKGSVRDVTLVSGFPADFSTERLEELDDESHVMVVSIIIGGNHRLVNYKSKTKVVAS
PEDMAKKTVVVESYVVDVPEGTSEEDTIFFVDNIIRYNL TSLAKLTKKMMK (SEQ ID
NO: 88).

PYL1:

MANSESSSPVNEEENSQRISTLHHQTMPSDLTQDEFTQLSQSIAEFHTYQLGNR
CSSLLAQRIHAPPETVWSVRRFDRPQIYKHFIKSCNVSEDFEMRVGCTRDNVISGLPA
NTRERLDLLDDRRVTGFSITGGEHRLRNYKSVTTVHRFEKEEEEEERIWTVVLESYVV
DVPEGNSEEDTRLFADTVIRLNLQKLASITEAMNRNNNNNNSSQVR (SEQ ID NO: 89).

PYL2:

MSSSPA VKGLTDEEQKTLEPVIKTYHQFEPDPTTCTSLITQRIHAPASVWVPLIRRF
DNPERYKHFVKRCRLISGDGDVGSVREVTVISGLPASTSTERLEFVDDDHRVLSFRVVG
GEHRLKNYKSVTSVNEFLNQDSGKVYTVVLESYTVDIPEGNTEEDTKMFVDTVVKLNL
QKLGVAATSAPMHDDE (SEQ ID NO: 90).

PYL3:

MNLAPIHDPSSSTTTTSSSTPYGLTKDEFSTLDSIIRTHHTFPRSPNTCTSLIAHRV
DAPAHAIWRFVRDFANPNKYKHFISCTIRVNGNGIKEIKVGTIREVSVVSGLPASTSVEI
LEVLDEEKRILSFRVLGGEHRLNNYRSVTSVNEFVLEKDKKKRVYSVVLESYIVDIPQG
NTEEDTRMFVDTVVKSNLQNLAVISTASPT (SEQ ID NO: 91).

PYL4:

MLAVHRPSSAVSDGDSVQIPMMIASFQKRFP SLSRDSTAARFHTHEVGPNQCCSA

VIQEISAPISTVWSVRRFDNPQAYKHFLKSCSVIGGDGDNVGLRQVHVVSGLPAASST
ERLDILDDERHVISFSVVGGDHRLSNYRSVTTLHPSPISGTVVVESYVVDVPPGNTKEET
CDFVDVIVRCNLQSLAKIAENTA AESKKKMSL (SEQ ID NO: 92).

PYL5:

MRSPVQLQHGSDATNGFHTLQPHDQTDGPIKRVCLTRGMHVPEHVAMHHTHDV
GPDQCCSSVVMIHAPPESVWALVRRFDNPKVYKNFIRQCRIVQGDGLHVGDREVMV
VSGLPAVSSTERLEILDEERHVISFSVVGGDHRLKNYRSVTTLHASDDEGTVVVESYIVD
VPPGNTTEEETLSFVDTIVRCNLQSLARSTNRQ (SEQ ID NO: 93).

PYL6:

MPTSIQFQRSSTAAEAANATVRNYPHHHQKQVQKVS LTRGMADVPEHVELSHT
HVVGPSQCFSVVVQDVEAPVSTVWSILSRFEHPQAYKHFVKSCHVVIGDGREVGSVREV
RVVSGLPAAFSLERLEIMDDDRHVISFSVVGGDHRLMNYKSVTTVHESEEDSDGKKRTR
VVESYVVDVPAGNDKEETCSFADTIVRCNLQSLAKLAENTSKFS (SEQ ID NO: 94).

PYL7:

MEMIGGDDTDTEMYGALVTAQSLRLRHLHHCRENQCTSVLVKYIQAPVHLVWS
LVRRFDQPQKYKPFISRCTVNGDPEIGCLREVN VKSGLPATTSTERLEQLDDEEHILGINII
GGDHRLKNYSSILTVHPEMIDGRSGTMVMESFVVDVPQGNTKDDTCYFVESLIKCNLKS
LACV SERLAAQDITNSIATFCNASNGYREKNHTETNL (SEQ ID NO: 95).

PYL8:

MEANGIENLTNPQNQEREFIRRHKKHELVDNQCSSTLVKHINAPVHIVWSLVRRFD
QPQKYKPFISRCVVKGNMEIGTVREVDVKSGLPATRSTERLELLDDNEHILSIRIVGGDH
RLKNYSSII SLHPETIEGRIGTLVIESFVVDVPEGNTKDETCYFVEALIKCNLKSLADISERL
AVQDTTESRV (SEQ ID NO: 96).

PYL9:

MMDGVEGGTAMYGGLETVQYVRTHHQHLCRENQCTSALVKHIKAPHLVWSL
VRRFDQPQKYKPFVSRCTVIGDPEIGSLREVN VKSGLPATTSTERLELLDDEEHILGIKIIG
GDHRLKNYSSILTVHPEIIEGRAGTMVIESFVVDVPEGNTKDETCYFVEALIRC NLKSLA
DVSERLASQDITQ (SEQ ID NO: 97).

PYR1:

MPSELTPEERSELKNSIAEFHTYQLDPGSCSSLHAQRIHAPPELVWSIVRRFDKPQ
TYKHFIKSCSVEQNFEMRVGCTRDVIVISGLPANTSTERLDILDDERRVTGFSIIGGEHRL
TNYKSVTTVHRFEKENRIWTVVLESYVVDMPGNS EDDTRMFADTVVKLNLQKLATV
AEAMARNSGDGSGSQVT (SEQ ID NO: 98).

В некоторых случаях элемент димера (например, пары, связывающей димеризатор) получен из белка ABI (также известного как нечувствительный к абсцизовой кислоте

белок). Например, элемент рассматриваемой пары, связывающей димеризатор, может быть получен из белков, таких как белки *Arabidopsis thaliana*: ABI1 (также известный как нечувствительный к абсцизовой кислоте белок 1, протеинфосфатаза 2С 56, AtPP2C56, P2C56 и PP2C ABI1) и/или ABI2 (также известный как P2C77, протеинфосфатаза 2С 77, AtPP2C77, нечувствительный к абсцизовой кислоте белок 2, протеинфосфатаза 2С ABI2 и PP2C ABI2). Например, подходящий элемент пары, связывающий димеризатор, может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична непрерывному участку аминокислотной последовательности от приблизительно 100 аминокислот до приблизительно 110 аминокислот (а.к.), от приблизительно 110 а.к. до приблизительно 115 а.к., от приблизительно 115 а.к. до приблизительно 120 а.к., от приблизительно 120 а.к. до приблизительно 130 а.к., от приблизительно 130 а.к. до приблизительно 140 а.к., от приблизительно 140 а.к. до приблизительно 150 а.к., от приблизительно 150 а.к. до приблизительно 160 а.к., от приблизительно 160 а.к. до приблизительно 170 а.к., от приблизительно 170 а.к. до приблизительно 180 а.к., от приблизительно 180 а.к. до приблизительно 190 а.к. или от приблизительно 190 а.к. до приблизительно 200 а.к. любой из следующих аминокислотных последовательностей:

ABI1:

MEEVSPAIAAGPFRPFSETQMDFTGIRLGKGYCNNQYSNQDSENGDLMVSLPETSSCSVSG
SHGSESRKVLISRINSPNLNMKESAAADIVVVDISAGDEINGS DITSEKKMISRTERSLFE
FKSVPLYGFTSICGRPEMEDAVSTIPRFLQSSSGSMLDGRFDPQSAAHFFGVYDGHGGS
QVANYCRERMHLALAEIEAKEKPMLCDGDTWLEKWKKALFNSFLRVDSEIESVAPETV
GSTSVVAVVFP SHIFVANCGDSRAVLCRGKTALPLSVDHKPDREDEAARIEAAGGKVIQ
WNGARVFGVLAMRSIGDRYLKPSIIPDPEVTAVKRVKEDDCLILASDGVDVMTDEE
ACEMARKRILLWHKKNVAGDASLLADERRKEGKDPAAAMSAEYLSKLAIQRGSKDN
ISVVVVDLKP RRKLKSKPLN (SEQ ID NO: 99).

ABi2:

MDEVSPA VAVPFRPFTDPHAGLRGYCNGESRVTLPESSCSGDGAMKDSSFEINTRQDSL
TSSSSAMAGVDISAGDEINGSDEFDPRSMNQSEKKVLSRTERSLFEFKCVPLYGVTSICG
RRPEMEDSVSTIPRFLQVSSSSLLDGRVTNGFNPHLSAHFFGVYDGHGGSQVANYCRER
MHLALTEEIVKEKPEFCDGDTWQEKWKKALFNSFMRVDSEIETVAHAPETV GSTSVVA
VVFP THIFVANCGDSRAVLCRGKTPLALSVDHKPDRDDEAARIEAAGGKVIRWNGARV
FGVLAMRSIGDRYLKPSVIPDPEVTSVRRVKEDDCLILASDGLWDVMTNEEVCDLARK

RILLWHKKNAMAGEALLPAEKRGEGKDPAAMSAAEYLSKMALQKGSKDNISVVVVDL
KGIRKFKSKSLN (SEQ ID NO: 100).

В некоторых случаях элемент димера (например, пары, связывающей димеризатор) получен из белка Cry2 (также известного как криптохром 2). Например, элемент рассматриваемого димера (например, пары, связывающей димеризатор) может быть получен из белков Cry2 из любого организма (например, растения), такого как, но не ограничиваясь ими, *Arabidopsis thaliana*. Например, подходящий элемент пары, связывающий димеризатор, может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична непрерывному участку аминокислотной последовательности от приблизительно 100 аминокислот до приблизительно 110 аминокислот (а.к.), от приблизительно 110 а.к. до приблизительно 115 а.к., от приблизительно 115 а.к. до приблизительно 120 а.к., от приблизительно 120 а.к. до приблизительно 130 а.к., от приблизительно 130 а.к. до приблизительно 140 а.к., от приблизительно 140 а.к. до приблизительно 150 а.к., от приблизительно 150 а.к. до приблизительно 160 а.к., от приблизительно 160 а.к. до приблизительно 170 а.к., от приблизительно 170 а.к. до приблизительно 180 а.к., от приблизительно 180 а.к. до приблизительно 190 а.к. или от приблизительно 190 а.к. до приблизительно 200 а.к. любой из следующих аминокислотных последовательностей:

Cry2 (*Arabidopsis thaliana*)

MKMDKKTIVWFRDLRIEDNPALAAAANEHGSVFPVFIWCPEEEGQFYPPGRASRW
WMKQSLAHLSQLKALGSDLTLIKTHNTISAILDCIRVTGATKVVFNHLYDPVSLVRDH
TVKEKLVERGISVQSYNGDLLYEPWEIYCEKGGKPFSTFNSYWKKCLDMSIESVMLPPPW
RLMPITAAAEAIWACSIIEELGLENEAEKPSNALLTRAWSPGWSNADKLLNEFIEKQLIDY
AKNSKKVVGNSTSLLSPYLHFGEISVRHVFQCARMKQIIWARDKNSEGEESADLFLRGIG
LREYSRYICFNFPFTHEQSLLSHLRFFPWDADVDKFKAWRQGRTGYPLVDAGMRELWA
TGWMHNRIRVIVSSFAVKFLLLPWKWGMKYFWDTLLDADLECDILGWQYISGSIIPDGH
ELDRLDNPALQGAKYDPEGEYIRQWLPELARLPTEWIHHPWDAPLTVLKASGVELGTN
YAKPIVDIDTARELLAKAISRTREAQIMIGAAPDEIVADSFEALGANTIKEPGLCPSVSSN
DQQVPSAVRYNGSKRVKPEEEEEERDMKKS RGFDERELFSTAESSSSSVFFVSQSCSLAS
EGKNLEGIQDSSDQITTSLGKNGCK (SEQ ID NO: 101)

В некоторых случаях элемент димера (например, пары, связывающей димеризатор) получен из белка CIB1 *Arabidopsis thaliana* (также известного как транскрипционный

фактор bHLH63). Например, подходящий элемент димера (например, пары, связывающий димеризатор) может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична непрерывному участку аминокислотной последовательности от приблизительно 100 аминокислот до приблизительно 110 аминокислот (а.к.), от приблизительно 110 а.к. до приблизительно 115 а.к., от приблизительно 115 а.к. до приблизительно 120 а.к., от приблизительно 120 а.к. до приблизительно 130 а.к., от приблизительно 130 а.к. до приблизительно 140 а.к., от приблизительно 140 а.к. до приблизительно 150 а.к., от приблизительно 150 а.к. до приблизительно 160 а.к., от приблизительно 160 а.к. до приблизительно 170 а.к., от приблизительно 170 а.к. до приблизительно 180 а.к., от приблизительно 180 а.к. до приблизительно 190 а.к. или от приблизительно 190 а.к. до приблизительно 200 а.к. следующей аминокислотной последовательности:

MNGAIGDLLLLNFPDMSVLERQRAHLKYLNPTFDSPLAGFFADSSMITGGEMDS
 YLSTAGLNLPMYGETTVEGDSRLSISPETTLGTGNFKKRKFDTETKDCNEKKKKMTM
 NRDDLVEEGEEEEKSKITEQNNGSTKSIKKMKHKAKKEENNFSNDSSKVTKELEKTDYIH
 VRARRGQATDSHSIAERVREKISERMKFLQDLVPGCDKITGKAGMLDEIINYVQSLQR
 QIEFLSMKLAIVNPRPDFDMDDDIFAKEVASTPMTVVPSPPEMVLSGYSHEMVHSGYSSEM
 VNSGYLHVNPMQQVNTSSDPLSCFNNGEAPSMWDSHVQNLYGNLGV (SEQ ID NO:
 102).

В некоторых случаях элемент димера (например, пары, связывающей димеризатор) получен из белка GAI *Arabidopsis thaliana* (также известного как нечувствительный к гибберелловой кислоте белок и белок GAI семейства DELLA). Например, подходящий элемент пары, связывающий димеризатор, может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична непрерывному участку аминокислотной последовательности от приблизительно 100 аминокислот до приблизительно 110 аминокислот (а.к.), от приблизительно 110 а.к. до приблизительно 115 а.к., от приблизительно 115 а.к. до приблизительно 120 а.к., от приблизительно 120 а.к. до приблизительно 130 а.к., от приблизительно 130 а.к. до приблизительно 140 а.к., от приблизительно 140 а.к. до приблизительно 150 а.к., от приблизительно 150 а.к. до приблизительно 160 а.к., от приблизительно 160 а.к. до приблизительно 170 а.к., от

приблизительно 170 а.к. до приблизительно 180 а.к., от приблизительно 180 а.к. до приблизительно 190 а.к. или от приблизительно 190 а.к. до приблизительно 200 а.к. следующей аминокислотной последовательности:

MKRDHHHHHHQDKKTMNNNEEDDGNMGDELLAVLGYKVRSSSEMADVAQKL
EQLEVMMNSNVQEDDLSQLATETVHYNPAELYTWLDSMLTDLNPPSSNAEYDLKAIPGD
AILNQFAIDSASSSNQGGGGDTYTTNKRLKCSNGVVETTTATAESTRHVVLVDSQENGV
RLVHALLACAEAVQKENLTVAEALVKQIGFLAVSQIGAMRKVATYFAEALARRIYRLSP
SQSPIDHSLSDTLQMHFYETCPYLKFAHFTANQAILEAFQGKKRVHVIDFSMSQGLQWP
ALMQALALRPGGPPVFRLTGIGPPAPDNFDYLHEVGCKLAHLAEAIHVEFEYRGFVANT
LADLDASMLELRPSEIESVAVNSVFELHKLLGRPGAIDKVLGVVNQIKPEIFTVVEQESN
HNSPIFLDRFTESLHYYSTLFDSLEGVPSGQDKVMSEVYLGKQICNVVACDGPDRVERH
ETLSQWRNRFGSAGFAAAHIGSNFAFKQASMLLALFNGGEGYRVEESDGCLMLGWHTR
PLIATSAWKLSTN (SEQ ID NO: 103).

В некоторых случаях элемент димера (например, пары, связывающей димеризатор) получен из белка *GID1 Arabidopsis thaliana* (также известного как рецептор гибберелловой кислоты *GID1*). Например, подходящий элемент димера может содержать аминокислотную последовательность, которая по меньшей мере приблизительно на 75%, по меньшей мере приблизительно на 80%, по меньшей мере приблизительно на 85%, по меньшей мере приблизительно на 90%, по меньшей мере приблизительно на 95%, по меньшей мере приблизительно на 98% или 100% идентична непрерывному участку аминокислотной последовательности от приблизительно 100 аминокислот до приблизительно 110 аминокислот (а.к.), от приблизительно 110 а.к. до приблизительно 115 а.к., от приблизительно 115 а.к. до приблизительно 120 а.к., от приблизительно 120 а.к. до приблизительно 130 а.к., от приблизительно 130 а.к. до приблизительно 140 а.к., от приблизительно 140 а.к. до приблизительно 150 а.к., от приблизительно 150 а.к. до приблизительно 160 а.к., от приблизительно 160 а.к. до приблизительно 170 а.к., от приблизительно 170 а.к. до приблизительно 180 а.к., от приблизительно 180 а.к. до приблизительно 190 а.к. или от приблизительно 190 а.к. до приблизительно 200 а.к. любой из следующих аминокислотных последовательностей:

GID1A:

MAASDEVNLIERSRTVVPLNTWVLISNFKVAYNILRRPDGTFNRHLAEYLDRKVT
ANANPVDGVFSFDVLIDRRINLLSRVYRPAYADQEQQPSILDLEKVPDGDIVPVILFFHGG
SFAHSSANSAYDTLCRRLVGLCKCVVSVNYRRAPENPYPCAYDDGWIALNWVNSRS
WLKSKKDSKVHIFLAGDSSGGNIAHNVALRAGESGIDVLGNILLNPMFGGNERTESEKS
LDGKYFVTVRDRDWWYWKAFLEPEDREHPACNPFSPRGKSLEGVSFPKSLVVVAGLDL

IRDWQLAYAEGCLKKAGQEVKLMHLEKATVGFYLLPNNNHFNVMDEISAFVNAEC
(SEQ ID NO: 104).

GID1B:

MAGGNEVNLNECKRIVPLNTWVLISNFKLAYKVLRRPDGSFNRLAEFLDRKVP
ANSFPLDGVFSFDHVDSTTNLLTRIQPASLLHQTRHGTLELTKPLSTTEIVPVLIFFHGGGS
FTHSSANSAYDTFCRRLVTICGVVVVSVDYRRSPEHRYPCAYDDGWNALNWVKS RVW
LQSGKDSNVYVYLAGDSSGGNIAHNVAVRATNEGKVLGNILLHPMFGGQERTQSEKT
LDGKYFVTIQDRDWYWRAYLPEGEDRDHPACNPFGRGQSLKGVNFPKSLVVVAGLD
LVQDWQLAYVDGLKKTGLEVNLLYLKQATIGFYFLPNNDHFHCLMEELNKFVHSIEDS
QSKSSPVLLTP (SEQ ID NO: 105).

GID1C:

MAGSEEVNLIESTKTVVPLNTWVLISNFKLAYNLLRRPDGTFNRHLAEFLDRKVPANANP
VNGVFSFDVIIDRQTNLLSRVYRPADAGTSPSITDLQNPVDGEIVPVIVFFHGGSF AHSSA
NSAIYDTLCRRLVGLCGAVVVSVNRYRRAPENRYPCAYDDGWAVLKWVNSSSWLRSKK
DSKVRIFLAGDSSGGNIVHNVAVRAMESRIDVLGNILLNPMFGGTERTESEKRLDGKYFV
TVRDRDWYWR AFLPEGEDREHPACSPFGPRSKSLEGLSFPKSLVVVAGLDLIQDWQLK
YAEGCLKKAGQEVKLLYLEQATIGFYLLPNNNHFTVMDEIAAFVNAECQ (SEQ ID NO:
106).

Как легко понять, CAR, идентифицированный путем скрининга библиотеки нуклеиновых кислот, кодирующих синтетические модульные полипептиды CAR, или библиотеки синтетических модульных полипептидов CAR, может быть модифицирован, например, путем добавления одного или более доменов (например, комодулирующих доменов), путем удаления одного или более доменов (например, путем удаления флуоресцентного репортера, используемого в процессе скрининга), путем разделения CAR на два полипептида (и добавления доменов димеризации), путем перегруппировки доменов и т.д.

В некоторых случаях библиотека может быть подвергнута скринингу для определения фенотипа, связанного с клеточным ответом на конкретную клеточную среду. В этой связи клеточный фенотип может быть определен по ответу клетки (например, на основании активации или ингибирования) на воздействие определенной среды. Например, ответ Т-клеток можно оценить в ответ на воздействие конкретной клеточной среды. Согласно некоторым вариантам реализации настоящего изобретения ингибирование Т-клеток можно оценить в ответ на воздействие микроокружения опухоли.

В некоторых случаях библиотека может быть подвергнута скринингу для определения фенотипа, связанного с локализацией клеток, например, под влиянием

хоуминга или нацеливания на клетку. Следовательно, может быть проведен скрининг, чтобы определить влияние элементов библиотеки на нацеливание на клетку. Например, клеточная библиотека может быть введена в организм-хозяина, и клетки библиотеки могут быть выделены из целевого положения организма-хозяина через некоторое время, чтобы оценить, какие клетки были успешно направлены к целевому положению. В некоторых случаях Т-клетки могут быть количественно исследованы, чтобы определить их нацеливание на опухоль в условиях *in vivo*.

В некоторых случаях библиотека может быть подвергнута скринингу для определения фенотипа, специфичного для состояния пациента. Специфичные для пациента состояния, подвергнутые скринингу указанным способом, будут сильно различаться и могут включать состояния, связанные с конкретным патологическим состоянием пациента, и могут включать скрининг библиотеки для идентификации конкретного элемента(ов) библиотеки, проявляющего усиленный или оптимальный фенотип, специфичный для состояния пациента. В некоторых случаях локализация элементов клеточной библиотеки может быть оценена после контакта с эксплантом или ксенотрансплантатом, полученным от пациента. Например, локализация Т-клеток клеточной библиотеки Т-клеток, экспрессирующих CAR, может быть оценена после контакта с ксенотрансплантатом опухоли, специфичной для пациента. В некоторых случаях пролиферация элементов клеточной библиотеки может быть оценена после контакта с эксплантом или ксенотрансплантатом, полученным от пациента. Например, пролиферация Т-клеток клеточной библиотеки Т-клеток, экспрессирующих CAR, может быть оценена после контакта с ксенотрансплантатом опухоли, специфичной для пациента. В некоторых случаях специфичный для пациента эксплантат или ксенотрансплантат может быть количественно исследован для определения увеличения или уменьшения жизнеспособности после контакта с клеточной библиотекой или конкретными элементами клеточной библиотеки. Например, гибель Т-клеток библиотеки Т-клеток, экспрессирующих CAR, может быть оценена после контакта с ксенотрансплантатом опухоли, специфичной для пациента. В этой связи библиотека может быть подвергнута скринингу для идентификации оптимального элемента(ов) библиотеки для лечения конкретного пациента.

В некоторых случаях фенотипы могут быть количественно исследованы в условиях *in vitro* с помощью динамической иммунизации антигеном. Под динамической иммунизацией антигеном подразумевают, что фенотип оценивают не только для определения присутствия или отсутствия антигена, и, следовательно, антиген может динамически варьироваться, например, динамически варьироваться в диапазоне

концентраций, динамически варьироваться в течение определенного периода времени и т.д. Например, уровни антигена могут быть протитрованы (например, с использованием различных концентраций), чтобы оценить фенотип, который подвергают скринингу, при различных дозах, т.е. для оценки ответа в зависимости от дозы. Антиген может быть представлен в разных концентрациях любым удобным способом. В качестве неограничивающего примера различные количества антигена могут быть представлены с использованием клеток, экспрессирующих антиген на разных уровнях, включая диапазон уровней. В некоторых случаях время применения антигена можно динамически варьировать, например, чтобы оценить фенотип в количественном исследовании в зависимости от времени или оценить кинетику фенотипа, который подвергают скринингу.

В некоторых случаях библиотека может быть подвергнута скринингу для определения отличительной черты фенотипа. В настоящей заявке термин «отличительная черта фенотипа» обычно относится к комбинации отдельных фенотипов. Например, в некоторых случаях клетка может иметь отличительную черту фенотипа, которая включает определенную морфологию в сочетании с экспрессией конкретного маркера. Отличительная черта фенотипа может сочетать фенотипы с аналогичными или различными фенотипическими категориями, например, отличительная черта фенотипа может включать экспрессию двух связанных, но разных маркеров клеточной поверхности, или отличительная черта фенотипа может включать экспрессию маркера клеточной поверхности и маркера клеточной пролиферации, или отличительная черта фенотипа может включать экспрессию маркера клеточной поверхности и конкретного секретлируемого маркера (например, цитокина), или отличительная черта фенотипа может включать экспрессию двух разных цитокинов и т.д. Любой подходящий фенотип, включая тот, который описан в настоящем документе, можно применять в качестве компонента отличительной черты фенотипа.

Идентификация синтетических модульных полипептидов, связанных с фенотипом

В настоящем изобретении предложены способы идентификации элементов библиотеки, которые связаны с определенным детектируемым фенотипом. Не желая быть связанными соответствием какой-либо теории, авторы настоящего изобретения полагают, что скоординированная сборка каждого мультимодульного синтетического полипептида наряду с каждой соответствующей многозвенной нуклеиновой последовательностью-штрихкодом обеспечивает сборку и последующую идентификацию каждого уникального синтетического модульного полипептида. Как описано выше, область нуклеиновой последовательности-штрихкода каждой нуклеиновой кислоты, кодирующей синтетический модульный полипептид, обеспечивает не только идентичность отдельных

модулей, которые составляют каждый синтетический модульный полипептид, но также специфичное расположение (именуемое в настоящем документе как архитектура) модулей. Следовательно, идентичность и архитектура каждого элемента библиотеки могут быть определены путем секвенирования области нуклеиновой последовательности-штрихкода.

Соответственно, скрининг библиотеки не обязательно должен быть выполнен с использованием физически разделенных элементов библиотеки, и элементы библиотеки можно объединять и подвергать скринингу одновременно. Объединение элементов библиотеки может быть выполнено в условиях *in vitro*, например, в пробирке или в образце клеток, выделенных из соответствующего организма, или в условиях *in vivo*, у животного или в ткани. После одновременного скрининга элементы библиотеки и/или их модули, связанные с фенотипом, могут быть идентифицированы путем идентификации и/или количественной оценки связанной области нуклеиновой последовательности-штрихкода. В некоторых случаях объединенный скрининг позволяет проводить скрининг большого количества уникальных элементов библиотеки, что нецелесообразно с помощью стандартного последовательного или параллельного скрининга. Количество уникальных элементов библиотеки, которые могут быть подвергнуты скринингу для определения фенотипа, будет зависеть от размера и сложности библиотеки и, следовательно, может варьироваться в диапазоне от 96 или менее до миллионов и более, включая, но не ограничиваясь перечисленными, например, 100 или более, 200 или более, 300 или более, 400 или более, 500 или более, 1000 или более, 2000 или более, 3000 или более, 4000 или более, 5000 или более, 6000 или более, 7000 или более, 8000 или более, 9000 или более, 10000 или более, 20000 или более, 30000 и более, 40000 и более, 50000 и более, 60000 и более, 70000 и более, 80000 или более, 90000 и более, 100000 и более и т.д.

В некоторых случаях количество или частота конкретной нуклеиновой последовательности-штрихкода может быть измерена для идентификации наиболее представленного модуля. Например, частота каждой нуклеиновой последовательности-штрихкода может быть количественно определена в объединенном образце, содержащем элементы библиотеки и связанные с ней нуклеиновые кислоты, кодирующие элементы библиотеки, так, что нуклеиновые последовательности-штрихкоды с наибольшей частотой обеспечат идентификацию тех модулей, которые наиболее представлены в образце. Согласно некоторым вариантам реализации настоящего изобретения такие образцы могут представлять собой клеточные образцы.

В некоторых случаях количество или частота конкретной многозвенной нуклеиновой последовательности-штрихкода (например, области последовательности-

штрихкода) может быть измерена для идентификации наиболее представленного модульного полипептида. Например, частота каждой многозвенной нуклеиновой последовательности-штрихкода может быть количественно определена в объединенном образце, содержащем элементы библиотеки и связанные с ней нуклеиновые кислоты, кодирующие элементы библиотеки, так, что многозвенные нуклеиновые последовательности-штрихкоды с наибольшей частотой обеспечивают идентификацию указанных модульных полипептидов, наиболее представленных в образце. Согласно некоторым вариантам реализации настоящего изобретения указанные образцы могут представлять собой клеточные образцы.

В некоторых случаях детектирование фенотипа и идентификация элементов библиотеки и их компонентов могут быть выполнены как часть интегрированного способа. Например, в некоторых случаях, фенотип можно детектировать методом проточной цитометрии, и элементы библиотеки могут быть идентифицированы с помощью секвенирования. Указанные интегрированные способы могут быть выполнены в комбинации с количественными исследованиями в условиях *in vitro* и/или *in vivo*, например, как показано на Фиг. 24, где FLOW-Seq используется как неограничивающий пример интегрированного способа.

В настоящей заявке термин «FLOW-seq» обычно относится к комбинации способов сортировки с помощью проточной цитометрии (например, FACS) с методами секвенирования (например, секвенирования следующего поколения) в одном связанном рабочем процессе. Любой удобный и подходящий способ сортировки с помощью проточной цитометрии и любой удобный и подходящий способ секвенирования можно применять в указанном методе FLOW-Seq. Например, в некоторых случаях, клеточная библиотека, экспрессирующая синтетические модульные полипептиды, содержащие нуклеиновые последовательности-штрихкоды, описанная в настоящем документе, может быть количественно исследована для определения фенотипа методом проточной цитометрии, и те клетки, которые имеют конкретный фенотип, могут быть отсортированы, и их нуклеиновые последовательности-штрихкоды впоследствии секвенированы для идентификации конкретных элементов библиотеки и/или для количественного определения частоты отдельных элементов библиотеки и/или их модулей, экспрессируемых в отсортированных клетках. Сортировка может быть выполнена любым удобным и подходящим способом, включая, например, сортировку в одну или более групп на основании фенотипа, детектированного с помощью проточной цитометрии. После сортировки секвенирование может быть выполнено непосредственно на отсортированном образце и/или отсортированной клетке, или отсортированный

образец и/или отсортированная клетка может быть размножена и/или культивирована перед сортировкой, например, для увеличения количества копий нуклеиновых кислот, кодирующих элементы библиотеки. Способы FLOW-seq были использованы, например, для фенотипического измерения уровней белка и идентификации родственных генетических элементов в бактериях (см., например, Kosuri et al. Proc Natl Acad Sci USA (2013) 110(34):14024-9 и Goodman et al. Science (2013) 342(6157):475-479), помимо этого секвенирование комбинировали с проточной цитометрией для корреляции Т-клеток, отсортированных на основании их функции, с соответствующими им секвенированными генами рецепторов Т-клеток (см., например, Han et al. Nature Biotechnology (2014) 32:684-692).

В некоторых случаях идентификация синтетического модульного полипептида, связанного с конкретным фенотипом, может включать хирургическое выделение ткани или органа из модели *in vivo*, в которую была введена библиотека. Например, в некоторых случаях, после периода времени, достаточного для проведения количественного исследования, орган или ткань могут быть выделены из животного-хозяина, и нуклеиновые кислоты, присутствующие в органе или ткани, могут быть секвенированы для идентификации отдельных элементов библиотеки, присутствующих в органе или ткани. В других случаях нуклеиновая кислота, выделенная из органа или ткани или животного-хозяина, может быть количественно определена, включая полуколичественную оценку, секвенирована для количественного определения относительной частоты или присутствия конкретного элемента библиотеки в органе или ткани. В других случаях нуклеиновая кислота, выделенная из органа или ткани или животного-хозяина, может быть количественно определена, включая полуколичественную оценку, секвенирована для количественного определения относительной частоты или присутствия конкретного модуля в органе или ткани.

В некоторых случаях, например, когда конкретный модуль высоко представлен после полуколичественной оценки или отдельный модуль идентифицирован как способствующий желательному фенотипу, может быть выполнен следующий раунд сборки новой библиотеки, в котором идентифицированный модуль включен во все вновь созданные элементы библиотеки (т.е. идентифицированный переменный модуль используется впоследствии как невариабельный модуль), и вновь созданную библиотеку подвергают скринингу для идентификации дополнительных модулей, которые влияют на фенотип совместно с первоначально идентифицированным модулем. Специалист в данной области техники легко поймет, в каких случаях итеративная сборка и скрининг библиотеки могут быть выполнены для развития библиотек и отдельных элементов

библиотеки с желательными фенотипами.

ПРИМЕРЫ

Следующие примеры приведены, чтобы предоставить специалистам в данной области техники полное описание настоящего изобретения, а также описание способов получения и применения настоящего изобретения, и не предназначены для ограничения объема настоящего изобретения согласно мнению авторов настоящего изобретения, также авторы настоящего изобретения не подразумевают, что эксперименты, приведенные ниже, представляют собой все или единственные эксперименты, которые были выполнены. Были предприняты усилия для обеспечения точности в отношении используемых чисел (например, количеств, температуры и т.д.), однако следует учитывать некоторые экспериментальные ошибки и отклонения. Если не указано иное, части представляют собой части по массе, молекулярная масса представляет среднюю молекулярную массу, температура представлена в градусах Цельсия и давление соответствует или приблизительно равно атмосферному давлению. Могут быть использованы стандартные аббревиатуры, например, п.о., пара (пары) оснований; тыс.п.о., тысяча (тысячи) пар оснований; пл, пиколитр (пиколитры); с или сек., секунда (секунды); мин, минута (минуты); ч или час, часы; а.к., аминокислота (аминокислоты); тыс.п.о., тысяча (тысячи) пар оснований; п.о., пара (пары) оснований; нук., нуклеотид (нуклеотиды); в/м, внутримышечный (внутримышечно); в/б, внутрибрюшинный (внутрибрюшинно); п/к, подкожный (подкожно); и т.п.

Пример 1: Конструирование и скрининг модульных библиотек комодулирующих доменов

Для получения фрагментов нуклеиновых кислот, кодирующих модули, содержащих нуклеиновые последовательности-штрихкоды, которые содержат необходимые элементы для конструирования библиотеки, нуклеиновые кислоты, кодирующие полипептидные модули (т.е. комодулирующие домены (т.е. костимулирующие или коингибирующие домены)), субклонировали в вектор для клонирования, подходящий для секвенирования и расщепления ферментами рестрикции типа PIS. После субклонирования вектор содержал: последовательность, кодирующую модуль, оптимальные линкерные последовательности Gly/Ser, фланкирующие последовательность, кодирующую модуль, специфичную для модуля нуклеиновую последовательность-штрихкод, подходящие для клонирования 3'-концевые гомологичные плечи, фланкирующие специфичную для модуля последовательность-штрихкод, сайт рестрикции BamHI между кодирующей модуль последовательностью и специфичной для модуля последовательностью-штрихкодом и сайты ферментов рестрикции типа PIS на 5'-

конце кодирующей модуль последовательности и на 3'-конце последовательности-штрихкода, специфичной для модуля (Фиг. 1). Субклонированные векторные вставки секвенировали для подтверждения идентичности вставленных комодулирующих доменов. Комодулирующие домены, используемые в данной комбинаторной библиотеке, и соответствующие им последовательности белка представлены в таблице 1.

Векторы для клонирования, содержащие кодирующую модуль последовательность, содержащую нуклеиновую последовательность-штрихкод, расщепляли с использованием фермента рестрикции типа Pst для высвобождения нуклеиновых кислот «с оптимальной последовательностью», кодирующих каждый полипептидный модуль (Фиг. 2). Представлен пример части вектора для клонирования, содержащей описанные элементы и последовательность, кодирующую костимулирующий домен CD28, до и после расщепления ферментом рестрикции типа Pst (Фиг. 3), который представляет общую конфигурацию каждой модульной плазмиды/фрагмента, использованных при конструировании библиотеки.

Вектор экспрессии (т.е. лентивирусный упаковочный вектор pHR (также называемый вектор-реципиент)) получали путем расщепления 3'-конца промотора вируса некроза селезенки (pSFFV) ферментом рестрикции в сайте BamHI. Конструкции элементов библиотеки собирали поэтапно (Фиг. 5). Клонирование In-Fusion[®] сначала использовали для вставки в вектор экспрессии последовательности, кодирующей одноцепочечный вариабельный фрагмент (scFv) и трансмембранные (TM) домены (scFV-TM), общие для всех элементов библиотеки. Реакционную смесь In-Fusion[®] трансформировали в компетентные клетки *E. coli*. Трансформированные клетки высаживали. Колонии отбирали и получали минипрепараты для экстракции ДНК, чтобы выделить сконструированную плазмиду. После успешного конструирования плазмиды для экспрессии, содержащей scFV-TM, каждый последовательный компонент многомодульного полипептида, включая репортер EGFP, вставляли в направлении 3'-конца относительно предыдущего компонента на отдельных стадиях в соответствии с Фиг. 5. Каждая стадия включала расщепление с помощью BamHI и сборку In-Fusion[®] с последующей трансфекцией реакционной смеси In-Fusion[®], отбор колонии и очистку плазмиды.

Как показано на Фиг. 5, каждый готовый элемент библиотеки заключительной реакции In-Fusion[®] содержал общий scFV-TM, соединенный со специфичной для элемента комбинацией двух комодулирующих доменов, соединенных с общим репортером (CD3z-EGFP), и пару специфичных для модуля нуклеиновых последовательностей-штрихкодов в обратной ориентации относительно комодулирующих доменов. Нуклеиновые

последовательности-штрихкоды фланкировали сайтами связывания праймеров, обеспечивающими амплификацию и/или секвенирование конкретной комбинации нуклеиновых последовательностей-штрихкодов, соответствующей комбинации комодулирующих доменов, специфичных для элементов библиотеки.

После завершения конструирования последней плазмиды для библиотеки отдельные элементы библиотеки трансфицировали в клетки НЕК-293, и лентивирусные частицы получали стандартными способами. Полученный лентивирус использовали для инфицирования Т-клеток, чтобы получить модифицированные иммунные клетки, экспрессирующие многомодульные полипептиды.

Модифицированные иммунные клетки сортировали методом проточной цитометрии на основании экспрессии репортера EGFP. Сортировку проводили для выделения популяции модифицированных клеток с одинаковыми уровнями экспрессии. Сортированные клетки с одинаковыми уровнями экспрессии многомодульных полипептидов использовали для последующих этапов функционального скрининга.

Данную общую стратегию использовали для создания четырех отдельных библиотек. Конструировали одномерные библиотеки (т.е. каждый элемент библиотеки содержал один комодулирующий домен) и двумерные библиотеки (т.е. каждый элемент библиотеки содержал два комодулирующих домена). Элементы двумерных библиотек собирали в соответствии с общей схемой, представленной на Фиг. 6. Четыре библиотеки и соответствующие им этапы скрининга были следующими:

1) Конструировали одномерную библиотеку из 20 элементов и использовали для проверки возможности количественного исследования функции Т-клеток с помощью FLOWseq. Каждый элемент библиотеки конфигурировали так, чтобы он содержал домен CD8, гибридный с модулем комодулирующего домена, гибридным с доменом CD3Z. Модифицированные Т-клетки распределяли на группы в соответствии с экспрессией ими репортера с помощью проточной цитометрии, и библиотеку подвергали скринингу для измерения дозозависимого ответа активации Т-клеток (CD69) на связанные с планшетом антигены.

2) Конструировали одномерную CD19-специфичную библиотеку из 62 элементов и подвергали скринингу в мышинной модели опухоли в условиях *in vivo*. Каждый элемент библиотеки конфигурировали так, чтобы он содержал CD19-специфичный домен, гибридный с модулем комодулирующего домена, гибридного с доменом CD3Z, который был гибризован с репортером EGFP.

3) Конструировали двумерную библиотеку, состоящую из 62×62 элементов. Каждый элемент библиотеки содержал CD19-специфичный домен, гибридный с

первым модулем комодулирующего домена, гибридованным со вторым модулем комодулирующего домена, гибридованным с доменом CD3Z, гибридованным с репортером EGFP.

4) Конструировали одномерную мезотелин-специфичную библиотеку из 62 элементов. Каждый элемент библиотеки содержал мезотелин-специфичный домен, гибридованный с доменом комодулирующего модуля, гибридованным с доменом CD3Z, гибридованным с репортером EGFP.

Объединенную одномерную CD19-специфичную библиотеку из 62 элементов подвергали функциональному скринингу для идентификации альтернативных комодулирующих последовательностей в химерных антигенных рецепторах (CAR). Каждый элемент библиотеки содержал CD19-специфичный scFv, который указывает на целевой антиген раковых клеток, один из модулей комодулирующих доменов, перечисленных в таблице 1, и первичный сигнальный домен CD3Z (Фиг. 7).

После антигенной стимуляции (при концентрации антигена 32 нг/мл, 125 нг/мл и 1000 нг/мл) модифицированных Т-клеток объединенной библиотеки, Т-клетки функционально сортировали методом проточной цитометрии на группы «высокой» или «низкой» стимуляции на основании экспрессии CD69 (Фиг. 8). Относительное обогащение специфичных последовательностей-штрихкодов, соответствующих индивидуальным комодулирующим доменам, определяли количественно в клетках с высокой и низкой стимуляцией путем секвенирования (при концентрации антигена 1000 нг/мл) (Фиг. 9). Данный подход позволил сравнить стимулирующий и ингибирующий эффект каждого комодулирующего домена в CD19-CD3Z-специфичных CAR и в отношении конкретной использованной антигенной стимуляции. Например, результаты скрининга дополнительно анализировали при разных уровнях вносимого антигена, что позволило быстро оценить зависимый от дозы ответ отдельных комодулирующих доменов (Фиг. 10).

Пример 2: Подтверждение полной сборки двумерной библиотеки из 61×61 элементов

Шестьдесят один переменный модуль CAR (таблица 2, представленная на Фиг. 25) собирали в библиотеку нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, кодирующих двумерные (2D) синтетические модульные полипептиды CAR, с помощью вложенной сборки. Библиотеку глубоко секвенировали с помощью метода секвенирования посредством синтеза (SBS) с использованием системы MiSeq (Illumina Inc., Хейворд, Калифорния, США) и определяли количество считываний каждого элемента собранной библиотеки (Фиг. 26).

Из всего 3721 возможного элемента библиотеки (т.е. вариантов CAR) все возможные элементы библиотеки детектировали путем секвенирования с максимальной частотой 1216 считываний и минимальной частотой 2 считывания. Среднее количество считываний по всей библиотеке составило 333 с медианой 311 считываний и стандартным отклонением 140. 90% элементов библиотеки были представлены считываниями, количество которых находилось в пределах двукратного значения медианы, и 98% элементов библиотеки были представлены считываниями, количество которых находилось в пределах трехкратного значения медианы.

Следовательно, результаты секвенирования подтвердили, что способ вложенной сборки способен обеспечить 2D-библиотеку, содержащую элементы библиотеки, которые представляют все возможные комбинации переменных модулей.

Пример 3: Нормирование частей библиотеки

Был разработан способ повышения распределения клонов в комбинаторной библиотеке с использованием нормирования частей библиотеки.

В любой реакции комбинаторной сборки некоторые части (например, модули) интегрированы в собранные продукты более эффективно, чем другие. Как следствие, собранные продукты (например, варианты белка), которые содержат части, которые интегрируются относительно неэффективно, представлены в недостаточном количестве или отсутствуют в готовой комбинаторной библиотеке. По аналогичным причинам другие собранные продукты (например, те варианты белка, которые содержат части, которые эффективно интегрируются) чрезмерно представлены и, следовательно, определяются с избыточной частотой во всех последующих количественных исследованиях.

Для решения этой проблемы был разработан способ улучшения сборки библиотек. Указанный способ включает проведение начальной сборки с использованием одинаковых объемов каждой части (вставки ДНК). Для этой цели создавали общую реакционную смесь, содержащую 1 мкл каждой части. После завершения комбинаторной сборки смесь продуктов определяли с помощью подходящего способа количественного исследования. Например, в данном примере использовали секвенирование следующего поколения (см. Фиг. 27), и рассчитывали долю каждой части в общей библиотеке (обозначенную как e_1, e_2, \dots в линейном уравнении, представленном на Фиг. 28). Линейную алгебру использовали для одновременного вычисления нового оптимизированного объема для каждой части (в общей реакционной смеси), и затем библиотеку повторно синтезировали, чтобы достичь более равномерного распределения вариантов.

Указанный подход подтверждали путем синтеза комбинаторной библиотеки CAR, в которой каждый CAR содержит два из 61 различного кодирующего домена,

соединенных в виде тандема (61^2 для 3721 варианта белка). Воспользовавшись преимуществом простой, линейной взаимосвязи между эффективностью, концентрацией и представленностью в собранной комбинаторной библиотеке, было продемонстрировано, что указанный способ можно применять для количественного контроля относительной частоты отдельных частей. Как показано на Фиг. 29, согласно прогнозам (черные столбики), 10-кратное изменение относительной концентрации набора частей в основной реакционной смеси для сборки приведет к 10-кратному изменению представленности частей в готовой композиции комбинаторной библиотеки. При измерении средней частоты 5 частей в собранной комбинаторной библиотеке (белые столбики) действительно наблюдали предсказанное влияние 10-кратного изменения. Полученная в результате нормированная большая комбинаторная библиотека (как показано на Фиг. 26) имела значительно улучшенное распределение вариантов в библиотеке (например, по сравнению с распределением варианта в библиотеке до нормирования (Фиг. 27)).

Несмотря на то, что настоящее изобретение было описано со ссылкой на конкретные варианты реализации, специалисты в данной области техники поймут, что могут быть осуществлены различные изменения, и эквиваленты могут быть заменены, не отступая от истинной сущности и объема настоящего изобретения. Помимо этого множество модификаций может быть выполнено для адаптации конкретной ситуации, материала, состава вещества, способа, этапа или этапов способа, чтобы достичь цели, сущности и объема настоящего изобретения. Подразумевается, что все указанные модификации включены в прилагаемую формулу изобретения.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ *ex vivo/in vitro* идентификации в клетке выбранного фенотипа, связанного с синтетическим модульным полипептидом, включающий:

а) введение библиотеки нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, в клетки-хозяева с получением гетерогенной популяции генетически модифицированных клеток-хозяев, причем указанная библиотека нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, содержит множество элементов, каждый из которых содержит нуклеотидную последовательность, кодирующую отличающийся синтетический модульный полипептид, где нуклеотидная последовательность содержит: кодирующую область, содержащую две или более кодирующих последовательности, кодирующих два или более переменных модуля, где две или более кодирующих последовательности находятся в одной рамке считывания друг с другом, и область нуклеиновой последовательности-штрихкода, содержащей две или более уникальных последовательности-штрихкода, где каждый переменный модуль в кодирующей области коррелирует с определенной уникальной последовательностью-штрихкодом в области нуклеиновой последовательности-штрихкода;

б) идентификацию генетически модифицированной клетки-хозяина в указанной гетерогенной популяции, которая экспрессирует выбранный фенотип в ответ на стимул;

в) из идентифицированной клетки-хозяина (б), идентификация и/или количественное определение области нуклеиновой последовательности-штрихкода путем секвенирования, тем самым идентифицируя синтетический модульный полипептид и/или его модуль, дающий выбранный фенотип;

где клетки-хозяева представляют собой Т-клетки, а синтетический модульный полипептид представляет собой модульный рецептор.

2. Способ получения библиотеки нуклеиновых кислот, содержащих нуклеиновые последовательности-штрихкоды, каждая из которых кодирует уникальный синтетический модульный полипептид, причем указанный способ включает:

приведение первого полинуклеотида, содержащего первую кодирующую модуль последовательность, соединенную с первой нуклеиновой последовательностью-штрихкодом, в контакт со вторым полинуклеотидом, содержащим вторую кодирующую модуль последовательность, соединенную со второй нуклеиновой последовательностью-штрихкодом, в условиях, достаточных для введения первого полинуклеотида во второй полинуклеотид на стыке между второй кодирующей последовательностью и второй нуклеиновой последовательностью-штрихкодом с получением бимодульного

полинуклеотида, содержащего нуклеиновые последовательности-штрихкоды,

причем указанный бимодульный полинуклеотид, содержащий нуклеиновые последовательности-штрихкоды, содержит вторую кодирующую модуль последовательность, соединенную с сохранением открытой рамки считывания с первой кодирующей модуль последовательностью, соединенной с первой нуклеиновой последовательностью-штрихкодом, соединенной со второй нуклеиновой последовательностью-штрихкодом,

где синтетический модульный полипептид представляет собой модульный рецептор.

3. Способ по пп. 1 или 2, отличающийся тем, что указанный модульный рецептор представляет собой полипептид химерного антигенного рецептора (CAR).

4. Способ по п.3, отличающийся тем, что указанное множество элементов содержит нуклеотидную последовательность, кодирующую модульный рецептор, содержащий модульный домен, выбранный из группы, включающей: антигенсвязывающий домен, специфичный связывающий домен или белок специфичного партнера по связыванию, костимулирующий домен, коингибирующий домен, внутриклеточный белок, трансмембранный домен и их комбинации.

5. Способ по п.3, отличающийся тем, что CAR содержит один или более ко-модулирующих домена, стимулирующих Т клетки или ингибирующих Т клетки.

6. Способ по п.5, отличающийся тем, что CAR содержит по меньшей мере один ко-модулирующий домен, имеющий по меньшей мере 85% идентичности последовательности с аминокислотной последовательностью SEQ ID NO: 19-42.

7. Способ по п. 3, отличающийся тем, что CAR содержит антигенсвязывающий домен, выбранный из анти-BCMA, анти-CD123, анти-CD138, анти-CD171, анти-CD19, анти-CD22, анти-CD30, анти-CD33, анти-CD7, анти-CD70, анти-CEA, анти-EGFRvIII, анти-EPSCAM, анти-EphA2, анти-ErbB, анти-FAP, анти-GD2, анти-GPC3, анти-HER2, анти-IL1RAP, анти-Kарра, анти-LeY, анти-Meso, анти-MG7, анти-MUC1, анти-NKG2D, анти-PSCA и анти-ROR1 антигенсвязывающих доменов.

8. Способ по п.3, отличающийся тем, что CAR содержит полипептид внутриклеточного сигнального домена содержащий один или более иммунорецепторных тирозиновых активирующих мотивов (ITAM).

9. Способ по п.8, отличающийся тем, что внутриклеточный полипептид сигнального домена имеет по меньшей мере 85% идентичность последовательности с аминокислотной последовательностью SEQ ID NO:107-132.

10. Способ по п.3, где CAR является условно-активным CAR содержащим домен димеризации в каждой своей цепи и где образующая димер пара выбрана из белка, связывающего FK506 (FKBP) и FKBP; FKBP и каталитической субъединицы кальцинейрина А; FKBP и циклофилина; FKBP и FKBP- рапамицин-ассоциированного белка; гиразы В (GyrB) и GyrA; дигидрофолат редуктазы (DHFR) и DHFR; DmrB и DmrA; Pyl и ABI; Cyt2 и CIB1; и GAI и GID1.

11. Способ по п. 1, где идентифицированный фенотип является клеточной локализацией или фенотипом, специфичным для состояния пациента.

12. Библиотека нуклеиновых кислот, содержащая нуклеиновые последовательности-штрихкоды, содержащая:

множество уникальных полинуклеотидов, каждый из которых содержит нуклеотидную последовательность, кодирующую уникальный синтетический модульный полипептид, причем каждый уникальный полинуклеотид содержит:

i) кодирующую область, кодирующую уникальный синтетический модульный полипептид, содержащий первую кодирующую последовательность, кодирующую первый модуль, соединенную с сохранением открытой рамки считывания со второй кодирующей последовательностью, кодирующей второй модуль,

ii) много-субъединичную область нуклеиновой последовательности-штрихкода, содержащую первую последовательность-штрихкод, специфичную для первой кодирующей последовательности, соединенную со второй нуклеиновой последовательностью-штрихкодом, специфичной для второй кодирующей последовательности, причем указанная первая и вторая нуклеиновые последовательности-штрихкоды расположены в обратном порядке от 5'-конца к 3'-концу по отношению к первой и второй кодирующим последовательностям; и

при этом секвенирование каждой субъединицы в много-субъединичной области нуклеиновой последовательности-штрихкода позволяет идентифицировать каждый уникальный синтетический модульный полипептид,

где синтетический модульный пептид является модульным рецептором.

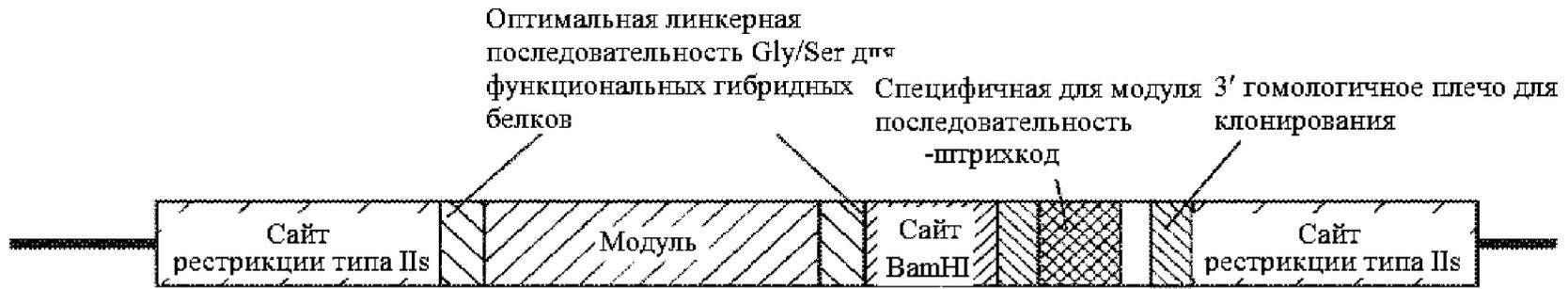
13. Библиотека по п. 12, отличающаяся тем, что указанная первая и вторая кодирующие последовательности соединены непосредственно без каких-либо промежуточных некодирующих нуклеотидов.

14. Библиотека по п. 12 или 13 которая объединена *in vitro*, причем объединенная библиотека содержит по меньшей мере 96 уникальных членов библиотеки.

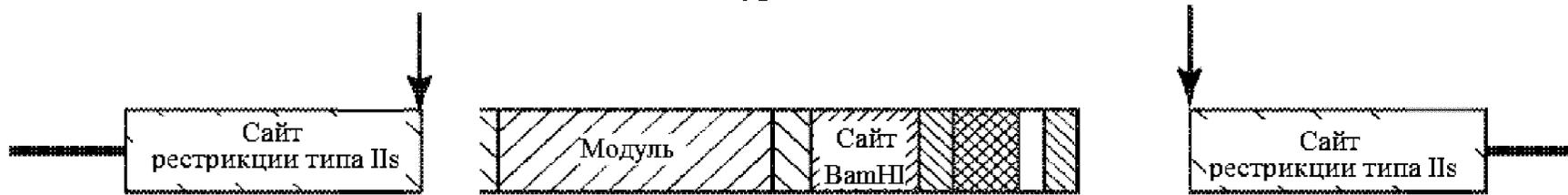
15. Клеточная библиотека, содержащая множество клеток, каждая из которых

содержит уникальный полинуклеотид из библиотека нуклеиновых кислот, содержащая нуклеиновые последовательности-штрихкоды, по любому из пп.12-14, при этом секвенирование каждой области нуклеиновой последовательности-штрихкода позволяет идентифицировать каждый уникальный синтетический модульный полипептид каждой клетки из библиотеки.

Фигура 1



Фигура 2



Фигура 3

До расщепления ферментом рестрикции типа IIS (ФР)

acctgcaaca**CTCCGGATCAGGATC**IGGTTTCGTGGGTGCGCAGCAAGAGAAGCAGACTGCT-
GCACAGCGACTACATGAACATGACCCCCAGACGGCCTGGCCCCACCAGAAAGCACTACCAGCCTTACGCC
CCTCCCAGAGACTTCGCCGCCTACAGATCTgg**CTCCGGATC**aggatcc**GGCAGTGGTAACACATAT**GTCTGCa
GATCCGGCAGTGGTAcaaggcaggt

После расщепления ФР IIS

CTCCGGATCAGGATCIGGTTTCGTGGGTGCGCAGCAAGAGAAGCAGACTGCTGCACAGCGAC
TACATGAACATGACCCCCAGACGGCCTGGCCCCACCAGAAAGCACTACCAGCCTTACGCCCTCCCAGAGA
CTTCGCCGCCTACAGATCTgg**CTCCGGATC**aggatcc**GGCAGTGGTAACACATAT**GTCTGCa**GATCCGGCAGT**
GGTA

Трансляция фрагмента, расщепленного ФР IIS

SGSGSGSWVRSKRSRLHSDYMNMTPRRPGPTRKHY-
QPYAPPRDFAAYRSGSGSGSGS...

Подпись

Курсивный шрифт: костимулирующий домен CD28

Жирный подчеркнутый шрифт: sdfsa: специфичный для модуля домен

Подчеркнутый шрифт: сайт фермента рестрикции типа IIS

Двойное подчеркивание: сайт BamHI

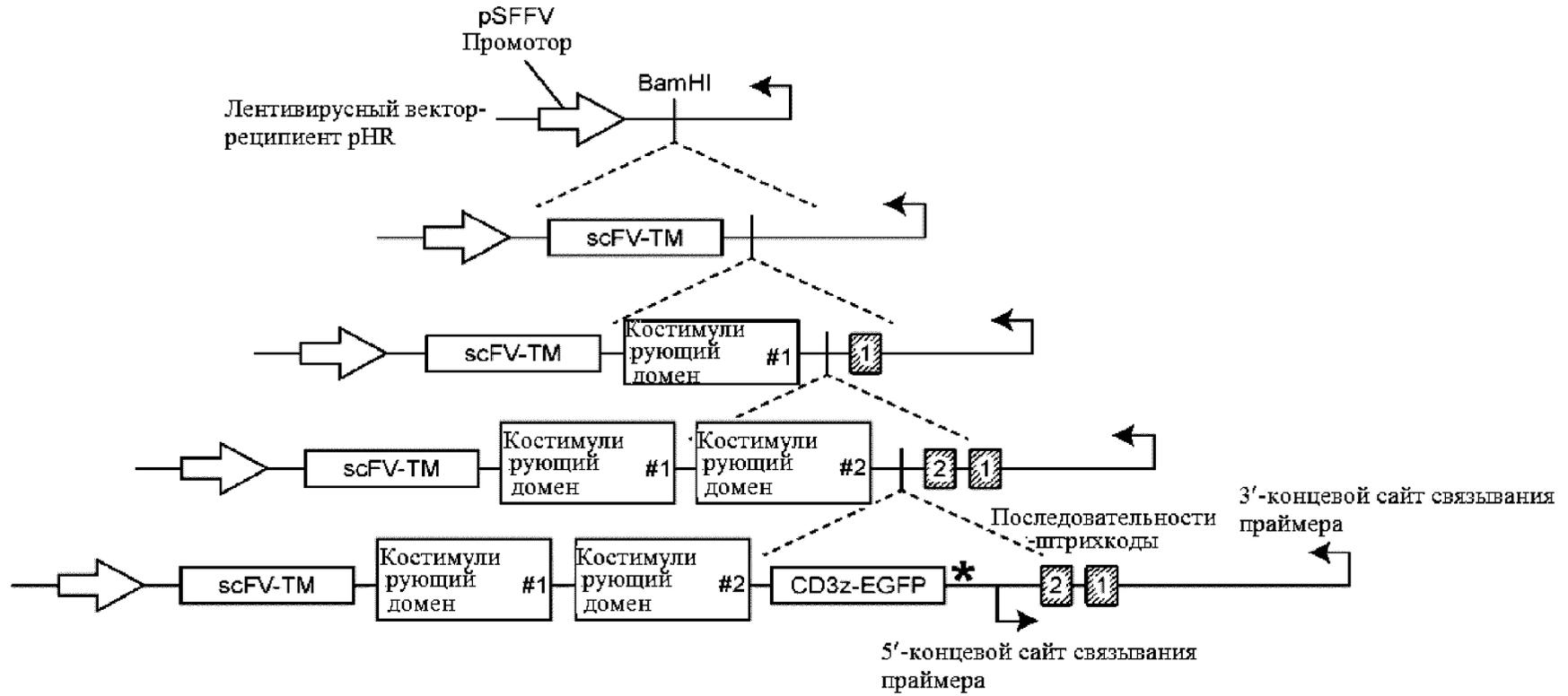
Жирный шрифт: линкерная последовательность Gly/Ser

Жирный курсивный шрифт: 3' гомологичное плечо для клонирования

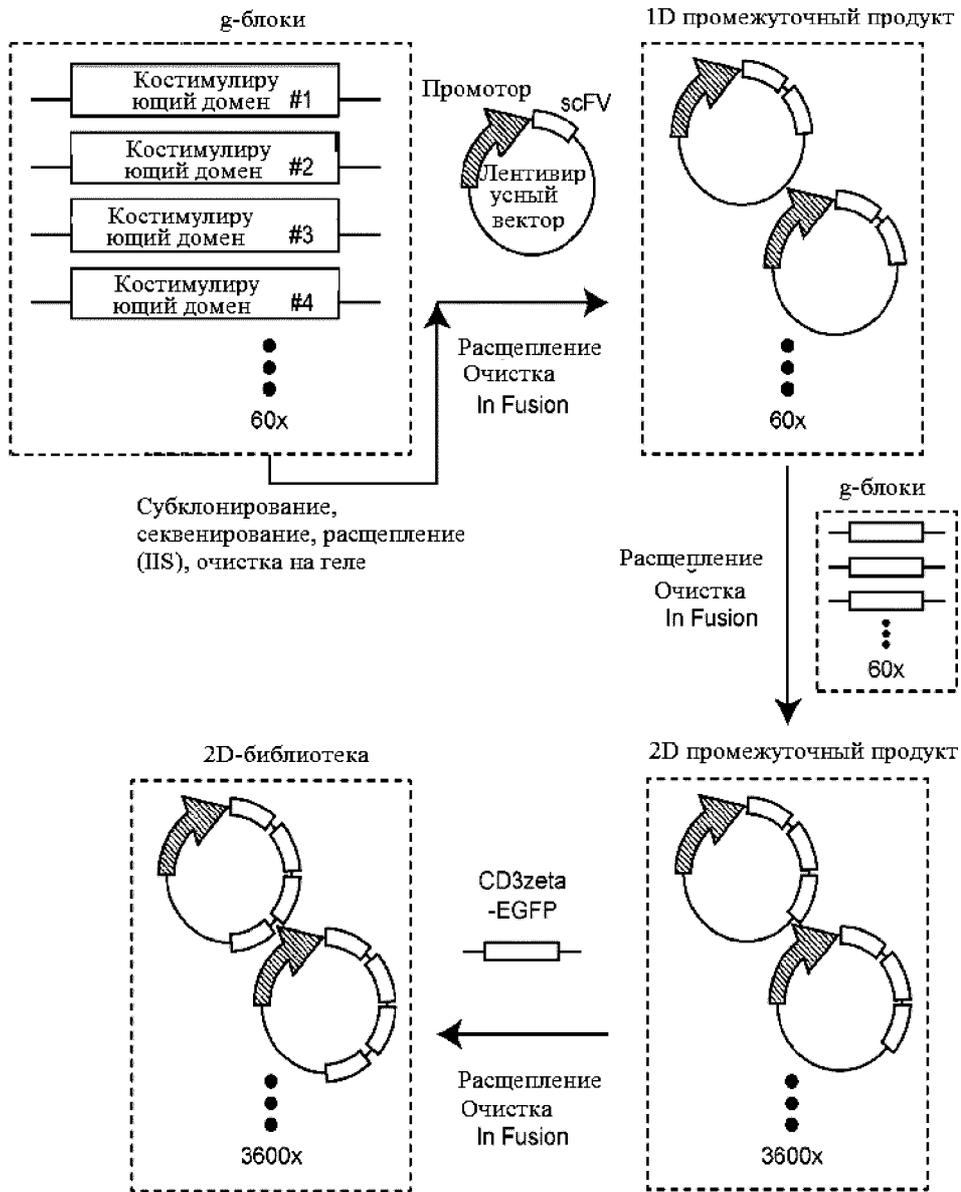
Фигура 4 (Cont.)

Индекс	Рецептор		Последовательность белка
24	TNR8_ человека (CD30)	187	RRACRKRIKQKLHLCYPVQTSQPKLELVDSRPRRSSTQLRSGASVTEP VAEERGLMSQPLMETCHSVGAAYLESPLQDASPAGGPPSSPRDLPEP RVSTEHTNNKIEKIYIMKADTVIVGTVKAELPEGRGLAGPAEPELEEELE ADHTPHYEQETEPPLGSCSDVMLSVEEEGKEDPLPTAASGK
25	TNR25_ человека (DR3)	108	TYRHOWPHKPLVTADEAGMEALTPPPATHLSPLDSAHTLLAPPDSSEK ICTVQLVGNWSWTPGYPETQEALCPQVTWSWDQLPSRALGPAAPTLS PESPAGSPAMMLQ
26	OX40	42	ALYLLRRDQRLPPDAHKKPPGGGSRFTPIQEEQADAHSTLAKI
27	NKG2D	51	MGWIRGRRSRHSWEMSEFHNYNLDLKKSDFSTRWQKQRCPPVVKSKC RENAS
28	TIM-1	48	KKYFFKKEVQQLSVSFSSLQIKALQNAVEKEVQAEDNIYIENSLYATD
29	CD160	22	SGFLQEKVWVMLVTSLVALQAL
30	BAFF-R	85	SWRRRQRRLRGASSAEAPDGDKDAPEPLDKVILSPGISDATAPAWPP PGEDPGTTPPGHMSVVPATELGSTELVTTKTAGPEQQ
31	TACI	111	ACFLKKRGDPCSCQPRSRPRQSPAKSSQDHAMEAGSPVSTSPPEVET CSFCFPECRAPTQESAVTPGTPDPTCAGRWGCHTRTTVLQPCPHIPD SGLGIVCVPAQEGGGPGA
32	CD72	95	MAEAITYADLRFVKAPLKKSISSRLGQDPGADDGGEITYENVOVPAVLG VPSSLASSVLGDKAARKSEQPTASWRAVTSFAVGRILPCRTTCLRY
33	CD22	141	KLORRWKRTQSQQGLQENSSGQSFVVRNKKVRRAPLSEPHSLGCV NPMMEDGISYTTLRFPENIPIRTGDAESSEMQRPPDCDDTVTYSALH KRQVGDYENVIPDFPEDEGIHYSELIQFGVGERPQAQENVVYVILKH
34	CD96	45	RKWCQYQKEIMERPPPFKPPPPPIKYTCIQEPNESDLPYHEMETL
35	NTB-A	84	LRKRRDSLSTQRTQGAESARNLEYVSVSPTNNTVYASVTHSNRET EIWTPRENDTITIVYSTINHESKESKPTFSRATALDNV
36	CRACC	88	WFLKREEQEEYIEKKRVDICRETPNICPHSGENTEYDITPHNTRILKE DPANTVYSTVEIPKKMENPHSLTMPDTPRLFAYENVI
37	Siglec-3.7.9	82	KTHRRKAARTAVGRNDTHPTTGSASPQHKKSKLHGPTETSSCSGAA PTVEMDEELHYASLNFHGMNPSKDTSTEYSEVRTQ
38	KLRG1	38	MTDSVIYSMLELPTATQAQNDYGPQQKSSSSSRPSCSCL
39	NKR-P1A	45	MDQQAIAELNLPDTSGPSSSPSSLPDVCQGGSPWHQFALKLSC
40	ILT2	168	LRHRRQGKHWTSTQRKADFQHPAGAVGPEPTDRGLQWRSSPAADA QEENLYAAVKHTQPEDGVEMDTRSPHDEDQAVTYAEVKHSRPRRE MASPPSPLSGEFLDTKDROAEEDRQMDTEAAASEAPQDVTYAQLHSL TLRREATEPPPSQEGPSPAVPSIYATLAIH
41	KIR2DL1	83	RWCSNKKNAAVMDQESAGNRTANSEDSDEQDPQEVYTYQLNHCVFT QRKITRPSQRPKTPPTDIIVYTELPNAESRSKVVSCP
42	KIR3DL1	84	HLWCSNKKNAAVMDQEPAGNRTANSEDSDEQDPPEEVTYAQLDH:CVF TQRKITRPSQRPKTPPTDTILYTELPNAKPRSKVVSCP
43	CD94-NKG2A	10	MAVFKTTLWR
44	CD300b	29	KGSQRVPEEPGEQPIYMNFSPLTKDMAT
45	CD300e	11	VNRPWAPPGR
46	TREM1	8	VTLRSFVP

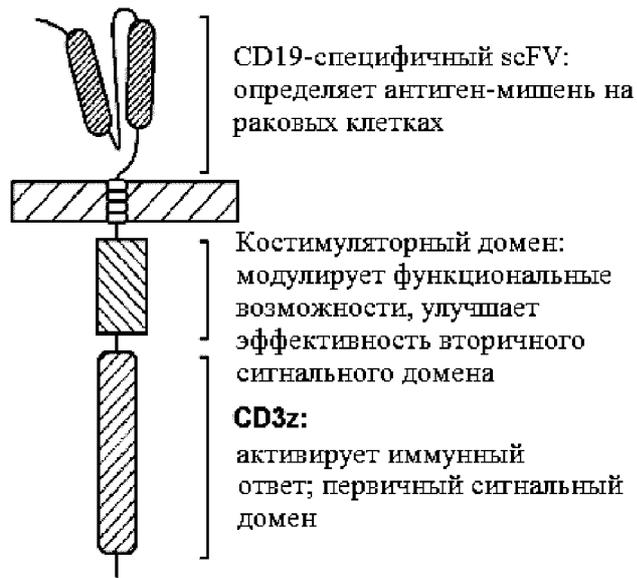
Фигура 5



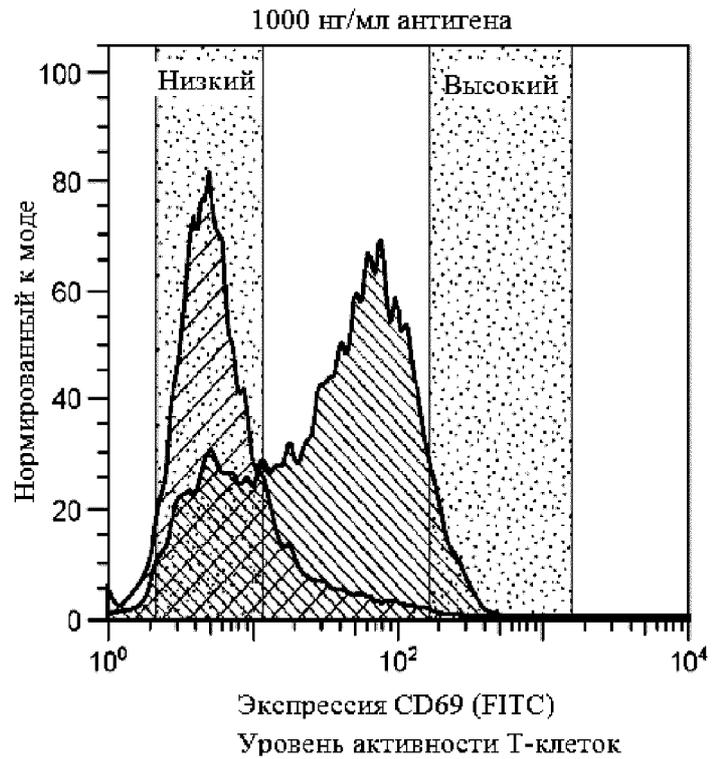
Фигура 6



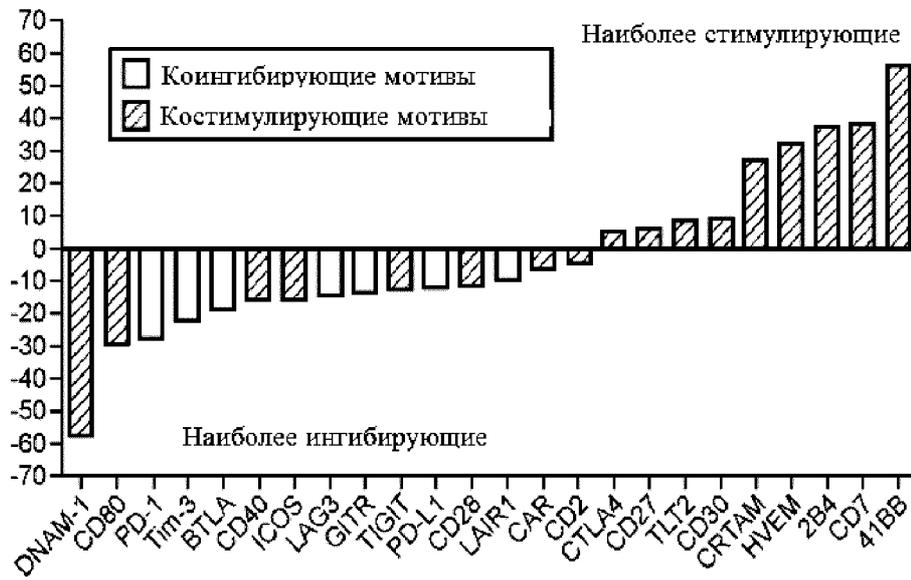
Фигура 7



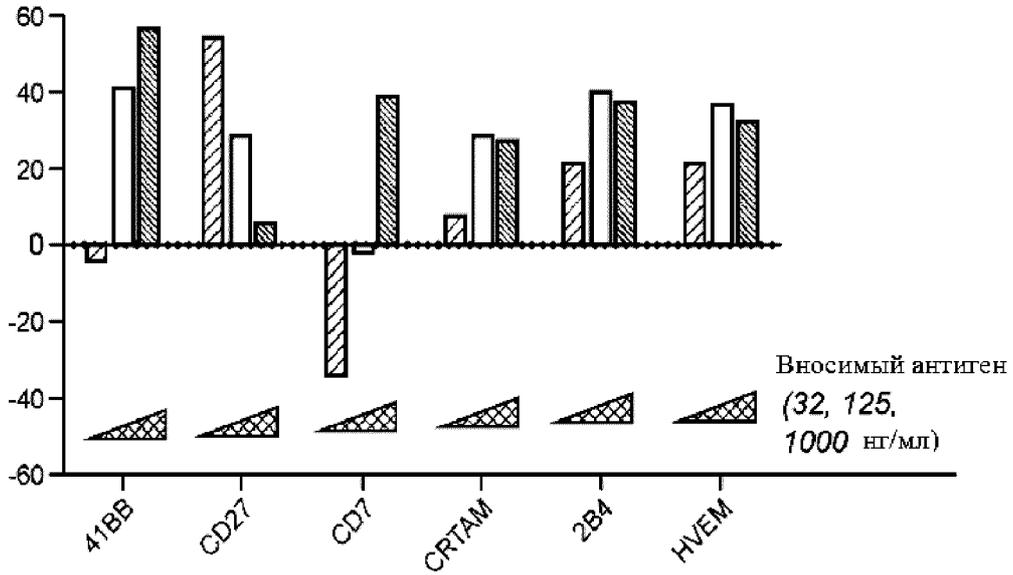
Фигура 8



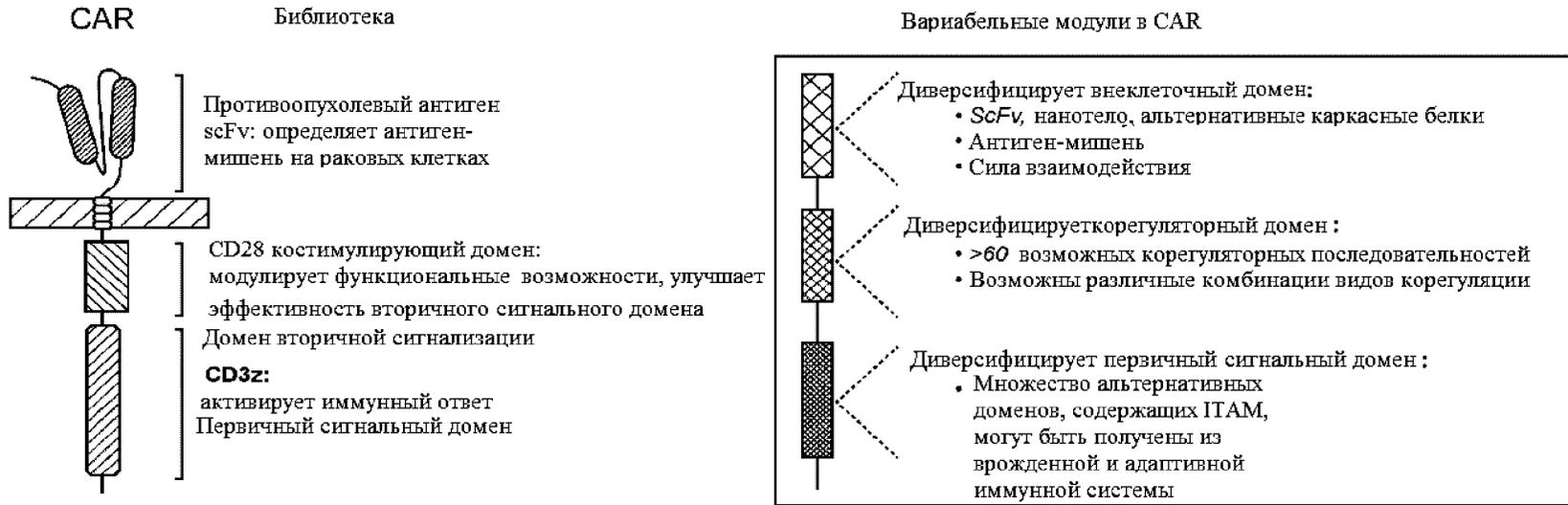
Фигура 9



Фигура 10



Фигура 11



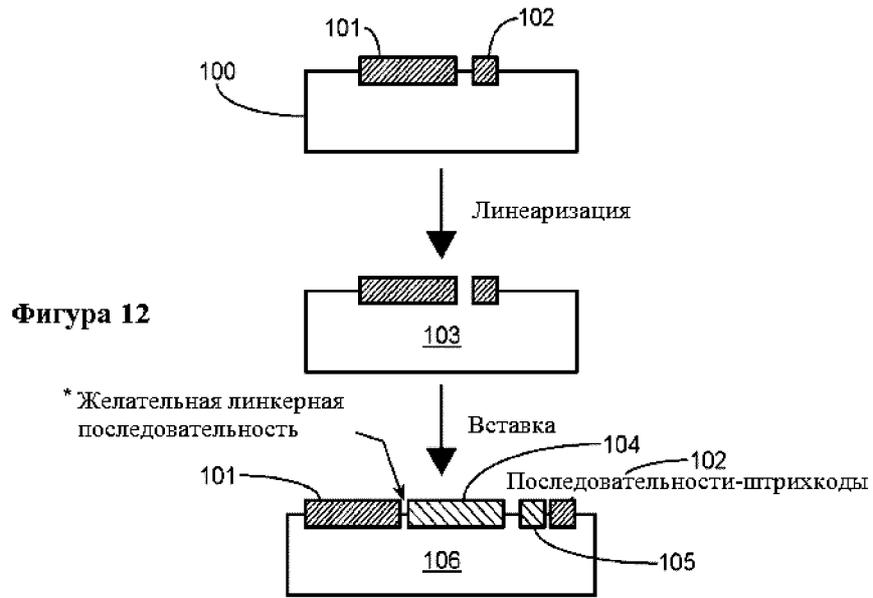
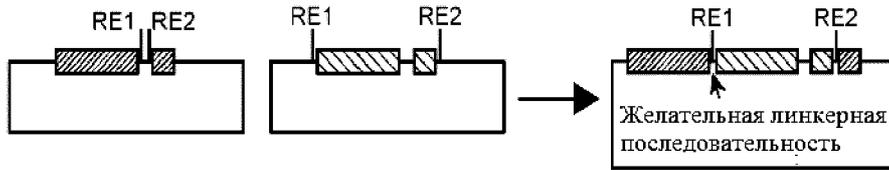
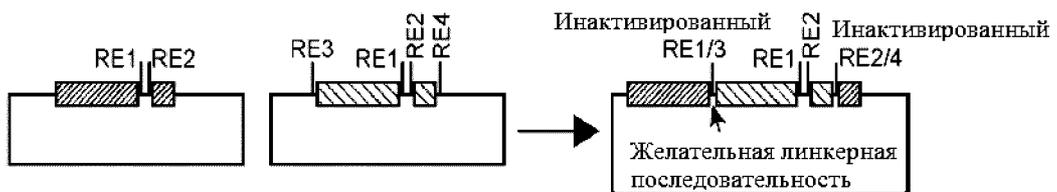


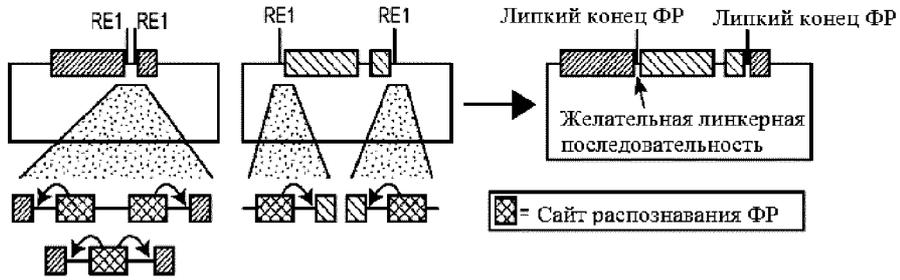
Figure 13



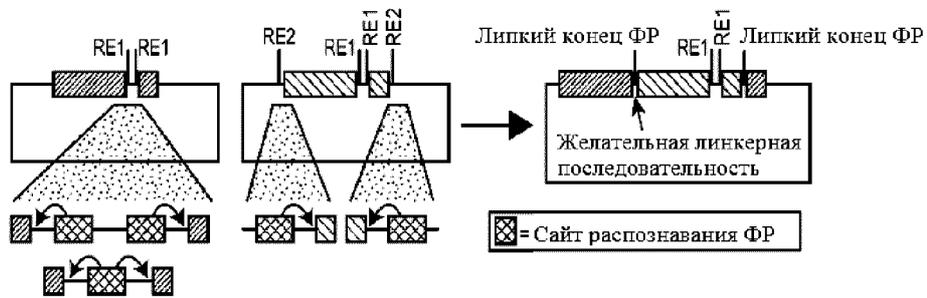
Фигура 14



Фигура 15

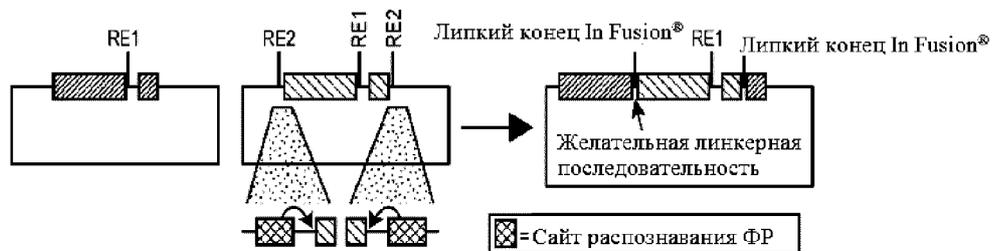


Фигура 16

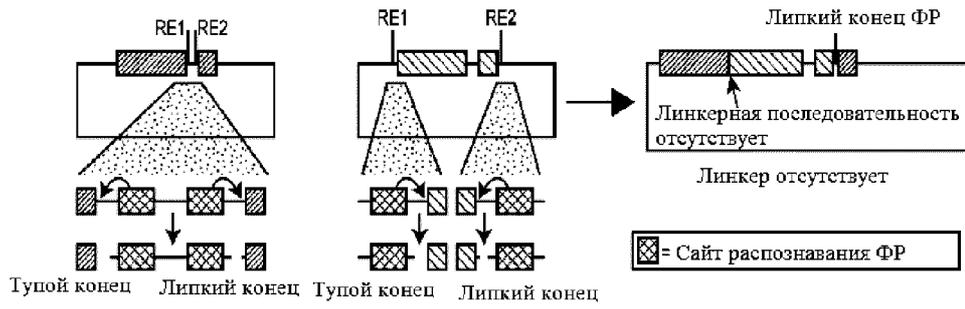


Фигура 17

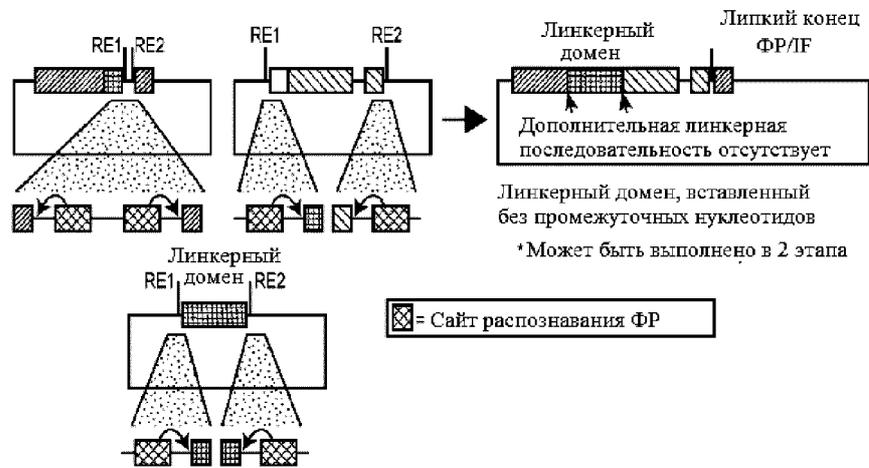
2 фермента рестрикции (тип IIs + In Fusion®, Gibson...)



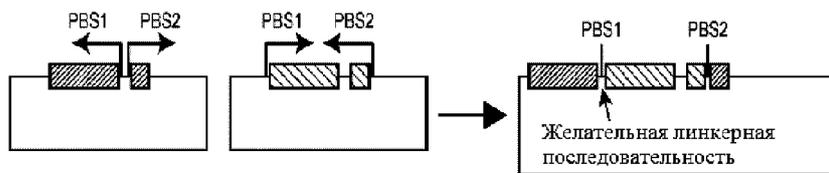
Фигура 18



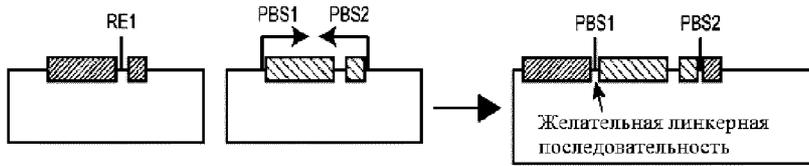
Фигура 19



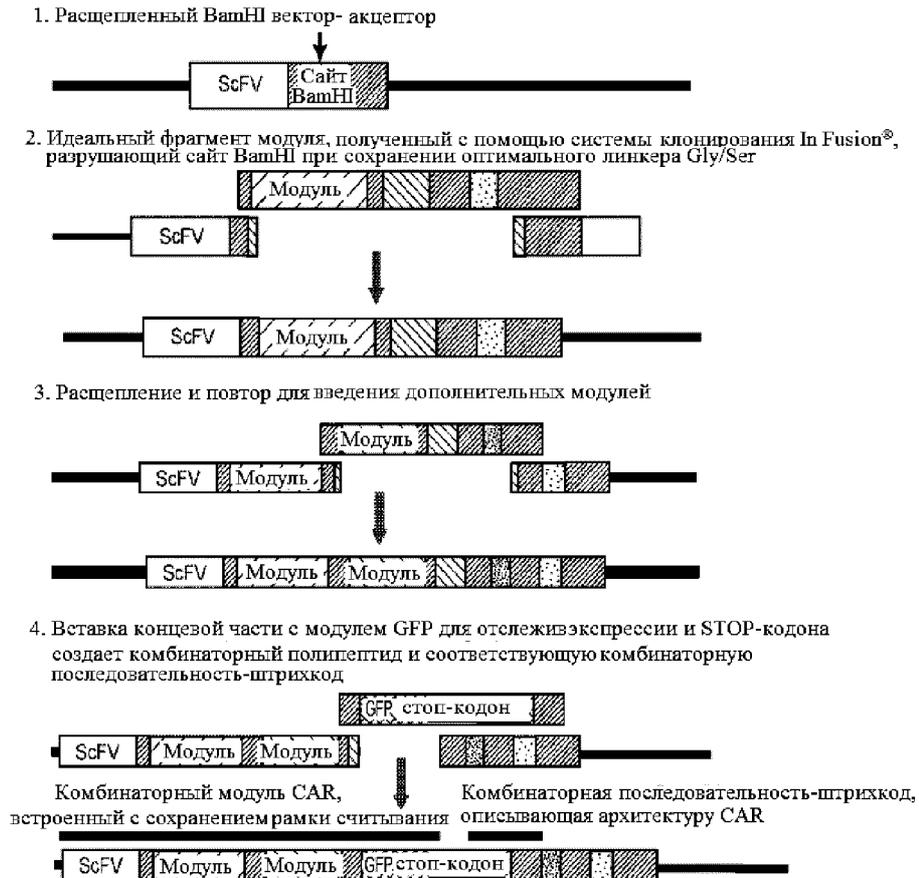
Фигура 20



Фигура 21

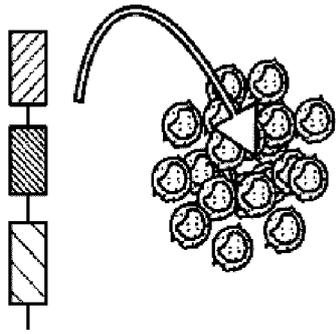


Фигура 22

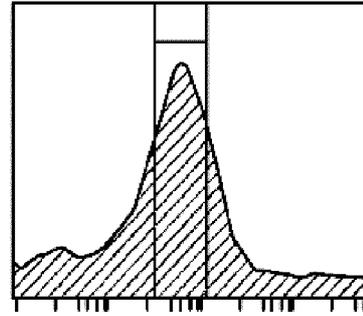


Фигура 23

Общая трансдукция библиотеки CAR в первичные T-клетки человека



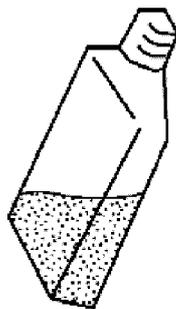
Сортировка для отбора одинакового уровня экспрессии CAR



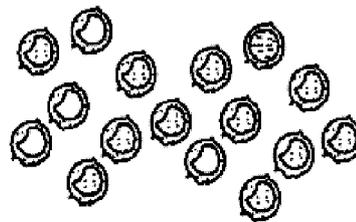
Уровень экспрессии CAR



Размножение популяции для последующих количественных исследований



Популяция первичных T-клеток человека, каждая из которых экспрессирует уникальный CAR

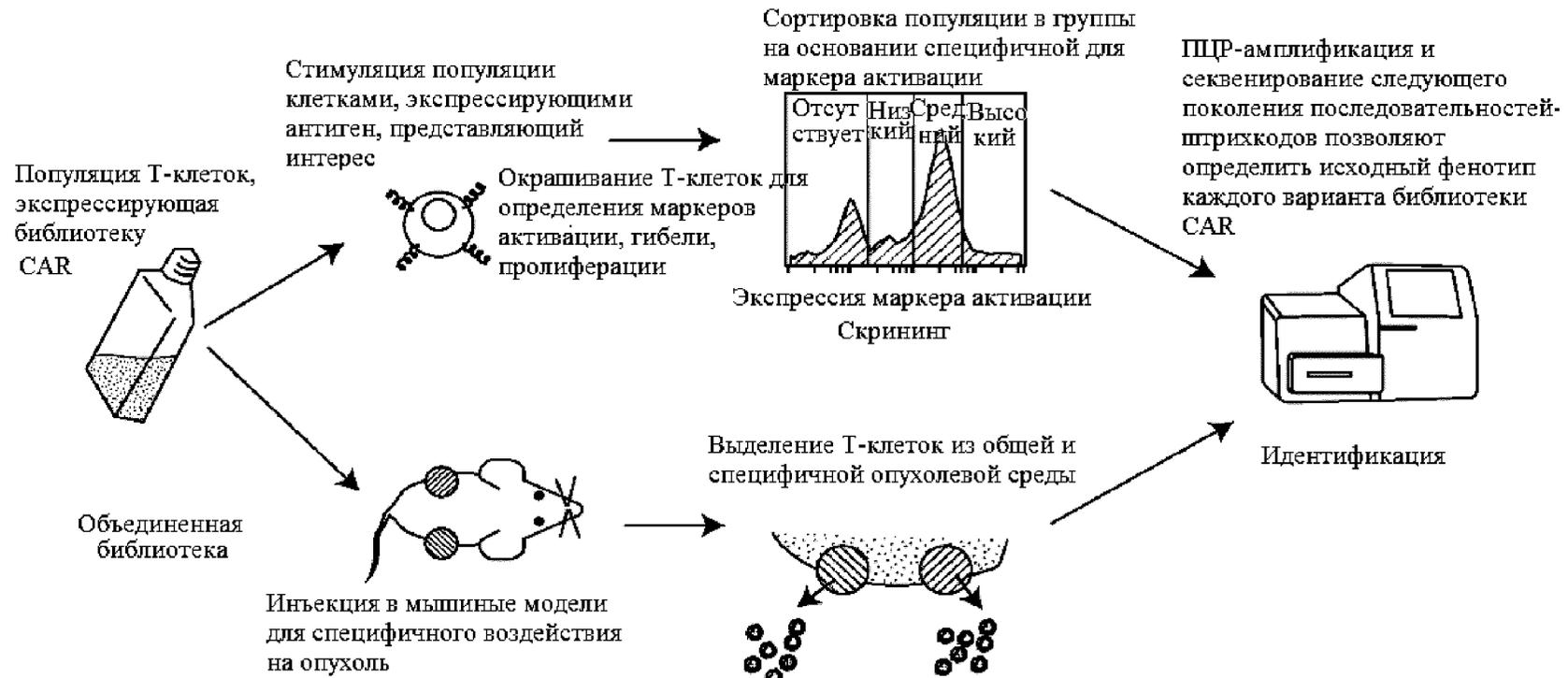


Объединенная библиотека



Фигура 24

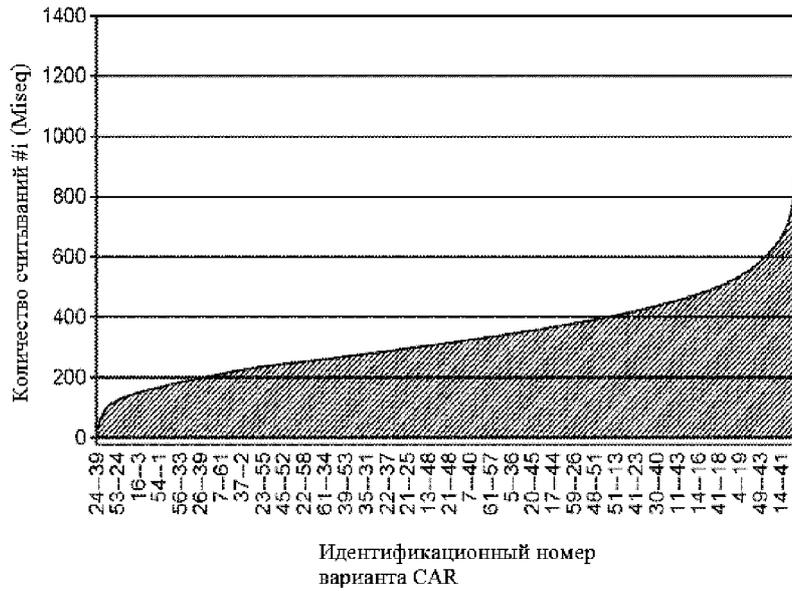
Анализ FLOW-seq позволяет определить микроскопические и макроскопические фенотипы, связанные с каждым членом библиотеки CAR к специфичному антигену или с опухолевой провокацией



Фигура 25

Номер	Модуль	Номер	Модуль	Номер	Модуль	Номер	Модуль
1	CD80	17	CRTAM	33	CD96	49	ILT4
2	PD1L	18	PDCD1	34	NTB-A	50	TLT-1
3	TRML2	19	LAIR1	35	CRACC	51	CD200R
4	ICOS	20	CXAR	36	Siglec-3.7.9	52	CD300a
5	TNF14	21	BTLA	37	KLRG1	53	CD300f
6	CTLA4	22	CD2	38	NKR-P1A	54	DC-SIGN
7	CD7	23	CD244	39	ILT2	55	B7-2
8	TNR9	24	TNR8	40	KIR2DL1	56	Allergin-1
9	CD28	25	TNR25	41	KIR3DL1	57	Pir-B
10	CD27	26	OX40	42	CD94-NKG2A	58	NC-14xSAG
11	LAG3	27	NKG2D	43	CD300b	59	TNR9_2
12	TNR18	28	TIM-1	44	CD300e	60	TNR9_3
13	CD226	29	CD160	45	TREM1	61	TNR9_4
14	TNR5	30	BAFF-R	46	TREM2		
15	HAVR	31	TAC1	47	ILT7		
16	TIGIT	32	CD22	48	ILT3		

Фигура 26



Фигура 29

