

(19)



Евразийское
патентное
ведомство

(21) 202090562 (13) A1

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ

(43) Дата публикации заявки
2020.08.10

(22) Дата подачи заявки
2018.09.11

(51) Int. Cl. C07K 16/00 (2006.01)
C07K 16/28 (2006.01)
C07K 16/46 (2006.01)
C07K 19/00 (2006.01)
C12N 15/13 (2006.01)
A61K 47/68 (2017.01)

(54) СВЯЗЫВАЮЩИЕ GPCR БЕЛКИ И ИХ СИНТЕЗ

(31) 62/556,863

(32) 2017.09.11

(33) US

(86) PCT/US2018/050511

(87) WO 2019/051501 2019.03.14

(71) Заявитель:

ТВИСТ БАЙОСАЙЕНС
КОРПОРЕЙШН (US)

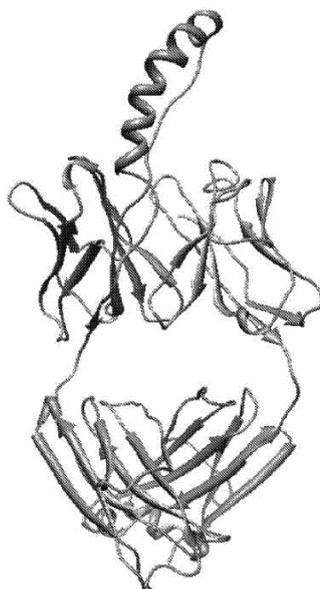
(72) Изобретатель:

Глэнвилль Джэйкоб (US)

(74) Представитель:

Строкова О.В., Глухарёва А.О., Лыгу
Т.Н., Угрюмов В.М., Христофоров
А.А., Гизатуллина Е.М., Гизатуллин
Ш.Ф., Костюшенкова М.Ю., Лебедев
В.В., Парамонова К.В. (RU)

(57) Настоящее изобретение относится к способам и композициям, относящимся к библиотекам сопряженных с G-белком рецепторов (GPCR), имеющих нуклеиновые кислоты, кодирующие остов, содержащий связывающий GPCR домен. Описанные в настоящем документе библиотеки включают мозаичные библиотеки, содержащие нуклеиновые кислоты, каждая из которых кодирует заданный вариант по меньшей мере одной заданной эталонной последовательности нуклеиновой кислоты. Дополнительно в настоящем документе описаны библиотеки белков, созданные при трансляции библиотек нуклеиновых кислот. Дополнительно в настоящем документе описаны клеточные библиотеки, экспрессирующие описанные в настоящем документе мозаичные библиотеки нуклеиновых кислот.



A1

202090562

202090562

A1

СВЯЗЫВАЮЩИЕ GPCR БЕЛКИ И ИХ СИНТЕЗ

ОПИСАНИЕ

Ссылка на родственную заявку

[0001] Согласно настоящей заявке испрашивается приоритет в соответствии с предварительной заявкой на выдачу патента США № 62/556863, поданной 11 сентября 2017 года, которая включена в настоящий документ посредством ссылки во всей своей полноте.

Область техники, к которой относится настоящее изобретение

[0002] Сопряженные с G-белком рецепторы (GPCR) вовлечены в широкий спектр заболеваний. Выработку антител к GPCR было сложно индуцировать из-за проблем с получением подходящего антигена, поскольку GPCR зачастую имеют низкий уровень экспрессии в клетках и очень нестабильны при очистке. Таким образом, существует потребность в улучшенных средствах для терапевтического вмешательства, целью которого являются GPCR.

Включение посредством ссылки

[0003] Все публикации, патенты и заявки на выдачу патента, упомянутые в настоящем описании, включены в настоящий документ посредством ссылки в той же степени, как если бы каждая отдельная публикация, патент или заявка на выдачу патента была специально и отдельно указана, как включенная посредством ссылки.

Краткое раскрытие настоящего изобретения

[0004] Настоящее изобретение относится к антителам, содержащим CDR-H3, который содержит последовательность под любым из SEQ ID NO: 2420-2436. Настоящее изобретение относится к антителам, содержащим CDR-H3, который содержит последовательность под любым из SEQ ID NO: 2420-2436; и причем антитело представляет собой моноклональное антитело, поликлональное антитело, биспецифическое антитело, мультиспецифическое антитело, привитое антитело, человеческое антитело, гуманизированное антитело, синтетическое антитело, химерное антитело, камелизированное антитело, одноцепочечные Fv (scFv), одноцепочечное антитело, Fab-фрагмент, F(ab')₂-фрагмент, Fd-фрагмент, Fv-фрагмент, однодоменное антитело, выделенный определяющий комплементарность участок (CDR), диатело, фрагмент, состоящий только из одного мономерного переменного домена, связанные

дисульфидными связями Fv (sdFv), интратело, антиидиотипическое (анти-Id) антитело или их антигенсвязывающие фрагменты. Настоящее изобретение относится к антителам, причем домен VH представляет собой IGHV1-18, IGHV1-69, IGHV1-8 IGHV3-21, IGHV3-23, IGHV3-30/33m, IGHV3-28, IGHV3-74, IGHV4-39 или IGHV4-59/61. Настоящее изобретение относится к антителам, причем домен VL представляет собой IGKV1-39, IGKV1-9, IGKV2-28, IGKV3-11, IGKV3-15, IGKV3-20, IGKV4-1, IGLV1-51 или IGLV2-14. Настоящее изобретение относится к способам ингибирования активности GLP1R, предусматривающим введение антител, которые описаны в настоящем документе. Настоящее изобретение относится к способам лечения метаболического нарушения, предусматривающим введение нуждающемуся в том субъекту антител, которые описаны в настоящем документе. В соответствии с некоторыми вариантами, антитело содержит CDR-Н3, содержащий последовательность под любым из SEQ ID NO: 2420-2436. Настоящее изобретение относится к способам лечения метаболического нарушения, причем метаболическое нарушение представляет собой сахарный диабет II типа или ожирение. Настоящее изобретение относится к нуклеиновым кислотам, кодирующим белок, содержащий последовательность под любым из SEQ ID NO: 2420-2436.

[0005] Настоящее изобретение относится к библиотекам нуклеиновых кислот, содержащим множество нуклеиновых кислот, причем каждая нуклеиновая кислота кодирует последовательность, которая при трансляции кодирует остов иммуноглобулина, причем остов иммуноглобулина содержит петлю CDR-Н3, которая содержит связывающий GPCR домен, и причем каждая нуклеиновая кислота содержит последовательность, кодирующую вариант последовательности связывающего GPCR домена. Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем длина петли CDR-Н3 составляет от приблизительно 20 до приблизительно 80 аминокислот. Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем длина петли CDR-Н3 составляет от приблизительно 80 до приблизительно 230 пар оснований. Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем остов иммуноглобулина дополнительно содержит один или несколько доменов, выбранных из переменного домена легкой цепи (VL), переменного домена тяжелой цепи (VH), константного домена легкой цепи (CL) и константного домена тяжелой цепи (CH). Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем домен VH представляет собой IGHV1-18, IGHV1-69, IGHV1-8 IGHV3-21, IGHV3-23, IGHV3-30/33m, IGHV3-28, IGHV3-74, IGHV4-39 или IGHV4-59/61. Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем домен VL представляет собой IGKV1-39, IGKV1-9, IGKV2-28, IGKV3-11, IGKV3-15, IGKV3-20, IGKV4-1, IGLV1-51 или IGLV2-14.

Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем длина домена VH составляет от приблизительно 90 до приблизительно 100 аминокислот. Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем длина домена VL составляет от приблизительно 90 до приблизительно 120 аминокислот. Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем длина домена VH составляет от приблизительно 280 до приблизительно 300 пар оснований. Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем длина домена VL составляет от приблизительно 300 до приблизительно 350 пар оснований. Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем библиотека содержит по меньшей мере 105 неидентичных нуклеиновых кислот. Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем остов иммуноглобулина содержит один домен иммуноглобулина. Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем остов иммуноглобулина содержит пептид не более чем из 100 аминокислот. Настоящее изобретение относится к библиотекам векторов, содержащим библиотеки нуклеиновых кислот, которые описаны в настоящем документе. Настоящее изобретение относится к клеточным библиотекам, содержащим библиотеки нуклеиновых кислот, которые описаны в настоящем документе.

[0006] Настоящее изобретение относится к библиотекам нуклеиновых кислот, содержащим множество нуклеиновых кислот, причем каждая нуклеиновая кислота кодирует последовательность, которая при трансляции кодирует связывающий GPCR домен, и причем каждая нуклеиновая кислота содержит последовательность, кодирующую другой связывающий GPCR домен длиной от приблизительно 20 до приблизительно 80 аминокислот. Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем длина связывающего GPCR домена составляет от приблизительно 80 до приблизительно 230 пар оснований. Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем связывающий GPCR домен сконструирован на основе конформационных взаимодействий лигандов, пептидных взаимодействий лигандов, низкомолекулярных взаимодействий лигандов, внеклеточных доменов GPCR или антител, целью которых являются GPCR. Настоящее изобретение относится к библиотекам векторов, содержащим библиотеки нуклеиновых кислот, которые описаны в настоящем документе. Настоящее изобретение относится к клеточным библиотекам, содержащим библиотеки нуклеиновых кислот, которые описаны в настоящем документе.

[0007] Настоящее изобретение относится к библиотекам белков, содержащим множество белков, причем каждый из этих белков из множества белков содержит остов иммуноглобулина, причем остов иммуноглобулина содержит петлю CDR-H3, которая

содержит вариант последовательности связывающего GPCR домена. Настоящее изобретение относится к библиотекам белков, причем длина петли CDR-H3 составляет от приблизительно 20 до приблизительно 80 аминокислот. Настоящее изобретение относится к библиотекам белков, причем остов иммуноглобулина дополнительно содержит один или несколько доменов, выбранных из переменного домена легкой цепи (VL), переменного домена тяжелой цепи (VH), константного домена легкой цепи (CL) и константного домена тяжелой цепи (CH). Настоящее изобретение относится к библиотекам белков, причем домен VH представляет собойIGHV1-18, IGHV1-69, IGHV1-8 IGHV3-21, IGHV3-23, IGHV3-30/33m, IGHV3-28, IGHV3-74, IGHV4-39 или IGHV4-59/61. Настоящее изобретение относится к библиотекам белков, причем домен VL представляет собойIGKV1-39, IGKV1-9, IGKV2-28, IGKV3-11, IGKV3-15, IGKV3-20, IGKV4-1, IGLV1-51 или IGLV2-14. Настоящее изобретение относится к библиотекам белков, причем длина домена VH составляет от приблизительно 90 до приблизительно 100 аминокислот. Настоящее изобретение относится к библиотекам белков, причем длина домена VL составляет от приблизительно 90 до приблизительно 120 аминокислот. Настоящее изобретение относится к библиотекам белков, причем множество белков применяют для получения библиотеки пептидомиметиков. Настоящее изобретение относится к библиотекам белков, причем библиотека белков содержит пептиды. Настоящее изобретение относится к библиотекам белков, причем библиотека белков содержит иммуноглобулины. Настоящее изобретение относится к библиотекам белков, причем библиотека белков содержит антитела. Настоящее изобретение относится к клеточным библиотекам, содержащим библиотеки белков, которые описаны в настоящем документе.

[0008] Настоящее изобретение относится к библиотекам белков, содержащим множество белков, причем множество белков содержит последовательность, содержащую различные связывающие GPCR домены, и причем длина каждого связывающего GPCR домена составляет от приблизительно 20 до приблизительно 80 аминокислот. Настоящее изобретение относится к библиотекам белков, причем библиотека белков содержит пептиды. Настоящее изобретение относится к библиотекам белков, причем библиотека белков содержит иммуноглобулины. Настоящее изобретение относится к библиотекам белков, причем библиотека белков содержит антитела. Настоящее изобретение относится к библиотекам белков, причем множество белков применяют для получения библиотеки пептидомиметиков. Настоящее изобретение относится к клеточным библиотекам, содержащим библиотеки белков, которые описаны в настоящем документе.

[0009] Настоящее изобретение относится к библиотекам векторов, содержащим описанную в настоящем документе библиотеку нуклеиновых кислот. Настоящее

изобретение относится к клеточным библиотекам, содержащим описанную в настоящем документе библиотеку нуклеиновых кислот. Настоящее изобретение относится к клеточным библиотекам, содержащим описанную в настоящем документе библиотеку белков.

Краткое описание чертежей

[0010] На **фиг. 1** изображена схема поверхностей взаимодействия лиганда сопряженного с G-белком рецептора (GPCR).

[0011] На **фиг. 2A** изображена первая схема остова иммуноглобулина.

[0012] На **фиг. 2B** изображена вторая схема остова иммуноглобулина.

[0013] На **фиг. 3** изображена схема мотива для размещения в остове.

[0014] На **фиг. 4** изображена схема GPCR.

[0015] На **фиг. 5** изображены схемы сегментов для сборки клональных фрагментов и неклональных фрагментов.

[0016] На **фиг. 6** изображены схемы сегментов для сборки клональных фрагментов и неклональных фрагментов.

[0017] На **фиг. 7** представлена схема стадий, на которой продемонстрирован иллюстративный технологический процесс синтеза генов, который раскрыт в настоящем документе.

[0018] На **фиг. 8** проиллюстрирован пример компьютерной системы.

[0019] **Фиг. 9** представляет собой блок-схему, иллюстрирующую архитектуру компьютерной системы.

[0020] **Фиг. 10** представляет собой схему, на которой продемонстрирована сеть, которая может включать множество компьютерных систем, множество сотовых телефонов и карманных компьютеров, а также сетевое хранилище данных (NAS).

[0021] **Фиг. 11** представляет собой блок-схему многопроцессорной компьютерной системы, использующей общее пространство виртуальных адресов памяти.

[0022] На **фиг. 12A-12C** изображены последовательности остовов иммуноглобулина.

[0023] На **фиг. 13** изображены последовательности остовов сопряженных с G-белком рецепторов.

[0024] **Фиг. 14** представляет собой график нормализованных результатов считывания для библиотеки переменного домена тяжелых цепей.

[0025] **Фиг. 15** представляет собой график нормализованных результатов считывания для библиотеки переменного домена легких цепей.

[0026] **Фиг. 16** представляет собой график нормализованных результатов считывания для библиотеки определяющего комплементарность участка 3 тяжелой цепи.

[0027] **Фиг. 17А** представляет собой диаграмму каркасных участков легкой цепи, проанализированных в отношении фолдинга.

[0028] **Фиг. 17В** представляет собой диаграмму каркасных участков легкой цепи, проанализированных в отношении термостабильности.

[0029] **Фиг. 17С** представляет собой диаграмму каркасных участков легкой цепи, проанализированных в отношении экспонирования мотивов с использованием метки FLAG.

[0030] **Фиг. 17D** представляет собой диаграмму каркасных участков легкой цепи, проанализированных в отношении экспонирования мотивов с использованием метки His.

[0031] **Фиг. 18А** представляет собой диаграмму каркасных участков тяжелой цепи, проанализированных в отношении фолдинга.

[0032] **Фиг. 18В** представляет собой диаграмму каркасных участков тяжелой цепи, проанализированных в отношении стабильности.

[0033] **Фиг. 18С** представляет собой диаграмму каркасных участков тяжелой цепи, проанализированных в отношении экспонирования мотивов с использованием метки FLAG.

[0034] **Фиг. 18D** представляет собой диаграмму каркасных участков тяжелой цепи, проанализированных в отношении экспонирования мотивов с использованием метки His.

[0035] **Фиг. 18Е** представляет собой диаграмму каркасных участков тяжелой цепи, проанализированных в отношении экспрессии.

[0036] **Фиг. 18F** представляет собой диаграмму каркасных участков тяжелой цепи, проанализированных в отношении специфичности отбора.

[0037] На **фиг. 19А-19С** показаны изображения сопряженных с G-белком рецепторов, визуализированных с помощью флуоресцирующих антител.

[0038] На **фиг. 20А-20С** показаны изображения сопряженных с G-белком рецепторов, визуализированных с помощью автофлуоресцирующих белков.

[0039] На **фиг. 21А** показана схема остова иммуноглобулина, содержащего домен VH, присоединенный к домену VL с помощью линкера.

[0040] На **фиг. 21В** показана схема полнодоменной архитектуры остова иммуноглобулина, содержащей домен VH, присоединенный к домену VL с помощью линкера, лидерную последовательность и последовательность рIII.

[0041] На **фиг. 21С** показана схема четырех каркасных элементов (FW1, FW2, FW3, FW4) и 3 переменных элементов (L1, L2, L3) CDR для домена VL или VH.

Краткое раскрытие настоящего изобретения

[0042] В настоящем раскрытии используют, если не указано иное, традиционные методики молекулярной биологии, которые известны специалистам в настоящей области техники. Если не указано иное, все технические и научные термины, используемые в настоящем документе, имеют то же значение, которое обычно понимают специалисты в настоящей области техники.

[0043] Определения

[0044] В настоящем раскрытии различные варианты осуществления представлены в формате диапазонов. Следует понимать, что описание в формате диапазонов приведено лишь для удобства и краткости, и его не следует рассматривать как жесткое ограничение объема каких-либо вариантов осуществления. Соответственно, описание диапазона следует рассматривать как конкретно раскрывающее все возможные поддиапазоны, а также отдельные числовые значения в пределах этого диапазона до десятой доли нижнего предела, если контекст явно не предписывает иное. Например, описание диапазона, такого как от 1 до 6, следует рассматривать как конкретно раскрытые поддиапазоны, такие как от 1 до 3, от 1 до 4, от 1 до 5, от 2 до 4, от 2 до 6, от 3 до 6 и т. д., а также отдельные значения в этом диапазоне, например, 1,1, 2, 2,3, 5 и 5,9. Это применимо независимо от широты диапазона. Верхний и нижний пределы этих промежуточных диапазонов могут быть независимо включены в меньшие диапазоны и также охватываются в настоящем раскрытии с учетом любого специально исключенного предела в указанном диапазоне. В тех случаях, когда указанный диапазон включает один или оба предела, диапазоны, исключаящие один или оба этих включенных пределов, также включены в настоящее раскрытие, если контекст явно не предписывает иное.

[0045] Применяемая в настоящем документе терминология предназначена лишь для описания конкретных вариантов осуществления и не предназначена для ограничения какого-либо варианта осуществления. Подразумевается, что применяемые в настоящем документе формы единственного числа также включают формы множественного числа, если контекст явно не предписывает иное. Дополнительно будет понятно, что термины «содержит» и/или «содержащий» при использовании в настоящем описании определяют наличие указанных признаков, целых чисел, стадий, операций, элементов и/или компонентов, но не исключают наличия или добавления одного или нескольких других признаков, целых чисел, стадий, операций, элементов, компонентов и/или их групп.

Применяемый в контексте настоящего документа термин «и/или» включает все без исключения комбинации одного или нескольких связанных перечисленных элементов.

[0046] Если специально не указано или не очевидно из контекста, в контексте настоящего документа термин «приблизительно» в отношении числа или диапазона чисел понимают как обозначающий число и числа +/- 10% от них, или на 10% ниже нижнего указанного предела и на 10% выше верхнего указанного предела для указанных для диапазона значений.

[0047] Если специально не указано, в контексте настоящего документа термин «нуклеиновая кислота» охватывает двухцепочечные или трехцепочечные нуклеиновые кислоты, а также одноцепочечные молекулы. В двухцепочечных или трехцепочечных нуклеиновых кислотах цепи нуклеиновой кислоты не обязательно должны быть параллельными (т. е. двухцепочечная нуклеиновая кислота не должна быть двухцепочечной по всей длине обеих цепей). Последовательности нуклеиновых кислот, если они имеются, представлены в направлении от 5' к 3', если не указано иное. Описанные в настоящем документе способы предусматривают получение выделенных нуклеиновых кислот. Описанные в настоящем документе способы дополнительно предусматривают получение выделенных и очищенных нуклеиновых кислот. «Нуклеиновая кислота» в контексте настоящего документа может составлять в длину по меньшей мере 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000 или более пар. Кроме того, в настоящем документе представлены способы синтеза любого числа нуклеотидных последовательностей, кодирующих полипептидные сегменты, включая последовательности, кодирующие нерибосомные пептиды (NRP), последовательности, кодирующие модули нерибосомный пептид-синтетаза (NRPS) и синтетические варианты, полипептидные сегменты других модульных белков, таких как антитела, полипептидные сегменты из белков других семейств, в том числе некодирующей ДНК или РНК, такой как регуляторные последовательности, например, промоторы, факторы транскрипции, энхансеры, siRNA, shRNA, RNAi, miRNA, малая ядрышковая РНК, полученная из микроРНК, или любая представляющая интерес функциональная или структурная единица ДНК или РНК. Ниже приведены неограничивающие примеры полинуклеотидов: кодирующие или некодирующие участки гена или фрагмента гена, межгенная ДНК, локусы (локус), определенные по результатам анализа связывания, экзоны, интроны, матричная РНК (мРНК), транспортная РНК, рибосомальная РНК, короткая интерферирующая РНК (siRNA), короткая шпилечная РНК (shRNA), микроРНК (miRNA), малая ядрышковая РНК, рибозимы, комплементарная ДНК (сDNA), которая

представляет собой представление ДНК с мРНК, обычно получаемое путем обратной транскрипции матричной РНК (мРНК) или путем амплификации; молекулы ДНК, полученные синтетическим путем или путем амплификации, геномная ДНК, рекомбинантные полинуклеотиды, разветвленные полинуклеотиды, плазмиды, векторы, выделенная ДНК с любой последовательностью, выделенная РНК с любой последовательностью, зонды нуклеиновой кислоты и праймеры. кДНК, кодирующая указанный в настоящем документе ген или фрагмент гена, может содержать по меньшей мере один участок, кодирующий последовательности экзонов без промежуточной последовательности интрона в эквивалентной геномной последовательности.

[0048] Библиотеки GPCR

[0049] Настоящее изобретение относится к способам и композициям, относящимся к библиотекам молекул, связывающихся с сопряженным с G-белком рецептором (GPCR), содержащим нуклеиновые кислоты, кодирующие остов, содержащий связывающий GPCR домен. Описываемые в настоящем документе остовы могут служить стабильной основой для связывающего GPCR домена. Связывающий GPCR домен может быть сконструирован на основе поверхностных взаимодействий лиганда GPCR и GPCR. Описанные в настоящем документе библиотеки могут дополнительно быть мозаичными для создания вариантов библиотек, содержащих нуклеиновые кислоты, каждая из которых кодирует заданный вариант по меньшей мере одной заданной эталонной последовательности нуклеиновой кислоты. Дополнительно в настоящем документе описаны библиотеки белков, которые можно получить при трансляции библиотек нуклеиновых кислот. В соответствии с некоторыми вариантами, описываемые в настоящем документе библиотеки нуклеиновых кислот переносят в клетки для создания клеточной библиотеки. Также в настоящем документе представлены последующие применения для библиотек, синтезированных с помощью описываемых в настоящем документе способов. Последующие применения включают идентификацию вариантов нуклеиновых кислот или белковых последовательностей с улучшенными биологически значимыми функциями, например, улучшенными стабильностью, аффинностью, связыванием, функциональной активностью, а также для лечения или предупреждения связанного с передачей сигналов GPCR болезненного состояния.

[0050] Библиотеки остовов

[0051] Настоящее изобретение относится к библиотекам, содержащим кодирующие остов нуклеиновые кислоты, причем последовательности связывающих GPCR доменов

размещены в остове. Описываемый в настоящем документе остов обеспечивает улучшенную стабильность для ряда последовательностей, кодирующих связывающий GPCR домен, который вставлен в остов, по сравнению с немодифицированным остовом. К иллюстративным остовам относятся без ограничения белок, пептид, иммуноглобулин, их производные или их комбинации. В соответствии с некоторыми вариантами, остов представляет собой иммуноглобулин. Описываемые в настоящем документе остовы характеризуются улучшенной функциональной активностью, структурной стабильностью, экспрессией, специфичностью или их комбинацией. В соответствии с некоторыми вариантами, остовы содержат длинные участки, служащие основой для связывающего GPCR домена.

[0052] Настоящее изобретение относится к библиотекам, содержащим нуклеиновые кислоты, кодирующие остов, причем остов представляет собой иммуноглобулин. В соответствии с некоторыми вариантами, иммуноглобулин представляет собой антитело. Применяемый в контексте настоящего документа термин «антитело» следует понимать как включающий белки, имеющие характерную двухплечевую Y-образную форму типичной молекулы антитела, а также один или несколько фрагментов антитела, которые сохраняют способность специфически связываться с антигеном. К иллюстративным антителам относятся без ограничения моноклональное антитело, поликлональное антитело, биспецифическое антитело, мультиспецифическое антитело, привитое антитело, человеческое антитело, гуманизированное антитело, синтетическое антитело, химерное антитело, камелизированное антитело, одноцепочечные Fv (scFv) (включая фрагменты, в которых VL и VH соединены с помощью рекомбинантных способов с помощью синтетического или природного линкера, который позволяет получить их в виде единой белковой цепи, в которой участки VL и VH соединены с образованием одновалентных молекул, в том числе одноцепочечных Fab и scFab), одноцепочечное антитело, Fab-фрагмент (включая одновалентные фрагменты, содержащие домены VL, VH, CL и CH1), F(ab')₂-фрагмент (включая двухвалентные фрагменты, содержащие два Fab-фрагмента, соединенные дисульфидным мостиком в шарнирном участке), Fd-фрагмент (включая фрагменты, содержащие фрагмент VH и CH1), Fv-фрагмент (включая фрагменты, содержащие домены VL и VH одного плеча антитела), однодоменное антитело (dAb или sdAb) (включая фрагменты, содержащие домен VH), выделенный определяющий комплементарность (CDR) участок, диатело (включая фрагменты, содержащие двухвалентные димеры, такие как два домена VL и VH, связанные друг с другом и распознающие два разных антигена), фрагмент, состоящий только из одного мономерного переменного домена, связанные дисульфидными связями Fv (sdFv), интратело,

антиидиотипическое (анти-Id) антитело или его антигенсвязывающие фрагменты ab. В соответствии с некоторыми вариантами, раскрываемые в настоящем документе библиотеки содержат нуклеиновые кислоты, кодирующие остов, причем остов представляет собой Fv-антитело, включая Fv-антитела, состоящие из минимального фрагмента антитела, который содержит полный антигенраспознающий и антигенсвязывающий сайт. В соответствии с некоторыми вариантами осуществления, Fv-антитело состоит из димера одного переменного домена тяжелой цепи и одного переменного домена легкой цепи в тесной нековалентной связи, и три гиперпеременных участка каждого переменного домена взаимодействуют, определяя антигенсвязывающий сайт на поверхности димера VH-VL. В соответствии с некоторыми вариантами осуществления, шесть гиперпеременных участков придают антителу антигенсвязывающую специфичность. В соответствии с некоторыми вариантами осуществления, один переменный домен (или половина Fv, содержащая только три гиперпеременных участка, специфичных к антигену, включая однодоменные антитела, выделенные из относящихся к верблюжьим животным, содержащие один переменный домен тяжелой цепи, такие как антитела VHH или нанотела) обладает способностью распознавать и связывать антиген. В соответствии с некоторыми вариантами, раскрываемые в настоящем документе библиотеки содержат нуклеиновые кислоты, кодирующие остов, причем остов представляет собой одноцепочечный Fv или scFv, включая фрагменты антител, содержащие домен VH, домен VL или как домен VH, так и домен VL, причем оба домена присутствуют в единой полипептидной цепи. В соответствии с некоторыми вариантами осуществления, полипептид Fv дополнительно содержит полипептидный линкер между доменами VH и VL, позволяющий scFv образовывать структуру, требуемую для связывания антигена. В соответствии с некоторыми вариантами, scFv связан с Fc-фрагментом или VHH связан с Fc-фрагментом (включая минитела). В соответствии с некоторыми вариантами, к антителу относятся молекулы иммуноглобулина и иммунологически активные фрагменты молекул иммуноглобулина, например, молекулы, которые содержат антигенсвязывающий сайт. Молекулы иммуноглобулина относятся к любому типу (например, IgG, IgE, IgM, IgD, IgA и IgY), классу (например, IgG 1, IgG 2, IgG 3, IgG 4, IgA 1 и IgA 2) или подклассу.

[0053] Описываемые в настоящем документе библиотеки, содержащие нуклеиновые кислоты, кодирующие остов, причем остов представляет собой иммуноглобулин, содержат варианты по меньшей мере в одном участке иммуноглобулина. Иллюстративные участки антитела для варианта включают без ограничения определяющий комплементарность участок (CDR), переменный домен или константный домен. В соответствии с некоторыми вариантами, CDR представляет собой CDR1, CDR2 или CDR3. В соответствии с

некоторыми вариантами, CDR представляет собой домен тяжелой цепи, включая без ограничения CDR-H1, CDR-H2 и CDR-H3. В соответствии с некоторыми вариантами, CDR представляет собой домен легкой цепи, включая без ограничения CDR-L1, CDR-L2 и CDR-L3. В соответствии с некоторыми вариантами, переменный домен представляет собой переменный домен легкой цепи (VL) или переменный домен тяжелой цепи (VH). В соответствии с некоторыми вариантами, домен VL содержит каппа или лямбда цепи. В соответствии с некоторыми вариантами, константный домен представляет собой константный домен легкой цепи (CL) или константный домен тяжелой цепи (CH).

[0054] Описываемые в настоящем документе способы предусматривают синтез библиотек, содержащих нуклеиновые кислоты, кодирующие остов, причем каждая нуклеиновая кислота кодирует заданный вариант по меньшей мере одной заданной эталонной последовательности нуклеиновой кислоты. В некоторых случаях заданная эталонная последовательность представляет собой последовательность нуклеиновой кислоты, кодирующую белок, а библиотека вариантов содержит последовательности, кодирующие вариант по меньшей мере отдельного кодона, так чтобы при стандартных процессах трансляции создавалось множество различных вариантов одного остатка в последующем белке, кодируемом синтезированной нуклеиновой кислотой. В соответствии с некоторыми вариантами, библиотека остовов содержит различные нуклеиновые кислоты, совместно кодирующие варианты в нескольких положениях. В соответствии с некоторыми вариантами, библиотека вариантов содержит последовательности, кодирующие вариант по меньшей мере отдельного кодона CDR-H1, CDR-H2, CDR-H3, CDR-L1, CDR-L2, CDR-L3 домена VL или VH. В соответствии с некоторыми вариантами, библиотека вариантов содержит последовательности, кодирующие варианты множества кодонов CDR-H1, CDR-H2, CDR-H3, CDR-L1, CDR-L2, CDR-L3 домена VL или VH. В соответствии с некоторыми вариантами, библиотека вариантов содержит последовательности, кодирующие варианты множества кодонов каркасного элемента 1 (FW1), каркасного элемента 2 (FW2), каркасного элемента 3 (FW3) или каркасного элемента 4 (FW4). Иллюстративное количество кодонов у варианта включает без ограничения по меньшей мере или приблизительно 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300 или более 300 кодонов.

[0055] В соответствии с некоторыми вариантами, по меньшей мере один участок иммуноглобулина у варианта получен из семейства V-генов тяжелой цепи, семейства D-генов тяжелой цепи, семейства J-генов тяжелой цепи, семейства V-генов легкой цепи или семейства J-генов легкой цепи. В соответствии с некоторыми вариантами, семейство V-генов легкой цепи включает ген каппа иммуноглобулина (IGK) или лямбда

иммуноглобулина (IGL). К иллюстративным генам относятся без ограничения IGHV1-18, IGHV1-69, IGHV1-8, IGHV3-21, IGHV3-23, IGHV3-30/33m, IGHV3-28, IGHV1-69, IGHV3-74, IGHV4-39, IGHV4-59/61, IGKV1-39, IGKV1-9, IGKV2-28, IGKV3-11, IGKV3-15, IGKV3-20, IGKV4-1, IGLV1-51 и IGLV2-14.

[0056] Настоящее изобретение относится к библиотекам, содержащим нуклеиновые кислоты, кодирующие остовы иммуноглобулина, причем библиотеки синтезированы с различными количествами фрагментов. В соответствии с некоторыми вариантами, фрагменты включают CDR-H1, CDR-H2, CDR-H3, CDR-L1, CDR-L2, CDR-L3 домена VL или VH. В соответствии с некоторыми вариантами, фрагменты содержат каркасный элемент 1 (FW1), каркасный элемент 2 (FW2), каркасный элемент 3 (FW3) или каркасный элемент 4 (FW4). В соответствии с некоторыми вариантами, библиотеки остовов синтезируют по меньшей мере или приблизительно с 2 фрагментами, с 3 фрагментами, с 4 фрагментами, с 5 фрагментами или более чем с 5 фрагментами. Длина каждого из фрагментов нуклеиновой кислоты или средняя длина синтезированных нуклеиновых кислот может составлять по меньшей мере или приблизительно 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600 или более 600 пар оснований. В соответствии с некоторыми вариантами, длина составляет приблизительно 50-600, 75-575, 100-550, 125-525, 150-500, 175-475, 200-450, 225-425, 250-400, 275-375 или 300-350 пар оснований.

[0057] Библиотеки, содержащие нуклеиновые кислоты, кодирующие остовы иммуноглобулинов, которые описаны в настоящем документе, при трансляции содержат различные по длине цепи аминокислот. В соответствии с некоторыми вариантами, длина каждого из аминокислотных фрагментов или средняя длина синтезированных цепей из аминокислот может составлять по меньшей мере или приблизительно 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150 или более 150 аминокислот. В соответствии с некоторыми вариантами, длина цепи из аминокислот составляет приблизительно 15-150, 20-145, 25-140, 30-135, 35-130, 40-125, 45-120, 50-115, 55-110, 60-110, 65-105, 70-100 или 75-95 аминокислот. В соответствии с некоторыми вариантами, длина цепи из аминокислот составляет от приблизительно 22 аминокислот до приблизительно 75 аминокислот. В соответствии с некоторыми вариантами, остовы иммуноглобулинов содержат по меньшей мере или приблизительно 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000 или более 5000 аминокислот.

[0058] Определенное количество вариантов последовательностей по меньшей мере для одного участка иммуноглобулина на вариант синтезируют de novo с помощью

способов, которые описаны в настоящем документе. В соответствии с некоторыми вариантами, определенное количество вариантов последовательностей синтезируют de novo для CDR-H1, CDR-H2, CDR-H3, CDR-L1, CDR-L2, CDR-L3, VL, VH или их комбинаций. В соответствии с некоторыми вариантами, определенное количество вариантов последовательностей синтезируют de novo для каркасного элемента 1 (FW1), каркасного элемента 2 (FW2), каркасного элемента 3 (FW3) или каркасного элемента 4 (FW4). Определенное количество вариантов последовательностей может составлять по меньшей мере или приблизительно 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500 или более 500 последовательностей. В соответствии с некоторыми вариантами, определенное количество вариантов последовательностей составляет по меньшей мере или приблизительно 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000 или более 8000 последовательностей. В соответствии с некоторыми вариантами, определенное количество вариантов последовательностей составляет приблизительно 10-500, 25-475, 50-450, 75-425, 100-400, 125-375, 150-350, 175-325, 200-300, 225-375, 250-350 или 275-325 последовательностей.

[0059] Варианты последовательностей по меньшей мере для одного участка иммуноглобулина, в соответствии с некоторыми вариантами, отличаются по длине и последовательности. В соответствии с некоторыми вариантами, по меньшей мере один участок, который синтезируют de novo, предназначен для CDR-H1, CDR-H2, CDR-H3, CDR-L1, CDR-L2, CDR-L3, VL, VH или их комбинаций. В соответствии с некоторыми вариантами, по меньшей мере один участок, который синтезируют de novo, предназначен для каркасного элемента 1 (FW1), каркасного элемента 2 (FW2), каркасного элемента 3 (FW3) или каркасного элемента 4 (FW4). В соответствии с некоторыми вариантами, вариант последовательности содержит по меньшей мере или приблизительно 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50 или более 50 вариантов нуклеотидов или аминокислот по сравнению с диким типом. В соответствии с некоторыми вариантами, вариант последовательности содержит по меньшей мере или приблизительно 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45 или 50 дополнительных нуклеотидов или аминокислот по сравнению с диким типом. В соответствии с некоторыми вариантами, вариант последовательности содержит по меньшей мере или приблизительно 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45 или 50 менее нуклеотидов или аминокислот по сравнению с диким типом. В соответствии с некоторыми вариантами, библиотеки содержат по меньшей мере или приблизительно 10^1 , 10^2 , 10^3 , 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} или более 10^{10} вариантов.

[0060] После синтеза библиотек остовов библиотеки остовов можно применять для скрининга и анализа. Например, библиотеки остовов анализируют на способность экспонирования и пэннинг. В соответствии с некоторыми вариантами, способность экспонирования оценивают с помощью селективируемой метки. Иллюстративные метки включают без ограничения радиоактивную метку, флуоресцентную метку, фермент, хемилюминесцентную метку, колориметрическую метку, аффинную метку или другие меченные элементы или метки, которые известны в настоящей области техники. В соответствии с некоторыми вариантами, меткой является гистидин, полигистидин, тус, гемагглютинин (НА) или FLAG. В соответствии с некоторыми вариантами, библиотеки остовов анализируют путем секвенирования с помощью различных способов, включая без ограничения секвенирование в режиме реального времени (SMRT) одиночной молекулы, секвенирование методом молекулярных колоний, секвенирование с помощью лигирования, секвенирование с обратимым терминатором, секвенирование с обнаружением протонов, секвенирование с ионными полупроводниками, секвенирование с помощью нанопор, электронное секвенирование, пиросеквенирование, секвенирование Максама - Гилберта, секвенирование с разрывом цепи (например, Сэнгера), секвенирование +S или секвенирование путем синтеза.

[0061] В соответствии с некоторыми вариантами, библиотеки остовов анализируют на функциональную активность, структурную стабильность (например, термостабильность или стабильность рН), экспрессию, специфичность или их комбинацию. В соответствии с некоторыми вариантами, библиотеки остовов анализируют на остовы, способные к фолдингу. В соответствии с некоторыми вариантами, участок антитела анализируют на функциональную активность, структурную стабильность, экспрессию, специфичность, фолдинг или их комбинацию. Например, участок VH или участок VL анализируют на функциональную активность, структурную стабильность, экспрессию, специфичность, фолдинг или их комбинацию.

[0062] Библиотеки GPCR

[0063] Настоящее изобретение относится к библиотекам молекул, связывающихся с сопряженным с G-белком рецептором (GPCR), содержащим нуклеиновые кислоты, кодирующие остовы, содержащие последовательности связывающих GPCR доменов. В соответствии с некоторыми вариантами, остовы представляют собой иммуноглобулины. В соответствии с некоторыми вариантами, остовы, содержащие последовательности связывающих GPCR доменов, определяют по взаимодействиям связывающих GPCR доменов с GPCR.

[0064] Настоящее изобретение относится к библиотекам, содержащим нуклеиновые кислоты, кодирующие остовы, содержащие связывающие GPCR домены, причем связывающие GPCR домены сконструированы на основе поверхностных взаимодействий на GPCR. Иллюстративные GPCR показаны в **таблице 1**. В соответствии с некоторыми вариантами, связывающие GPCR домены взаимодействуют с N-концом или C-концом GPCR. В соответствии с некоторыми вариантами, связывающие GPCR домены взаимодействуют по меньшей мере с одним трансмембранным доменом, в том числе без ограничения трансмембранным доменом 1 (TM1), трансмембранным доменом 2 (TM2), трансмембранным доменом 3 (TM3), трансмембранным доменом 4 (TM4), трансмембранным доменом 5 (TM5), трансмембранным доменом 6 (TM6) и трансмембранным доменом 7 (TM7). В соответствии с некоторыми вариантами, связывающие GPCR домены взаимодействуют с внутриклеточной поверхностью GPCR. Например, связывающие GPCR домены взаимодействуют по меньшей мере с одной внутриклеточной петлей, в том числе без ограничения внутриклеточной петлей 1 (ICL1), внутриклеточной петлей 2 (ICL2) и внутриклеточной петлей 3 (ICL3). В соответствии с некоторыми вариантами, связывающие GPCR домены взаимодействуют с внеклеточной поверхностью GPCR. См. **фиг. 1**. Например, связывающие GPCR домены взаимодействуют по меньшей мере с одним внеклеточным доменом (ECD) или внеклеточной петлей (ECL) GPCR. Внеклеточные петли включают без ограничения внеклеточную петлю 1 (ECL1), внеклеточную петлю 2 (ECL2) и внеклеточную петлю 3 (ECL3).

Таблица 1. Перечень GPCR

GPCR	Название гена	Регистрационный номер
5-гидрокситриптаминовый рецептор 1A	HTR1A	P08908
5-гидрокситриптаминовый рецептор 1B	HTR1B	P28222
5-гидрокситриптаминовый рецептор 1D	HTR1D	P28221
5-гидрокситриптаминовый рецептор 1E	HTR1E	P28566
5-гидрокситриптаминовый рецептор 1F	HTR1F	P30939
5-гидрокситриптаминовый рецептор 2A	HTR2A	P28223
5-гидрокситриптаминовый рецептор 2B	HTR2B	P41595
5-гидрокситриптаминовый рецептор 2C	HTR2C	P28335
5-гидрокситриптаминовый рецептор 4	HTR4	Q13639
5-гидрокситриптаминовый рецептор 5A	HTR5A	P47898

5-гидрокситриптаминовый рецептор 6	HTR6	P50406
5-гидрокситриптаминовый рецептор	HTR7	P34969
Аденозиновый рецептор A1	ADORA1	P30542
Аденозиновый рецептор A2a	ADORA2A	P29274
Аденозиновый рецептор A2b	ADORA2B	P29275
Аденозиновый рецептор A3	ADORA3	P33765
Мускариновый ацетилхолиновый рецептор M1	CHRM1	P11229
Мускариновый ацетилхолиновый рецептор M2	CHRM2	P08172
Мускариновый ацетилхолиновый рецептор M3	CHRM3	P20309
Мускариновый ацетилхолиновый рецептор M4	CHRM4	P08173
Мускариновый ацетилхолиновый рецептор M5	CHRM5	P08912
Рецептор адренкортикотропного гормона	MC2R	Q01718
α -1A адренергический рецептор	ADRA1A	P35348
α -1B адренергический рецептор	ADRA1B	P35368
α -1D адренергический рецептор	ADRA1D	P25100
α -2A адренергический рецептор	ADRA2A	P08913
α -2B адренергический рецептор	ADRA2B	P18089
α -2C адренергический рецептор	ADRA2C	P18825
β -1 адренергический рецептор	ADRB1	P08588
β -2 адренергический рецептор	ADRB2	P07550
β -3 адренергический рецептор	ADRB3	P13945
Рецептор ангиотензина II 1-го типа	AGTR1	P30556
Рецептор антигена Даффи/хемокиновый рецептор	DARC	Q16570
Рецептор эндотелина 1	EDNRA	P25101
Рецептор эндотелина B	EDNRB	P24530
N-формилпептидный рецептор 2	FPR2	P25090

Рецептор фолликулостимулирующего гормона	FSHR	P23945
Галаниновый рецептор 1-го типа	GALR1	P47211
Галаниновый рецептор 2-го типа	GALR2	O43603
Галаниновый рецептор 3-го типа	GALR3	O60755
Рецептор гастрин/холецистокинина типа В	CCKBR	P32239
Рецептор гонадотропин-высвобождающего гормона	GNRHR	P30968
Предполагаемый рецептор гонадотропин-высвобождающего гормона II	GNRHR2	Q96P88
Сопряженный с G-белком рецептор эстрогена 1	GPER	Q99527
Рецептор урациловых нуклеотидов/цистеинил-лейкотриена	GPR17	Q13304
Предполагаемый сопряженный с G-белком рецептор 44	GPR44	Q9Y5Y4
Сопряженный с G-белком рецептор 55	GPR55	Q9Y2T6
Рецептор гастрин-высвобождающего пептида	GRPR	P30550
H1-гистаминовый рецептор	HRH1	P35367
H2-гистаминовый рецептор	HRH2	P25021
H3-гистаминовый рецептор	HRH3	Q9Y5N1
H4-гистаминовый рецептор	HRH4	Q9H3N8
Рецептор KiSS-1	KISS1R	Q969F8
Рецептор 1 лизофосфатидной кислоты	LPAR1	Q92633
Рецептор 2 лизофосфатидной кислоты	LPAR2	Q9HBW0
Рецептор 3 лизофосфатидной кислоты	LPAR3	Q9UBY5
Рецептор 4 лизофосфатидной кислоты	LPAR4	Q99677
Рецептор 6 лизофосфатидной кислоты	LPAR6	P43657
Рецептор лютропина-хориогонадотропина	LHCGR	P22888
Рецептор 1 лейкотриена B4	LTB4R	Q15722
Рецептор 2 лейкотриена B4	LTB4R2	Q9NPC1
Рецептор 3 меланокортина	MC3R	P41968
Рецептор 4 меланокортина	MC4R	P32245

Рецептор 5 меланокортина	MC5R	P33032
Обонятельный рецептор 10G9	OR10G9	Q8NGN4
Обонятельный рецептор 10H1	OR10H1	Q9Y4A9
Обонятельный рецептор 10H2	OR10H2	O60403
Обонятельный рецептор 10H3	OR10H3	O60404
Обонятельный рецептор 10H4	OR10H4	Q8NGA5
Обонятельный рецептор 10H5	OR10H5	Q8NGA6
Обонятельный рецептор 10J1	OR10J1	P30954
Обонятельный рецептор 10J3	OR10J3	Q5JRS4
Обонятельный рецептор 10J5	OR10J5	Q8NHC4
Обонятельный рецептор 10K1	OR10K1	Q8NGX5
Обонятельный рецептор 10K2	OR10K2	Q6IF99
Обонятельный рецептор 10P1	OR10P1	Q8NGE3
Обонятельный рецептор 10Q1	OR10Q1	Q8NGQ4
Обонятельный рецептор 10R2	OR10R2	Q8NGX6
Обонятельный рецептор 10S1	OR10S1	Q8NGN2
Обонятельный рецептор 10T2	OR10T2	Q8NGX3
Обонятельный рецептор 10V1	OR10V1	Q8NGI7
Обонятельный рецептор 10W1	OR10W1	Q8NGF6
Обонятельный рецептор 14A2	OR14A2	Q96R54
Обонятельный рецептор 14C36	OR14C36	Q8NHC7
Обонятельный рецептор 14I1	OR14I1	A6ND48
Обонятельный рецептор 14J1	OR14J1	Q9UGF5
Обонятельный рецептор 14K1	OR14K1	Q8NGZ2
Обонятельный рецептор 2A12	OR2A12	Q8NGT7
Обонятельный рецептор 2A14	OR2A14	Q96R47
Обонятельный рецептор 2A25	OR2A25	A4D2G3
Обонятельный рецептор 2AG1	OR2AG1	Q9H205
Обонятельный рецептор 2AG2	OR2AG2	A6NM03
Обонятельный рецептор 2AJ1	OR2AJ1	Q8NGZ0
Обонятельный рецептор 2AK2	OR2AK2	Q8NG84
Обонятельный рецептор 2AP1	OR2AP1	Q8NGE2
Обонятельный рецептор 2AT4	OR2AT4	A6NND4

Обонятельный рецептор 51I2	OR51I2	Q9H344
Обонятельный рецептор 51J1	OR51J1	Q9H342
Обонятельный рецептор 51L1	OR51L1	Q8NGJ5
Обонятельный рецептор 51M1	OR51M1	Q9H341
Обонятельный рецептор 51Q1	OR51Q1	Q8NH59
Обонятельный рецептор 51S1	OR51S1	Q8NGJ8
Обонятельный рецептор 51T1	OR51T1	Q8NGJ9
Обонятельный рецептор 51V1	OR51V1	Q9H2C8
Обонятельный рецептор 52A1	OR52A1	Q9UKL2
Обонятельный рецептор 52A5	OR52A5	Q9H2C5
Обонятельный рецептор 52B2	OR52B2	Q96RD2
Обонятельный рецептор 52B4	OR52B4	Q8NGK2
Обонятельный рецептор 52B6	OR52B6	Q8NGF0
Обонятельный рецептор 52D1	OR52D1	Q9H346
Обонятельный рецептор 52E2	OR52E2	Q8NGJ4
Обонятельный рецептор 52E4	OR52E4	Q8NGH9
Обонятельный рецептор 52E5	OR52E5	Q8NH55
Обонятельный рецептор 52E6	OR52E6	Q96RD3
Обонятельный рецептор 52E8	OR52E8	Q6IFG1
Обонятельный рецептор 52H1	OR52H1	Q8NGJ2
Обонятельный рецептор 52I1	OR52I1	Q8NGK6
Обонятельный рецептор 52I2	OR52I2	Q8NH67
Обонятельный рецептор 52K1	OR52K1	Q8NGK4
Обонятельный рецептор 52K2	OR52K2	Q8NGK3
Обонятельный рецептор 52L1	OR52L1	Q8NGH7
Обонятельный рецептор 52M1	OR52M1	Q8NGK5
Обонятельный рецептор 52N1	OR52N1	Q8NH53
Обонятельный рецептор 52N2	OR52N2	Q8NGI0
Обонятельный рецептор 52N4	OR52N4	Q8NGI2
Обонятельный рецептор 52N5	OR52N5	Q8NH56
Обонятельный рецептор 52R1	OR52R1	Q8NGF1
Обонятельный рецептор 52W1	OR52W1	Q6IF63
Чувствительный к красному опсин	OPN1LW	P04000

Подобный визуальному пигменту рецептор перопсин	RRH	O14718
Обонятельный рецептор 1A1	OR1A1	Q9P1Q5
Обонятельный рецептор 1A2	OR1A2	Q9Y585
Обонятельный рецептор 1B1	OR1B1	Q8NGR6
Обонятельный рецептор 1C1	OR1C1	Q15619
Обонятельный рецептор 1D2	OR1D2	P34982
Обонятельный рецептор 1F1	OR1F1	O43749
Обонятельный рецептор 1F12	OR1F12	Q8NHA8
Обонятельный рецептор 1G1	OR1G1	P47890
Обонятельный рецептор 1H	OR1H	O60431
Обонятельный рецептор 1J1	OR1J1	Q8NGS3
Обонятельный рецептор 1J2	OR1J2	Q8NGS2
Обонятельный рецептор 1J4	OR1J4	Q8NGS1
Обонятельный рецептор 1K1	OR1K1	Q8NGR3
Обонятельный рецептор 1L1	OR1L1	Q8NH94
Обонятельный рецептор 1L3	OR1L3	Q8NH93
Обонятельный рецептор 1L4	OR1L4	Q8NGR5
Обонятельный рецептор 1L6	OR1L6	Q8NGR2
Обонятельный рецептор 1L8	OR1L8	Q8NGR8
Обонятельный рецептор 1M1	OR1M1	Q8NGA1
Обонятельный рецептор 1N1	OR1N1	Q8NGS0
Обонятельный рецептор 1N2	OR1N2	Q8NGR9
Обонятельный рецептор 1Q1	OR1Q1	Q15612
Обонятельный рецептор 1S1	OR1S1	Q8NH92
Обонятельный рецептор 1S2	OR1S2	Q8NGQ3
Обонятельный рецептор 2A2	OR2A2	Q6IF42
Обонятельный рецептор 2A4	OR2A4	O95047
Обонятельный рецептор 2B2	OR2B2	Q9GZK3
Предполагаемый обонятельный рецептор 2B3	OR2B3	O76000
Обонятельный рецептор 2B6	OR2B6	P58173
Предполагаемый обонятельный рецептор 2B8	OR2B8P	P59922
Обонятельный рецептор 2T5	OR2T5	Q6IEZ7

Обонятельный рецептор 2Т6	OR2Т6	Q8NHC8
Обонятельный рецептор 2Т8	OR2Т8	A6NH00
Обонятельный рецептор 2V1	OR2V1	Q8NHB1
Обонятельный рецептор 2V2	OR2V2	Q96R30
Обонятельный рецептор 2W1	OR2W1	Q9Y3N9
Обонятельный рецептор 2W3	OR2W3	Q7Z3T1
Обонятельный рецептор 2Y1	OR2Y1	Q8NGV0
Обонятельный рецептор 2Z1	OR2Z1	Q8NG97
Обонятельный рецептор 3А1	OR3А1	P47881
Обонятельный рецептор 3А2	OR3А2	P47893
Обонятельный рецептор 3А3	OR3А3	P47888
Обонятельный рецептор 3А4	OR3А4	P47883
Предполагаемый обонятельный рецептор 4А4	OR4А4P	Q8NGN8
Обонятельный рецептор 4А5	OR4А5	Q8NH83
Обонятельный рецептор 4А8	OR4А8P	P0C604
Обонятельный рецептор 4В1	OR4В1	Q8NGF8
Обонятельный рецептор 4С3	OR4С3	Q8NH37
Обонятельный рецептор 4С5	OR4С5	Q8NGB2
Обонятельный рецептор 4С6	OR4С6	Q8NH72
Обонятельный рецептор 4С11	OR4С11	Q6IEV9
Обонятельный рецептор 4С12	OR4С12	Q96R67
Обонятельный рецептор 4С13	OR4С13	Q8NGP0
Обонятельный рецептор 4С15	OR4С15	Q8NGM1
Обонятельный рецептор 4С16	OR4С16	Q8NGL9
Обонятельный рецептор 4D1	OR4D1	Q15615
Обонятельный рецептор 4D2	OR4D2	P58180
Обонятельный рецептор 4D5	OR4D5	Q8NGN0
Обонятельный рецептор 4D6	OR4D6	Q8NGJ1
Обонятельный рецептор 4D9	OR4D9	Q8NGE8
Обонятельный рецептор 4D10	OR4D10	Q8NGI6
Обонятельный рецептор 4D11	OR4D11	Q8NGI4
Обонятельный рецептор 5В17	OR5В17	Q8NGF7
Обонятельный рецептор 5В21	OR5В21	A6NL26

Обонятельный рецептор 5C1	OR5C1	Q8NGR4
Обонятельный рецептор 5D13	OR5D13	Q8NGL4
Обонятельный рецептор 5D14	OR5D14	Q8NGL3
Обонятельный рецептор 5D16	OR5D16	Q8NGK9
Обонятельный рецептор 5D18	OR5D18	Q8NGL1
Обонятельный рецептор 5F1	OR5F1	O95221
Обонятельный рецептор 5H1	OR5H1	A6NKK0
Обонятельный рецептор 5H2	OR5H2	Q8NGV7
Обонятельный рецептор 5H6	OR5H6	Q8NGV6
Обонятельный рецептор 5I1	OR5I1	Q13606
Обонятельный рецептор 5J2	OR5J2	Q8NH18
Обонятельный рецептор 5K1	OR5K1	Q8NHB7
Обонятельный рецептор 5K2	OR5K2	Q8NHB8
Обонятельный рецептор 5K3	OR5K3	A6NET4
Обонятельный рецептор 5K4	OR5K4	A6NMS3
Обонятельный рецептор 5L1	OR5L1	Q8NGL2
Обонятельный рецептор 5L2	OR5L2	Q8NGL0
Обонятельный рецептор 5M1	OR5M1	Q8NGP8
Обонятельный рецептор 5M3	OR5M3	Q8NGP4
Обонятельный рецептор 5M8	OR5M8	Q8NGP6
Обонятельный рецептор 5M9	OR5M9	Q8NGP3
Обонятельный рецептор 5M10	OR5M10	Q6IEU7
Обонятельный рецептор 5M11	OR5M11	Q96RB7
Обонятельный рецептор 5P2	OR5P2	Q8WZ92
Обонятельный рецептор 5P3	OR5P3	Q8WZ94
Обонятельный рецептор 5R1	OR5R1	Q8NH85
Обонятельный рецептор 5T1	OR5T1	Q8NG75
Обонятельный рецептор 5T2	OR5T2	Q8NGG2
Обонятельный рецептор 5T3	OR5T3	Q8NGG3
Обонятельный рецептор 5W2	OR5W2	Q8NH69
Обонятельный рецептор 7G3	OR7G3	Q8NG95
Обонятельный рецептор 8A1	OR8A1	Q8NGG7
Обонятельный рецептор 8B3	OR8B3	Q8NGG8

Обонятельный рецептор 8B4	OR8B4	Q96RC9
Обонятельный рецептор 8B8	OR8B8	Q15620
Обонятельный рецептор 8B12	OR8B12	Q8NGG6
Обонятельный рецептор 8D1	OR8D1	Q8WZ84
Обонятельный рецептор 8D2	OR8D2	Q9GZM6
Обонятельный рецептор 8D4	OR8D4	Q8NGM9
Орексиновый рецептор 1-го типа	HCRTR1	O43613
Орексиновый рецептор 2-го типа	HCRTR2	O43614
Оксозйкозанный рецептор 1	OXER1	Q8TDS5
Окситоциновый рецептор	OXTR	P30559
P2Y пуринергический рецептор 1	P2RY1	P47900
P2Y пуринергический рецептор 2	P2RY2	P41231
P2Y пуринергический рецептор 4	P2RY4	P51582
P2Y пуринергический рецептор 6	P2RY6	Q15077
P2Y пуринергический рецептор 8	P2RY8	Q86VZ1
Предполагаемый P2Y пуринергический рецептор 10	P2RY10	O00398
P2Y пуринергический рецептор 11	P2RY11	Q96G91
P2Y пуринергический рецептор 12	P2RY12	Q9H244
P2Y пуринергический рецептор 13	P2RY13	Q9BPV8
P2Y пуринергический рецептор 14	P2RY14	Q15391
Активируемый протеиназой рецептор 1	F2R	P25116
Активируемый протеиназой рецептор 2	F2RL1	P55085
Активируемый протеиназой рецептор 3	F2RL2	O00254
Активируемый протеиназой рецептор 4	F2RL3	Q96RI0
Рецептор простагландина D2	PTGDR	Q13258
Рецептор простагландина E2 подтипа EP1	PTGER1	P34995
Рецептор простагландина E2 подтипа EP2	PTGER2	P43116
Рецептор простагландина E2 подтипа EP3	PTGER3	P43115
Рецептор простагландина E2 подтипа EP4	PTGER4	P35408
Рецептор ангиотензина II 2-го типа (AT2)	AGTR2	P50052
Апелиновый рецептор	APLNR	P35414
Рецептор брадикинина B1	BDKRB1	P46663

Рецептор брадикинина B2	BDKRB2	P30411
Анафилатаксиновый хемотактический рецептор C5a	C5AR1	P21730
Холецистокининовый рецептор типа А	CCKAR	P32238
С-С хемокиновый рецептор 10	CCR10	P46092
С-С хемокиновый рецептор 1	CCR1	P32246
С-С хемокиновый рецептор 2	CCR2	P41597
С-С хемокиновый рецептор 3	CCR3	P51677
С-С хемокиновый рецептор 4	CCR4	P51679
С-С хемокиновый рецептор 5	CCR5	P51681
С-С хемокиновый рецептор 6	CCR6	P51684
С-С хемокиновый рецептор 7	CCR7	P32248
С-С хемокиновый рецептор 8	CCR8	P51685
С-С хемокиновый рецептор 9	CCR9	P51686
Цистеинил-лейкотриеновый рецептор 1	CYSLTR1	Q9Y271
Цистеинил-лейкотриеновый рецептор 2	CYSLTR2	Q9NS75
Каннабиноидный рецептор 1-го типа	CNR1	P21554
Каннабиноидный рецептор 2-го типа	CNR2	P34972
СХ3С хемокиновый рецептор 1	CX3CR1	P49238
Высокоаффинный рецептор А интерлейкина-8	IL8RA	P25024
Высокоаффинный рецептор В интерлейкина-8	IL8RB	P25025
С-Х-С хемокиновый рецептор 3	CXCR3	P49682
С-Х-С хемокиновый рецептор 4	CXCR4	P61073
С-Х-С хемокиновый рецептор 6	CXCR6	O00574
С-Х-С хемокиновый рецептор 7	CXCR7	P25106
Дофаминовый рецептор D(1A)	DRD1	P21728
Дофаминовый рецептор D(2)	DRD2	P14416
Дофаминовый рецептор D(3)	DRD3	P35462
Дофаминовый рецептор D(4)	DRD4	P21917
Дофаминовый рецептор D(1B)	DRD5	P21918

Рецептор мелатонинстимулирующего гормона	MC1R	Q01726
Мелатониновый рецептор типа 1A	MTNR1A	P48039
Мелатониновый рецептор типа 1B	MTNR1B	P49286
Рецептор вещества Р	TACR1	P25103
Рецептор вещества К	TACR2	P21452
Рецептор нейромедина К	TACR3	P29371
Рецептор нейромедина В	NMBR	P28336
Рецептор нейропептидов В/W 1-го типа	NPBWR1	P48145
Рецептор нейропептидов В/W 2-го типа	NPBWR2	P48146
Рецептор нейропептида FF 1	NPFFR1	Q9GZQ6
Рецептор нейропептида FF 2	NPFFR2	Q9Y5X5
Рецептор нейропептида Y 1-го типа	NPY1R	P25929
Рецептор нейропептида Y 2-го типа	NPY2R	P49146
Рецептор нейропептида Y 4-го типа	PPYR1	P50391
Рецептор нейропептида Y 5-го типа	NPY5R	Q15761
Нейротензиновый рецептор 1-го типа	NTSR1	P30989
Нейротензиновый рецептор 2-го типа	NTSR2	O95665
Обонятельный рецептор 10A2	OR10A2	Q9H208
Обонятельный рецептор 10A3	OR10A3	P58181
Обонятельный рецептор 10A4	OR10A4	Q9H209
Обонятельный рецептор 10A5	OR10A5	Q9H207
Обонятельный рецептор 10A6	OR10A6	Q8NH74
Обонятельный рецептор 10A7	OR10A7	Q8NGE5
Обонятельный рецептор 10AD1	OR10AD1	Q8NGE0
Обонятельный рецептор 10AG1	OR10AG1	Q8NH19
Обонятельный рецептор 10C1	OR10C1	Q96KK4
Обонятельный рецептор 10G2	OR10G2	Q8NGC3
Обонятельный рецептор 10G3	OR10G3	Q8NGC4
Обонятельный рецептор 10G4	OR10G4	Q8NGN3
Обонятельный рецептор 10G6	OR10G6	Q8NH81
Обонятельный рецептор 10G7	OR10G7	Q8NGN6
Обонятельный рецептор 10G8	OR10G8	Q8NGN5

Обонятельный рецептор 2T10	OR2T10	Q8NGZ9
Обонятельный рецептор 2T11	OR2T11	Q8NH01
Обонятельный рецептор 2T12	OR2T12	Q8NG77
Обонятельный рецептор 2T27	OR2T27	Q8NH04
Обонятельный рецептор 2T29	OR2T29	Q8NH02
Обонятельный рецептор 2T33	OR2T33	Q8NG76
Обонятельный рецептор 2T34	OR2T34	Q8NGX1
Обонятельный рецептор 2T35	OR2T35	Q8NGX2
Обонятельный рецептор 4A15	OR4A15	Q8NGL6
Обонятельный рецептор 4A16	OR4A16	Q8NH70
Обонятельный рецептор 4A47	OR4A47	Q6IF82
Обонятельный рецептор 4C45	OR4C45	A6NMZ5
Обонятельный рецептор 4C46	OR4C46	A6NHA9
Обонятельный рецептор 4F15	OR4F15	Q8NGB8
Обонятельный рецептор 4F17	OR4F17	Q8NGA8
Обонятельный рецептор 4F21	OR4F21	O95013
Обонятельный рецептор 51A2	OR51A2	Q8NGJ7
Обонятельный рецептор 51A4	OR51A4	Q8NGJ6
Обонятельный рецептор 51A7	OR51A7	Q8NH64
Обонятельный рецептор 51B2	OR51B2	Q9Y5P1
Обонятельный рецептор 51B4	OR51B4	Q9Y5P0
Обонятельный рецептор 51B5	OR51B5	Q9H339
Обонятельный рецептор 51B5	OR51B6	Q9H340
Обонятельный рецептор 51D1	OR51D1	Q8NGF3
Обонятельный рецептор 51E1	OR51E1	Q8TCB6
Обонятельный рецептор 51E2	OR51E2	Q9H255
Обонятельный рецептор 51F1	OR51F1	A6NGY5
Обонятельный рецептор 51F2	OR51F2	Q8NH61
Обонятельный рецептор 51G1	OR51G1	Q8NGK1
Обонятельный рецептор 51G2	OR51G2	Q8NGK0
Предполагаемый обонятельный рецептор 51H1	OR51H1P	Q8NH63
Обонятельный рецептор 51I1	OR51I1	Q9H343

Обонятельный рецептор 56A1	OR56A1	Q8NGH5
Обонятельный рецептор 56A3	OR56A3	Q8NH54
Обонятельный рецептор 56A4	OR56A4	Q8NGH8
Обонятельный рецептор 56A5	OR56A5	P0C7T3
Обонятельный рецептор 56B1	OR56B1	Q8NGI3
Обонятельный рецептор 56B4	OR56B4	Q8NH76
Обонятельный рецептор 5AC2	OR5AC2	Q9NZP5
Обонятельный рецептор 5AK2	OR5AK2	Q8NH90
Обонятельный рецептор 5AN1	OR5AN1	Q8NGI8
Обонятельный рецептор 5AP2	OR5AP2	Q8NGF4
Обонятельный рецептор 5AR1	OR5AR1	Q8NGP9
Обонятельный рецептор 5AS1	OR5AS1	Q8N127
Обонятельный рецептор 5AU1	OR5AU1	Q8NGC0
Обонятельный рецептор 5H14	OR5H14	A6NHG9
Обонятельный рецептор 5H15	OR5H15	A6NDH6
Обонятельный рецептор 6C65	OR6C65	A6NJZ3
Обонятельный рецептор 6C68	OR6C68	A6NDL8
Обонятельный рецептор 6C70	OR6C70	A6NIJ9
Обонятельный рецептор 6C74	OR6C74	A6NCV1
Обонятельный рецептор 6C75	OR6C75	A6NL08
Обонятельный рецептор 6C76	OR6C76	A6NM76
Обонятельный рецептор 7E24	OR7E24	Q6IFN5
Опсин-3	OPN3	Q9H1Y3
Меланопсин	OPN4	Q9UHM6
Опсин-5	OPN5	Q6U736
Опиоидный рецептор δ -типа	OPRD1	P41143
Опиоидный рецептор κ -типа	OPRK1	P41145
Опиоидный рецептор μ -типа	OPRM1	P35372
Ноцептиновый рецептор	OPRL1	P41146
Чувствительный к синему опсин	OPN1SW	P03999
Родопсин	RHO	P08100
Чувствительный к зеленому опсин	OPN1MW	P04001
Обонятельный рецептор 2B11	OR2B11	Q5JQS5

Обонятельный рецептор 2C1	OR2C1	O95371
Обонятельный рецептор 2C3	OR2C3	Q8N628
Обонятельный рецептор 2D2	OR2D2	Q9H210
Обонятельный рецептор 2D3	OR2D3	Q8NGH3
Обонятельный рецептор 2F1	OR2F1	Q13607
Обонятельный рецептор 2F2	OR2F2	O95006
Обонятельный рецептор 2G2	OR2G2	Q8NGZ5
Обонятельный рецептор 2G3	OR2G3	Q8NGZ4
Обонятельный рецептор 2G6	OR2G6	Q5TZ20
Обонятельный рецептор 2H1	OR2H1	Q9GZK4
Обонятельный рецептор 2H2	OR2H2	O95918
Предполагаемый обонятельный рецептор 2I1	OR2I1P	Q8NGU4
Обонятельный рецептор 2J1	OR2J1	Q9GZK6
Обонятельный рецептор 2J2	OR2J2	O76002
Обонятельный рецептор 2J3	OR2J3	O76001
Обонятельный рецептор 2K2	OR2K2	Q8NGT1
Обонятельный рецептор 2L2	OR2L2	Q8NH16
Обонятельный рецептор 2L3	OR2L3	Q8NG85
Обонятельный рецептор 2L5	OR2L5	Q8NG80
Обонятельный рецептор 2L8	OR2L8	Q8NGY9
Обонятельный рецептор 2L13	OR2L13	Q8N349
Обонятельный рецептор 2M2	OR2M2	Q96R28
Обонятельный рецептор 2M3	OR2M3	Q8NG83
Обонятельный рецептор 2M4	OR2M4	Q96R27
Обонятельный рецептор 2M5	OR2M5	A3KFT3
Обонятельный рецептор 2M7	OR2M7	Q8NG81
Обонятельный рецептор 2S2	OR2S2	Q9NQN1
Обонятельный рецептор 2T1	OR2T1	O43869
Обонятельный рецептор 2T2	OR2T2	Q6IF00
Обонятельный рецептор 2T3	OR2T3	Q8NH03
Обонятельный рецептор 2T4	OR2T4	Q8NH00
Обонятельный рецептор 4E1	OR4E1	P0C645
Обонятельный рецептор 4E2	OR4E2	Q8NGC2

Обонятельный рецептор 4F3/4F16/4F29	OR4F3	Q6IEY1
Обонятельный рецептор 4F4	OR4F4	Q96R69
Обонятельный рецептор 4F5	OR4F5	Q8NH21
Обонятельный рецептор 4F6	OR4F6	Q8NGB9
Обонятельный рецептор 4K1	OR4K1	Q8NGD4
Обонятельный рецептор 4K2	OR4K2	Q8NGD2
Обонятельный рецептор 4K3	OR4K3	Q96R72
Обонятельный рецептор 4K5	OR4K5	Q8NGD3
Обонятельный рецептор 4K13	OR4K13	Q8NH42
Обонятельный рецептор 4K14	OR4K14	Q8NGD5
Обонятельный рецептор 4K15	OR4K15	Q8NH41
Обонятельный рецептор 4K17	OR4K17	Q8NGC6
Обонятельный рецептор 4L1	OR4L1	Q8NH43
Обонятельный рецептор 4M1	OR4M1	Q8NGD0
Обонятельный рецептор 4M2	OR4M2	Q8NGB6
Обонятельный рецептор 4N2	OR4N2	Q8NGD1
Обонятельный рецептор 4N4	OR4N4	Q8N0Y3
Обонятельный рецептор 4N5	OR4N5	Q8IXE1
Обонятельный рецептор 4P4	OR4P4	Q8NGL7
Обонятельный рецептор 4Q2	OR4Q2	P0C623
Обонятельный рецептор 4Q3	OR4Q3	Q8NH05
Обонятельный рецептор 4S1	OR4S1	Q8NGB4
Обонятельный рецептор 4S2	OR4S2	Q8NH73
Обонятельный рецептор 4X1	OR4X1	Q8NH49
Обонятельный рецептор 4X2	OR4X2	Q8NGF9
Обонятельный рецептор 5A1	OR5A1	Q8NGJ0
Обонятельный рецептор 5A2	OR5A2	Q8NGI9
Обонятельный рецептор 5B2	OR5B2	Q96R09
Обонятельный рецептор 5B3	OR5B3	Q8NH48
Обонятельный рецептор 5B12	OR5B12	Q96R08
Обонятельный рецептор 6A2	OR6A2	O95222
Обонятельный рецептор 6B1	OR6B1	O95007
Обонятельный рецептор 6B2	OR6B2	Q6IFH4

Обонятельный рецептор 6B3	OR6B3	Q8NGW1
Обонятельный рецептор 6C1	OR6C1	Q96RD1
Обонятельный рецептор 6C2	OR6C2	Q9NZP2
Обонятельный рецептор 6C3	OR6C3	Q9NZP0
Обонятельный рецептор 6C4	OR6C4	Q8NGE1
Обонятельный рецептор 6C6	OR6C6	A6NF89
Обонятельный рецептор 6F1	OR6F1	Q8NGZ6
Обонятельный рецептор 6J1	OR6J1	Q8NGC5
Обонятельный рецептор 6K2	OR6K2	Q8NGY2
Обонятельный рецептор 6K3	OR6K3	Q8NGY3
Обонятельный рецептор 6K6	OR6K6	Q8NGW6
Обонятельный рецептор 6M1	OR6M1	Q8NGM8
Обонятельный рецептор 6N1	OR6N1	Q8NGY5
Обонятельный рецептор 6N2	OR6N2	Q8NGY6
Обонятельный рецептор 6P1	OR6P1	Q8NGX9
Обонятельный рецептор 6Q1	OR6Q1	Q8NGQ2
Обонятельный рецептор 6S1	OR6S1	Q8NH40
Обонятельный рецептор 6T1	OR6T1	Q8NGN1
Обонятельный рецептор 6V1	OR6V1	Q8N148
Обонятельный рецептор 6X1	OR6X1	Q8NH79
Обонятельный рецептор 6Y1	OR6Y1	Q8NGX8
Обонятельный рецептор 7A5	OR7A5	Q15622
Обонятельный рецептор 7A10	OR7A10	O76100
Обонятельный рецептор 7A17	OR7A17	O14581
Обонятельный рецептор 7C1	OR7C1	O76099
Обонятельный рецептор 7C2	OR7C2	O60412
Обонятельный рецептор 7D4	OR7D4	Q8NG98
Обонятельный рецептор 7G1	OR7G1	Q8NGA0
Обонятельный рецептор 7G2	OR7G2	Q8NG99
Рецептор простагландина F2- α	PTGFR	P43088
Простациклиновый рецептор	PTGIR	P43119
Рецептор пролактин-высвобождающего пептида	PRLHR	P49683

Рецептор активирующего тромбоциты фактора	PTAFR	P25105
Рецептор пироглутамилированного RFамидного пептида	QRFPR	Q96P65
Сопряженный с G-белком RPE-ретикулярный рецептор	RGR	P47804
Сфингозин-1-фосфатный рецептор 1	S1PR1	P21453
Сфингозин-1-фосфатный рецептор 2	S1PR2	O95136
Сфингозин-1-фосфатный рецептор 3	S1PR3	Q99500
Сфингозин-1-фосфатный рецептор 4	S1PR4	O95977
Сфингозин-1-фосфатный рецептор 5	S1PR5	Q9H228
Соматостатиновый рецептор 1-го типа	SSTR1	P30872
Соматостатиновый рецептор 2-го типа	SSTR2	P30874
Соматостатиновый рецептор 3-го типа	SSTR3	P32745
Соматостатиновый рецептор 4-го типа	SSTR4	P31391
Соматостатиновый рецептор 5-го типа	SSTR5	P35346
Рецептор тромбоксана A2	TBXA2R	P21731
Рецептор 1 следовых аминов	TAAR1	Q96RJ0
Рецептор 2 следовых аминов	TAAR2	Q9P1P5
Предполагаемый рецептор 3 следовых аминов	TAAR3	Q9P1P4
Рецептор 5 следовых аминов	TAAR5	O14804
Рецептор 6 следовых аминов	TAAR6	Q96RI8
Рецептор 8 следовых аминов	TAAR8	Q969N4
Рецептор 9 следовых аминов	TAAR9	Q96RI9
Рецептор тиреотропного гормона	TSHR	P16473
Рецептор вазопрессина V1a	AVPR1A	P37288
Рецептор вазопрессина V1b	AVPR1B	P47901
Рецептор вазопрессина V2	AVPR2	P30518
Рецептор 1 хемокина XC	XCR1	P46094
Специфичный для головного мозга ингибитор 1 ангиогенеза	BAI1	O14514

Специфичный для головного мозга ингибитор 2 ангиогенеза	BAI2	O60241
Специфичный для головного мозга ингибитор 3 ангиогенеза	BAI3	O60242
Кальцитониновый рецептор	CALCR	P30988
Рецептор 1-го типа кальцитонин-ген-родственного пептида	CALCRL	Q16602
Рецептор 1 кортикотропин-высвобождающего фактора	CRHR1	P34998
Рецептор 2 кортикотропин-высвобождающего фактора	CRHR2	Q13324
Рецептор соматотропин-высвобождающего гормона	GHRHR	Q02643
Рецептор желудочного ингибиторного полипептида	GIPR	P48546
Рецептор глюкагоноподобного пептида-1	GLP1R	P43220
Рецептор глюкагоноподобного пептида-2	GLP2R	O95838
Глюкагоновый рецептор	GCGR	P47871
Рецептор активирующего гипофизную аденилатциклазу полипептида I типа	ADCYAP1R1	P41586
Представитель 2 семейства вкусовых рецепторов 1-го типа	TAS1R2	Q8TE23
Рецептор 1 паратиреоидного гормона	PTH1R	Q03431
Рецептор паратиреоидного гормона 2	PTH2R	P49190
Рецептор секретина	SCTR	P47872
Рецептор вазоактивного пептида кишечника 1	VIPR1	P32241
Рецептор вазоактивного пептида кишечника 2	VIPR2	P41587
Frizzled-10	FZD10	Q9ULW2
Frizzled-1	FZD1	Q9UP38
Frizzled-2	FZD2	Q14332
Frizzled-3	FZD3	Q9NPG1
Frizzled-4	FZD4	Q9ULV1
Frizzled-5	FZD5	Q13467

Frizzled-6	FZD6	O60353
Frizzled-7	FZD7	O75084
Frizzled-8	FZD8	Q9H461
Frizzled-9	FZD9 (FZD3)	O00144
Гомолог Smoothened	SMO (SMOH)	Q99835
Внеклеточный кальцийчувствительный рецептор	CASR	P41180
GABA-рецептор типа В, субъединица 1	GABBR1	Q9UBS5
GABA-рецептор типа В, субъединица 2	GABBR2	O75899
Представитель А группы 6 семейства С GPCR	GPRC6A	Q5T6X5
Метаботропный рецептор 1 глутамата	GRM1	Q13255
Метаботропный рецептор 2 глутамата	GRM2	Q14416
Метаботропный рецептор 3 глутамата	GRM3	Q14832
Метаботропный рецептор 4 глутамата	GRM4	Q14833
Метаботропный рецептор 5 глутамата	GRM5	P41594
Метаботропный рецептор 6 глутамата	GRM6	O15303
Метаботропный рецептор 7 глутамата	GRM7	Q14831
Метаботропный рецептор 8 глутамата	GRM8	O00222
Представитель 1 семейства вкусовых рецепторов 1-го типа	TAS1R1	Q7RTX1
Представитель 2 семейства вкусовых рецепторов 1-го типа	TAS1R2	Q8TE23
Представитель 3 семейства вкусовых рецепторов 1-го типа	TAS1R3	Q7RTX0

[0065] В настоящем документе описаны связывающие GPCR домены, причем связывающие GPCR домены сконструированы на основе поверхностных взаимодействий лиганда GPCR с GPCR. В соответствии с некоторыми вариантами, лиганд представляет собой субатомную частицу (например, фотон), ион, органическую молекулу, пептид и белок. Неограничивающие примеры лигандов, которые могут быть связаны с помощью GPCR, включают (-)-адреналин, (-)-норадреналин, (лизо)фосфолипидные медиаторы, [дез-Arg10]каллидин, [дез-Arg9]брадикинин, [дез-Gln14]грелин, [Нур3]брадикинин, [Leu]энкефалин, [Met]энкефалин, 12-гидроксигептадекатриеновую кислоту, 12R-НЕТЕ,

12S-НЕТЕ, 12S-НРЕТЕ, 15S-НЕТЕ, 17 β -эстрадиол, 20-гидрокси-ЛТВ4, 2-арахидоноилглицерин, 2-олеоил-LPA, 3-гидроксиоктановую кислоту, 5-гидрокситриптамиин, 5-оксо-15-НЕТЕ, 5-оксо-ЕТЕ, 5-оксо-ЕТrЕ, 5-оксо-ОДЕ, 5S-НЕТЕ, 5S-НРЕТЕ, 7 α ,25-дигидроксихолестерин, ацетилхолин, АСТН, аденозиндифосфат, аденозин, адреномедуллин 2/интермедин, адреномедуллин, амилин, анандамид, ангиотензин II, ангиотензин III, аннексин I, ранний эндогенный лиганд аполинового рецептора, аполин-13, аполин-17, аполин-36, аспирин-зависимый липоксин А4, аспирин-зависимый резолвин D1, АТФ, бета-дефензин 4А, большой диноρφин, пептид 8-22 мозгового вещества надпочечника крупного рогатого скота, брадикинин, С3а, С5а, Са2+, пептид, связанный с геном кальцитонина, катепсин, кальцитонин G, ССК-33, ССК-4, ССК-8, ССL1, ССL11, ССL13, ССL14, ССL15, ССL16, ССL17, ССL19, ССL2, ССL20, ССL21, ССL22, ССL23, ССL24, ССL25, ССL26, ССL27, ССL28, ССL3, ССL4, ССL5, ССL7, ССL8, хемерин, хенодесоксихолевую кислоту, холевую кислоту, кортикотрофин-высвобождающий гормон, ССТ-17, СХ3СL1, СХСL1, СХСL10, СХСL11, СХСL12 α , СХСL12 β , СХСL13, СХСL16, СХСL2, СХСL3, СХСL5, СХСL6, СХСL7, СХСL8, СХСL9, цистеиниллейкотриены (CysLT), урациловые нуклеотиды, дезоксихолевую кислоту, дигидрофингозин-1-фосфат, диолеоилфосфатидную кислоту, дофамин, диноρφин А, диноρφин А-(1-13), диноρφин А-(1-8), диноρφин В, эндоморфин-1, эндотелин-1, эндотелин-2, эндотелин-3, эндотелин-3, F2L, свободные жирные кислоты, FSH, GABA, галанин, галанин-подобный пептид, желудочный ингибиторный полипептид, гастрин-17, гастрин-высвобождающий пептид, грелин, GHRH, глюкагон, амид глюкагоноподобного пептида 1-(7-36), глюкагоноподобный пептид 1-(7-37), глюкагоноподобный пептид 2, глюкагоноподобный пептид 2-(3-33), GnRH I, GnRH II, GRP-(18-27), hCG, гистамин, гуманин, INSL3, INSL5, каллидин, кисспептин-10, кисспептин-13, кисспептин-14, кисспептин-54, кинуреновую кислоту, большой нейромедин N, большой нейротензин, L-глутаминовую кислоту, LH, литохоловую кислоту, L-молочную кислоту, карбоновые кислоты с длинной цепью, LPA, ЛТВ4, LTC4, LTD4, LTE4, LXA4, Lys-[Нур3]-брадикинин, лизофосфатидилинозитол, лизофосфатидилсерин, жирные кислоты со средней длиной цепи, меланин-концентрирующий гормон, мелатонин, метилкарбамил PAF, Mg2+, мотилин, N-арахидоноилглицин, нейрокинин А, нейрокинин В, нейромедин В, нейромедин N, нейромедин S-33, нейромедин U-25, нейроностагин, нейропептид AF, нейропептид В-23, нейропептид В-29, нейропептид FF, нейропептид S, нейропептид SF, нейропептид W-23, нейропептид W-30, нейропептид Y, нейропептид Y-(3-36), нейротензин, ноцицептин/орфанин FQ, N-олеоилэтаноламид, обестатин, октопамин, орексин-А, орексин-В, оксистеролы, окситоцин, PACAP-27, PACAP-38, PAF, полипептид поджелудочной железы, пептид YY, PGD2, PGE2, PGF2 α , PGI2, PGJ2, PHM,

фосфатидилсерин, PHV, прокинетицин-1, прокинетицин-2, прокинетицин-2 β , просапозин, PrRP-20, PrRP-31, PTH, PTHrP, PTHrP-(1-36), QRFP43, релаксин, релаксин-1, релаксин-3, резолвин D1, резолвин E1, RFRP-1, RFRP-3, R-спондины, секретин, сериновые протеазы, сфингозин-1-фосфат, сфингозилфосфорилхолин, SRIF-14, SRIF-28, вещество P, янтарную кислоту, тромбин, тромбоксан A2, TIP39, Т-кинин, TRH, TSH, тирамин, UDP-глюкозу, уридиндифосфат, урокортин 1, урокортин 2, урокортин 3, родственный уротензину II пептид, уротензин-II, вазопрессин, VIP, Wnt, Wnt-1, Wnt-10a, Wnt-10b, Wnt-11, Wnt-16, Wnt-2, Wnt-2b, Wnt-3, Wnt-3a, Wnt-4, Wnt-5a, Wnt-5b, Wnt-6, Wnt-7a, Wnt-7b, Wnt-8a, Wnt-8b, Wnt-9a, Wnt-9b, XCL1, XCL2, Zn²⁺, α -CGRP, α -кетоглутаровую кислоту, α -MSH, α -неоэндофин, β -аланин, β -CGRP, β -D-гидроксималяную кислоту, β -эндофин, β -MSH, β -неоэндофин, β -фенилэтиламин и γ -MSH.

[0066] С помощью различных способов на основе поверхностных взаимодействий лиганда GPCR с GPCR анализировали последовательности связывающих GPCR доменов. Например, проводят мультивидовой вычислительный анализ. В соответствии с некоторыми вариантами, проводят структурный анализ. В соответствии с некоторыми вариантами, проводят анализ последовательностей. Анализ последовательностей можно проводить с использованием базы данных, известной в настоящей области техники. Неограничивающие примеры баз данных включают без ограничения NCBI BLAST (blast.ncbi.nlm.nih.gov/Blast.cgi), UCSC Genome Browser (genome.ucsc.edu/), UniProt (www.uniprot.org/) и IUPHAR/BPS Guide to PHARMACOLOGY (guidetopharmacology.org/).

[0067] В настоящем документе описаны связывающие GPCR домены, сконструированы на основе анализа последовательностей среди различных организмов. Например, анализ последовательностей проводят для выявления гомологичных последовательностей у разных организмов. Иллюстративные организмы включают без ограничения мышь, крысу, лошадь, овцу, корову, примата (например, шимпанзе, павиана, гориллу, орангутанга, мартышку), собаку, кошку, свинью, осла, кролика, рыбу, муху и человека.

[0068] После выявления связывающих GPCR доменов можно создать библиотеки, содержащие нуклеиновые кислоты, кодирующие связывающие GPCR домены. В соответствии с некоторыми вариантами, библиотеки связывающих GPCR доменов содержат последовательности связывающих GPCR доменов, сконструированные на основе конформационных взаимодействий лигандов, пептидных взаимодействий лигандов, низкомолекулярных взаимодействий лигандов, внеклеточных доменов GPCR или антител, целью которых являются GPCR. Для создания библиотек белков можно транслировать библиотеки связывающих GPCR доменов. В соответствии с некоторыми вариантами,

библиотеки связывающих GPCR доменов транслируют для создания библиотек пептидов, библиотек иммуноглобулинов, их производных или их комбинаций. В соответствии с некоторыми вариантами, библиотеки связывающих GPCR доменов транслируют для создания библиотек белков, которые дополнительно модифицируют для создания библиотек пептидомиметиков. В соответствии с некоторыми вариантами, библиотеки связывающих GPCR доменов транслируют для создания библиотек белков, которые применяют для создания библиотек малых молекул.

[0069] Описанные в настоящем документе способы предусматривают синтез библиотек связывающих GPCR доменов, содержащих нуклеиновые кислоты, каждая из которых кодирует заданный вариант по меньшей мере одной заданной эталонной последовательности нуклеиновой кислоты. В некоторых случаях заданная эталонная последовательность представляет собой последовательность нуклеиновой кислоты, кодирующую белок, а библиотека вариантов содержит последовательности, кодирующие вариант по меньшей мере отдельного кодона, так чтобы при стандартных процессах трансляции создавалось множество различных вариантов одного остатка в последующем белке, кодируемом синтезированной нуклеиновой кислотой. В соответствии с некоторыми вариантами, библиотеки связывающих GPCR доменов содержат различные нуклеиновые кислоты, совместно кодирующие варианты в нескольких положениях. В соответствии с некоторыми вариантами, библиотека вариантов содержит последовательности, кодирующие вариант по меньшей мере отдельного кодона домена в связывающем GPCR домене. В соответствии с некоторыми вариантами, библиотека вариантов содержит последовательности, кодирующие варианты множества кодонов в связывающем GPCR домене. Иллюстративное количество кодонов у варианта включает без ограничения по меньшей мере или приблизительно 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300 или более 300 кодонов.

[0070] Описанные в настоящем документе способы предусматривают синтез библиотек, содержащих нуклеиновые кислоты, кодирующие связывающие GPCR домены, причем библиотеки содержат последовательности, кодирующие вариант цепи связывающих GPCR доменов. В соответствии с некоторыми вариантами, библиотека содержит последовательности, кодирующие вариант с длиной, которая по меньшей мере или приблизительно на 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300 или более 300 кодонов меньше по сравнению с заданной эталонной последовательностью. В соответствии с некоторыми вариантами, библиотека содержит последовательности, кодирующие вариант с длиной, которая по меньшей мере или приблизительно на 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75,

80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300 или более 300 кодонов больше по сравнению с заданной эталонной последовательностью.

[0071] После выявления связывающих GPCR доменов связывающие GPCR домены можно поместить в остовы, как описано в настоящем документе. В соответствии с некоторыми вариантами, остовы представляют собой иммуноглобулины. В соответствии с некоторыми вариантами, связывающие GPCR домены помещают в участок CDR-H3. Связывающие GPCR домены, которые можно поместить в остовы, также можно назвать мотивом. Остовы, содержащие связывающие GPCR домены, можно сконструировать на основе связывания, специфичности, стабильности, экспрессии, фолдинга или последующей активности. В соответствии с некоторыми вариантами, остовы, содержащие связывающие GPCR домены, могут контактировать с GPCR. В соответствии с некоторыми вариантами, остовы, содержащие связывающие GPCR домены, могут обеспечивать высокоаффинное связывание с GPCR. Иллюстративные аминокислотные последовательности связывающих GPCR доменов описаны в **таблице 2**.

Таблица 2. Аминокислотные последовательности GPCR

SEQ ID NO	GPCR	Аминокислотная последовательность
1	CXCR4	MEGISIYTSDNYTEEMGSGDYDSMKEPCFREENANFNKIFLP TIYSIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVAD LLFVITLPFWAVDAVANWYFGNFLCKAVHVIYTVNLYSSVL ILAFISLDRYLAIVHATNSQRPRKLLAEKVVYVGVWIPALLL TIPDFIFANVSEADDRYICDRFYPNDLWVVVFQFQHIMVGLI LPGIVILSCYCIISKLSHSHKQKRRKALKTTVILILAFFACWL PYYIGISIDSFILLEIKQGCEFENTVHKWISITEALAFFHCCLN PILYAFLGAKFKTSAQHALTSVSRGSSLKILSKGKRGGHSSVS TESESSSFHSS
2	CCR4	MNPTDIADTTLDESIYSNYLYESIPKCTKEGIKAFGELFLPP LYSLVVFVGLLGNSVVVLVLFKYKRLRSMTDVYLLNLAISD LLFVFSLPFWGYAADQWVFGGLGLCKMISWMYLVGFYSGI FFVMLMSIDRYLAIVHAVFSLRARTLTYGVITSLATWSVAVF ASLPGFLFSTCYTERNHTYCKTKYSLNSTTWKVLSSLEINILG LVIPLGIMLFCYSMIIRTLQHCKNEKKNKAVKMIFAVVVLFL GFWTPYNIVLFLETLVELEVLQDCTFERYLDYAIQATETLAF

		VHCCLNPIIYFFLGEKFRKYILQLFKTCRGLFVLCQYCGLLQI YSADTPSSSYTQSTMDHDLHDAL
3	GCGR	MPPCQPQRPLLLLLLLACQPQVPSAQVMDFLFEKWKLYGD QCHHNLSELLPPTELVCNRTFDKYSWPDTPANTTANISCPW YLPWHHKVQHRFVFKRCGPDGQWVRGPRGQPWRDASQCQ MDGEEIEVQKEVAKMYSSFQVMYTVGYSLSLGALLLAIL GGLSKLHCTRNAIHANLFAFVLKASSVLVIDGLLRTRYSQK IGDDLSVSTWLSDGAVAGCRVAAVFMQYGIVANYCWLLVE GLYLHNLLGLATLPERSSFFSLYLIGIGWGAPMLFVVPWAVVK CLFENVQCWTSNDNMGFWWILRFPVFLAILINFFIFVRIVQLL VAKLRARQMHTDYKFRLLAKSTLTLIPLLGVHEVVFVAVTD EHAQGTLSAKLFFDLFLSSFQGLLVAVLYCFLNKEVQSELR RRWHRWRLGKVLWEERNTSNHRASSSPGHGPPSKELQFGR GGGSQDSSAETPLAGGLPRLAESPF
4	mGluR5	MVLLILSVLLKEDVRGSAQSSERRVVAHMPGDIIIIGALFSV HHQPTVDKVERKCGAVREYGIQRVEAMLHTLERINSDPT LLPNITLGCEIRDSCWHSVAVALEQSIEFIRDSLISSEEEGLVR CVDGSSSSFRSKKPIVGVIGPGSSSVAIQVQNLLQLFNIPQIAY SATSMDSLKTLFKYFMRVVPDAQQARAMVDIVKRYNWT YVSAVHTEGNYGESGMEAFKDMAKEGICIAHSYKIYSNAG EQSFDKLLKKTSHLPKARVVACFCEGMTVRGLLMAMRRL GLAGEFLLGSDGWADRYDVTGQYQREAVGGITIKLQSPDV KWFDDYYLKLREPETNHRNPWFQEFWQHRFQCRLEGFPQEN SKYNKTCNSSLTLKTHHVQDSKMGFVINAIYSMAYGLHNM QMSLCPGYAGLCDAMKPIDGRKLLLESLMKTNFTGVSAGDIL FDENGDSPGRYEIMNFKEMGKDYFDYINVGSWDNGELKMD DDEVWSKKSNIIRSVCSEPCEKGQIKVIRKGEVSCCWTCPC KENEYVFDEYTCACQLGSWPTDDLGTCDLIPVQYLRWGD PEPIAAVVFACLGLLATLFTVVFIIYRDTPVVKSSSRELCYII LAGICLGYLCTFCLIAKPKQIYCYLQRIGIGLSPAMSYSALVT KTNRIRILAGSKKICKKPRFMSACAQLVIAFILICIQLGIIV ALFIMEPPDIMHDYPSIREVYLCNTTNLGVVTPPLGYNGLLIL SCTFYAFKTRNVPANFNKAKYIAFTMYTTCTIWLAFVPIYFGS NYKIITMCFVSLSATVALGCMFVPKVYIILAKPERNVRSAFT TSTVVRMHVGDGKSSSAASRSSLVNLWKRRGSSGETLRYK

		<p>DRRLAQHKSEIECFTPKGSMGNGGRATMSSSNGKSVTWAQ NEKSSRGQHLWQRLSIHINKKENPNQTAVIKPFPKSTESRGL GAGAGAGGSAGGVGATGGAGCAGAGPGGPESPDAGPKAL YDVAEAEHFAPARPRSPSPISTLSHRAGSASRTDDDVP SLH SEPVARSSSSQGSLMEQISSVVTRFTANISELNSMMLSTAAPS PGVGAPLCSSYLIPKEIQLPTTMTTFAEIQLPAIEVTGGAQPA AGAQAAGDAARESPAAGPEAAAAPDLEELVALTPPSPFRD SVDSGSTTPNSPVSEALCIPSSPKYDTLIIRDYTSSSSL</p>
5	GLP-1R	<p>RPQGATVSLWETVQKWREYRRQCQRSLTEDPPPATDLFCNR TFDEYACWPDGEPGSFVNVSCPWYLPWASSVPQGHVYRFC TAEGLWLQKDNSSLPWRDLSECEESKRGRSSPEEQLLFLYII YTVGYALSFSALVIASAILLGFRHLHCTRNYIHLNLFASFILR ALSVFIKDAALKWMYSTAAQQHQWDGLLSYQDSLSCRLVF LLMQYCVAANYWLLVEGVLYTLLAFSVLSEQWIFRLYV SIGWGVPLLFVVPWGIVKYL YEDEGCWTRNSNMNYWLIIRL PILFAIGVNFLIFVRVICIVVSKLKANLMCKTDIKCRLAKSTLT LIPLLGTHEVIFAFVMDEHARGTLRFIKLFTELSFTSFQGLMV AILYCFVNNEVQLEFRKSWERWRLEHLHIQRDSSMKPLKCP TSSLSSGATAGSSMYTATCQASCS</p>
6	GABAB	<p>MLLLLLLAPLFLRPPGAGGAQTPNATSEGCQIIHPPWEGGIR YRGLTRDQVKAINFLPVDYEIEYVCRGEREVVGPKVRKCLA NGSWTDMDTPSRCVRIKSKSYLTLENGKVFLTGGDLPALDG ARVDFRCDFHLVGSSRSICSQGWSTPKPHCQVNRTPHSE RRAVYIGALFPMMSGGWPGGQACQPAVEMALEDVNSRRDILP DYELKLIHHSKCDPGQATKYL YELLYNDPIKIILMPGCSSVS TLVAEAARMWNLIVLSYGSSSPALSNRQRFPTFFRTHPSATL HNPTRVKLFKKGWKKIATIQQTTEVFTSTLDDLEERVKEA GIEITFRQSFFSDPAVPVKNLKRQDARIIVGLFYETEARKVFC EVYKERLFGKKYVWFLIGWYADNWFKIYDPSINCTVDEMT EAVEGHITTEIVMLNPANTRISISNMTSQUEFVEKLTKRLKRHP EETGGFQEAPLAYDAIWALALALNKTSGGGGRSGVRLEDFN YNNQTITDQIYRAMNSSFEGVSGHVVDASGSRMAWTLIE QLQGGSYKKIGYYDSTKDDL SWSKTDKWIGGSPADQTLVI KTRFRLSQKLFISVSVLSSLGIVLAVVCLSFNIYN SHVRYIQNS QPNNLNLTA VGCSLALAAVFPLGLDGYHIGRNQFPFVCQAR</p>

		LWLLGLGFSLGYGSMFTKIWWVHTVFTKKEEKKEWRKTLE PWKLYATVGLLVGMDVLT LAIWQIVDPLHRTIETFAKEPK EDIDVSILPQLEHCSSRKMNTWLGIFYGYKGLLLLLGIFLAYE TKSVSTEKINDHRAVGMAIYNVAVLCLITAPVTMILSSQQDA AFAFASLAIVFSSYITLVVLFVPMRRLITRGEWQSEAQDTM KTGSSTNNNEEEKSRLLKENRELEKIIAEKEERVSELRHQLQ SRQQLRSRRHPPTPPEPSGGLPRGPPEPPDRLSCDGSRVHLLY K
7	OPRM1	MDSSAAPTNASNCTDALAYSSCSPAPSPGSWVNLSHLDGNL SDPCGPNRTDLGGRDSLCPPTGSPSMITAITIMALYSIVCVVG LFGNFLVMYVIVRYTKMKTATNIYIFNLALADALATSTLPFQ SVNYLMGTWPFGTILCKIVISIDYYNMFTSIFTLCTMSVDRYI AVCHPVKALDFRTPRNAKIINVCNWILSSAIGLPVMFMATTK YRQGSIDCTLTFSHPTWYWENLLKICVFIFAFIMPVLIITVCY GLMILRLKSVRMLSGSKEKDRNLRRITRMVLVVAVFIVCW TPIHIYVIAKALVTIPETTFQTVSWHFCIALGYTNSCLNPVLYA FLDENFKRCFREFCIPTSSNIEQQNSTRIRQNTRDHPSTANTV DRTNHQLENLEAETAPLP
8	OPRK1	MDSPIQIFRGEPGPTCAPSACLPPNSSAWFPGWAEPDSNGSA GSEDAQLEPAHISPAIPVIITAVYSVVFVVGLVGNLSLVMFVIIR YTKMKTATNIYIFNLALADALVTTTMPFQSTVYLMNSWPFG DVLCKIVISIDYYNMFTSIFTLTMMSVDRYIAVCHPVKALDF RTPKAKIINICIWLLSSSVGISAIVLGGTKVREDVDVIECSLQ FPDDDYSSWDLFMKICVFIFAFVIPVLIIVCYTLMILRLKSV RLLSGSREKDRNLRRITRLVLVVAVFVVCWTPIHIFILVEAL GSTSHSTAALSSYYFCIALGYTNSLNPILYAFLDENFKRCFR DFCFPLKMRMERQSTSRVRNTVQDPAYLRDIDGMNKPV
9	C5aR	MDSFNYTTPDYGHYDDKDTLDLNTVPDKTSNTLRVPDILAL VIFAVVFLVGVGNALVVWVTAFAEAKRTINAIWFLNLAVAD FLSCLALPILFTSIVQHHPFPGGAACSILPSLILLNMYASILL LATISADRFLLVFKPIWCQNFAGLAWIACAVAWGLALLL TIPSFLYRVVREEYFPPKVLGVDYSHDKRRERAVAIVRLVL GFLWPLLTLTICYTFILLRTWSRRATRSTKTLKVVAVVASF FIFWLPYQVTGIMMSFLEPSSPTFLLLKKLDSLCSFAYINCCI

		NPIIYVVAGQGFQGRRLRKSLLRNVLTEESVVRESKSFTRS TVDTMAQKTQAV
10	CGRP	ELEESPEDSIQLGVTRNKIMTAQYECYQKIMQDPIQQAEGVY CNRTWDGWLCWNDVAAGTESMQLCPDYFQDFDPSEKVTKI CDQDGNWFRHPASNRTWTNYTQCNVNTHEKVKTALNLFY LTIIGHGLSIASLLISLGIFFYFKSLSCQRITLHKNLFFSFVCNS VVTIIIHLTAVANNQALVATNPVSCKVSQFIHLYLMGCNYFW MLCEGIYLHTLIVVAVFAEKQHLMWYYFLGWGFPLIPACIH AIARSLYYNDNCWISSDTHLLYIIHGPICAALLVNLFFLLNIV RVLITKLVTHQAESNLYMKAVRATLILVPLLGIEFVLIPWR PEGKIAEEVYDYIMHILMHFQGLLVSTIFCFFNGEVQAILRRN WNQYKIQFGNSFSNSEALRSASYTVSTISDGGPGYSHDCPSEH LNGKSIHDIENVLLKPENLYN
11	Мускариновый M1	MNTSAPPAVSPNITVLAPGKGPWQVAFIGITTGLLSLATVTG NLLVLISFKVNTELKTVNNYFLLSLACADLIIGTFSMNLYTTY LLMGHWALGTLACDLWLALDYVASNASVMNLLLISFDRYF SVTRPLSYRAKRTPRRAALMIGLAWLVSFVLWAPAILFWQY LVGERTVLAGQCYIQFLSQPIITFGTAMAAFYLPVTVMCTLY WRIYRETENRARELAALQGSETPGKGGSSSSSERSQPGAEG SPETPPGRCCRCCRAPRLLQAYSWKEEEEEDEGSMESLTSSE GEEPGSEVVIKMPMVDPEAQAPTKQPPRSSPNTVKRPTKKG RDRAGKGQKPRGKEQLAKRKTFSLVKEKKAARTLSAILLAF ILTWTPYNIMVLVSTFCKDCVPETLWELGYWLCYVNSTINP MCYALCNKAFRDTRLLLLCRWDKRRWRKIPKRPGSVHRT PSRQC
12	Мускариновый M4	MANFTPVNGSSGNQSVRLVTSSSHNRYETVEMVFIATVTGS LSLVTVVGNILVMLSIKVNRLQTVNNYFLFSLACADLIIGA FSMNLYTVYIIKGYWPLGAVVCDLWLALDYVVSNASVMNL LIISFDRYFCVTKPLTYPARRTTKMAGLMIAAAWVLSFVLW APAILFWQFVVGKRTVPDNQCFIQFLSNPAVTFGTAAAFYL PVVIMTVLYIHISLASRSRVHKHRPEGPKKAKTLAFLKSPL MKQSVKPPPPGEAAREELRNGKLEEAPPPALPPPPRPVADK DTSNESSSGSATQNTKERPADELSTTEATTPAMPAPPLQPRA LNPASRWSKIQIVTKQTGNECVTAIEIVPATPAGMRPAANVA RKFASIARNQVRKKRQMAARERKVTRTIFAILLAFILTWTPY

		NVMVLVNTFCQSCIPDTVWSIGYWLCYVNSTINPACYALCN ATFKKTRHLLLCQYRNIGTAR
13	CCR2	MLSTSRSRFIRNTNESGEEVTTFFDYDYGAPCHKFDVKQIGA QLLPPLYSLVFIFGFVGNMLVVLILINCKKLKCLTDIYLLNLA ISDLLFLITLPLWAHSAANEWVFGNAMCKLFTGLYHIGYFG GIFFIILLTIDRYLAIVHAVFALKARTVTFGVVTSVITWLVAV FASVPGIIFTKCQKEDSVYVCGPYFPRGWNNFHTIMRNILGL VLPLLMVICYSGILKTLLRCRNEKKRHRAVRVIFTIMIVYFL FWTPYNIVILLNTFQEFFGLSNCESTSQLDQATQVTETLGMT HCCINPIIYAFVGEKFRSLFHIALGCRIAPLQKPVCGGPGVRP GKNVKVTTQGLLDGRGKKGKSIGRAPEASLQDKEGA
14	CCR9	MTPTDFTSPIPNMADDYGSESTSSMEDYVNFNFTDFYCEKN NVRQFASHFLPPLYWLVFIVGALGNSLVILVYWYCTRVKTM TDMFLLNLAIADLLFLVTLPFWAIAAADQWKQTFMCKVV NSMYKMFYSCVLLIMCISVDRYIAIAQAMRAHTWREKRL YSKMVCFTIWVLAALCIPEILYSQIKEESGIAICTMVYPSDE STKLKSAVLTCLKVILGFFLPFVVMACCYTIIIHTLIQAKKSSK HKALKVTITVLTVFLVLSQFPYNCILLVQTIDAYAMFISNCAVS TNIDICFQVTQTIAFFHSCLNPVLYVFGFRFRDLVKTLKNL GCISQAQWVSFTRREGSLKLSSMLLETTSGALS
15	GPR174	MPANYTCTRPDGDNTDFRYFIYAVTYTVILVPGLIGNILALW VFYGYMKETKRAVIFMINLAIADLLQVLSLPLRIFYLNDHW PFGPGLCMFCFYLYVNMYASIFLVCISVRRFWFLMYPFRF HDCKQKYDLYISIAGWLIICLACVLFPLLRTSDDTSGNRTKCF VDLPTRNVNLAQSVVMMTIGELIGFVTPLLIVLYCTWKTVLS LQDKYPMAQDLGEKQKALKMILTCAGVFLICFAPYHFSFPL DFLVKSNEIKSCLARRVILIFHSVALCLASLNSCLDPVIYYFST NEFRRRLSRQDLHDSIQLHAKSFVSNHTASTMTPELC
16	MASP-2	TPLGPKWPEPVFGRLASPGFPGEYANDQERRWTLTAPPGYR LRLYFTHFDLELSHLCEYDFVKLSSGAKVLATLCGQESTDTE RAPGKDTFYSLGSSLDITFRSDYSNEKPFTGFEAFYAAEDIDE CQVAPGEAPTCDHHCHNHLGGFYCSCRAGYVLHRNKRTCS ALCSGQVFTQRSGELSSPEYPRYPKLSSTYSISLEEGFSVIL DFVESFDVETHPETLCPYDFLKIQTDREEHGPF CGKTLPHRIE TKSNTVTITFVTDESGDHTGWKIHYTSTAQPCPYPMAPPNGH

		VSPVQAKYILKDSFSIFCETGYELLQGHLPLKSFTAVCQKDG SWDRPMPACSIWDCGPPDDLPSGRVEYITGPGVTTYKAVIQY SCEETFYTMKVNDGKYVCEADGFWTSSKGEKSLPVCEPVC GLSARTTGGRIYGGQKAKPGDFWQVLILGGTTAAGALLYD NWVLTAAHAVYEQKHDASALDIRMGTCLKRLSPHYTQAWSE AVFIHEGYTHDAGFDNDIALIKLNNKVVINSNITPICLPRKEA ESFMRTDDIGTASGWGLTQRGFLARNLMYVDIPIVDHQKCT AAYEKPPYPRGSVTANMLCAGLESGGKDSCRGDSSGALVF LDSETERWFVGGIVSWGSMNCGEAGQYGVYTKVINYIPWIE NIISDF
17	CCR5	MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIF GFVGNMLVILILINCKRLKSMTDIYLLNLAISDLFFLLTVPFW AHYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLA VVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQK EGLHYTCSSHPYSQYQFWKNFQTLKIVILGLVPLLMVIC YSGILKTLRCRNEKKRHRAVRLIFTIMIVYFLFWAPYNIVLL LNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCINPIIYA FVGEKFRNYLLVFFQKHIAKRFCKCCSIFQQEAPERASSVYT RSTGEQEISVGL
18	FSHR	CHHRICHCSNRVFLCQESKVTEIPSDLPRNAIELRFVLTKLRV IQKGA FSGFGDLEKIEISQNDVLEVIEADVFSNLPKLHEIRIEK ANNLLYINPEAFQNLPNLQYLLISNTGIKHLDPVHKIHS LQKV LLDIQDNINIHTIERN SFVGLSFESVILWLNKNGIQEIHNCAFN GTQLDELNLSDNNNLEELPNDVFHGASGPVILDISRTRIHSLP SYGLENLKKLRARSTYNLKKLPTLEKLVALMEASLTYP SHC CAFANWRRQISELHPICNKSILRQEVDYMTQARGQRSSLAED NESSYSRGFDMTYTEFDYDLCNEVVDVTCSPKPD AFNPCEDI MGYNILRVLIWFISILAITGNIIVLVILTTSQYKLTVPFLMCN LAFADLCIGIYLLLIASVDIHTKSQYHNYAIDWQTGAGCDA A GFFT VFA SELSVYTLT AITLERWHTITHAMQLDCKVQLRHA ASVMVMGWIFAFAAALFPIFGISSYMKVSICLPM DIDSPLSQL YVMSLLVLNVLAFVVICGCYIHIYLTVRNPNIVSSSSDTRIAK RMAMLIFTDFLCMAPISFFAISASLKVPLITVSKAKILLVLFHP INSCANPFLYAIFTKNFRRDFFILLSKCGCYEMQAQIYRTETS STVHNTHPRNGHCSSAPRV TNGSTYILVPLSHLAQN

19	mGluR2 PAM	<p>EGPAKKVLTLEGDLVLGGLFPVHQKGGPAEDCGPVNEHRGI QRLEAMLFALDRINRDPHLLPGVRLGAHILDSCSKDTHALEQ ALDFVRASLSRGADGSRHICPDGSYATHGDAPTAITGVIGGS YSDVSIQVANLLRLFQIPQISYASTSAKLSDKSRYDYFARTVP PDFFQAKAMAEILRFFNWTYVSTVASEGDYGETGIEAFELEA RARNICVATSEKVGRAMSRAAFEGVVRALLQKPSARVAVLF TRSEDARELLAASQRLNASFTWVASDGGWGALESVVAGSEG AAEGAITIELASYPISDFASYFQSLDPWNNSRNPWFREFWEQ RFRCFSFRQRDCAAHSLRAVPFEQESKIMFVVNAVYAMAHAL HNMHRALCPNTRLCDAMRPVNGRRLYKDFVLNVKFDAPF RPADTHNEVRFDRFGDGIGRYNIFTYLRAGSGRYRYQKVG WAEGLTLDTSIPWASPSAGPLPASRCSEPCLQNEVKSVPQPG EVCCWLCIPCQPYEYRLDEFTCADCGLYWPNASLTGCFEL PQEYIRWGDAAVGPVTIACLGALATLFLVGVFVRHNATPV VKASGRELCYILLGGVFLCYCMTFIFIAPSTAVCTLRRLGL GTAFSVCYSALLTKTNRIARIFGGAREGAQRPRFISPASQVAI CLALISGQLLIVVAWLVEAPGTGKETAPERREVVTLCNH RDASMLGSLAYNVLLIALCTLYAFKTRKCPENFNEAKFIGFT MYTTCIIWLAFLPIFYVTSSDYRVQTTTMCVSVSLSGSVVLG CLFAPKLHILFQPQKNVVSHRAPTSRFGSAAARASSSLGQGS GSQFVPTVCNGREVVDSTTSSL</p>
20	mGluR3	<p>LGDHNFLRREIKIEGDLVLGGLFPINEKGTGTEECGRINEDRG IQRLEAMLF AIDEINKDDYLLPGVKLG VHILDTCSRDTYALE QSLEFVRASLTKVDEAEYMCPDGSYAIQENIPLLIAGVIGGSY SSVSIQVANLLRLFQIPQISYASTSAKLSDKSRYDYFARTVPP DFYQAKAMAEILRFFNWTYVSTVASEGDYGETGIEAFEQEA RLRNICIAAEKVGRSNIRKSYDSVIRELLQKPNARVVVLFM RSDDSRELIAAASRANASFTWVASDGGWGAQESIIGSEHVA YGAITLELASQPVRQFDRYFQSLNPYNNHRNPWFRDFWEQK FQCSLQNKRNHRRVCDKHLAIDSSNYEQESKIMFVVNAVYA MAHALHKMQRTLCPNTTKLCDAMKILDGKKLYKDYLLKIN FTAPFNPNDADSIVKFDTFDGMGRYNVFNFNQNVGGKYS YLKVGHWAE TSLDVNSIHW SRNSVPTSQCSDPCAPNEMKN MQPGDVCCWICIPCEPYEYLADEFTCMDCGSGQWPTADLTG CYDLPEDYIRWEDAWAIGPVTIACLGFMCTCMVVTVFIKHN</p>

		<p>NTPLVKASGRELCYILLFGVGLSYCMTFFFIKPSPVICALRR LGLGSSFAICYALLTKTNCIARIFDGVKNGAQRPKFISPSSQ VFICLGLLIVQIVMVS VWLILEAPGTRRYTLAEKRETVILKCN VKDSSMLISLTYDVILVILCTVYAFKTRKCPENFNEAKFIGFT MYTTCIIWLAFLPIFYVTSSDYRVQTTTMCISVSLSGFVVLGC LFAPKVHILFQPQKNV VTHRLHLNRFSVSGTGTTYSQSSAST YVPTVCNGREVLDSTTSSL</p>
21	mGluR4	<p>KPKGHPHMNSIRIDGITLGGFLFPVHGRGSEGKPCGELKKEK GIHRLEAMLFALDRINNDPDLNITLGARILDTC SRDTHALE QSLTFVQALIEKDGTEVRCGSGGPPITKPERVVG VIGASGSS VSIMVANILRLFQIPQISYASTAPDLSDNSRYDFFSRV VPSDT YQAQAMVDIVRALKWNYVSTVASEGSYGESGVEAFIQKSR EDGGVCIAQSVKIPREPKAGEFDKIIRLLETSNARAVIIFANE DDIRRVLEAARRANQTGHFFWMGSDSWGSKIAPVLHLEE V AEGAVTILPKRMSVRGFDRYFSSRTL DNNRRIWFAEFWED NFHCKLSRHALKKGSHVKKCTNRERIGQDSAYEQEGKVQF VIDAVYAMGHALHAMHRDLCPRVGLCPRMDPVDGTQLL KYIRNVNFSGIAGNPVTFNENGDAPGRYDIYQYQLRND SAE YKVGISWTDHLHLRIERMHWPGSGQQLPR SICSLPCQPGERK KTVKGMPCWHCEPCTGYQYQVD RYTCKTCPYDMRPTEN RTGCRPIPIIKLEWGSPWAVLPLFLAVV GIIAATLFVVITFVRY NDTPIVKASGRELSYVLLAGIFLCYATTF LMAEPDLGTCSLR RIFLGLGMSISYAALLTKTNRIYRIFEQ GKRSVSAPRFISPASQ LAITFSLISLQLLGICVWFVVDPSHSV VDFQDQRTLDP RFARG VLKCDISDLSLICLLGYSMLLMVTCTVY AIKTRGVPETFNEA KPIGFTMYTTCIVWLAFIPIFFGTSQSADKLYIQTTTLTVSVSL SASVSLGMLYMPKVYIILFHPEQNV PKRKRSLKAVVTAATM SNKFTQKGNFRPNGEAKSEL CENLEAPALATKQTYV TYTNH AI</p>
22	mGluR7	<p>QEMYAPHSIRIEGDVTLGGFLFPVHAKGPSGVPCGDIKRENGI HRLEAMLYALDQINSDPNLLPNVTLGARILDTC SRDITYALEQ SLTFVQALIQKDTSDVRCTNGEPPVFVKPEK VVG VIGASGSS VSIMVANILRLFQIPQISYASTAPELSDDRRYDFFSRV VPPDSF QAQAMVDIVKALGWNYVSTLASEGSYGEKGVESFTQISKEA GGLCIAQSVRIPQERKDR TIDFDRIKQLLDTPNSRAVVIFAN</p>

		<p> DEDIKQILAAAKRADQVGHFLWVGSDSWGSKINPLHQHEDI AEGAITIQPKRATVEGFDAYFTSRTLENNRRNVWFAEYWEE NFNCKLTISGSKKEDTDRKCTGQERIGKDSNYEQEGKVQFVI DAVYAMAHALHHMNKDLCADYRGVCPHEMEQAGGKLLK YIRNVNFNGSAGTPVMFNKNGDAPGRYDIFQYQTTNTSNPG YRLIGQWTDELQLNIEDMQWGKGVREIPASVCTLPCKPGQR KKTQKGTGCCWTCEPCDGYQYQFDEMTCQHCPYDQRPEN RTGCQDIPKLEWHSPWAVIPVFLAMLGIIATIFVMATFIRYN DTPIVRASGRELSYVLLTGIFLCYIITFLMIAKPDVAVCSFRRV FLGLGMCISYAALLTKTNRIYRIFEQGKKSVTAPRLISPTSQ AITSSLISVQLLGVFIWFGVDPPNIIIDYDEHKTMNPEQARGV LKCDITDLQIICSLGYSILLMVTCTVYAIKTRGVPENFNEAKPI GFTMYTTCIVWLAFIPIFFGTAQSAEKLYIQTTTLTISMNLSAS VALGMLYMPKVYIIIFHPELNVQKRKRSFKAVVTAATMSSR LSHKPSDRPNGEAKTELCENVDPNSPAAKKKYVSYNNLVI </p>
23	CXCR3	<p> MVLEVS DHQVLNDAEVAALLENFSSSYDYGENESDSCCTSP PCPQDFSLNFDRAFLPALYSLLFLLGLLGNGAVAAVLLSRRT ALSSTDTFLLHLAVADTLLVLTPLWAVDAAVQWVFGSGL CKVAGALFNINFYAGALLACISFDRYLNIVHATQLYRRGPP ARVTLTCLAVWGLCLL FALPDFIFLSAHHDERLNATHCQYN FPQVGR TALRVLQLVAGFLLPLLVMAYCYAHILAVLLVSRG QRRLRAMRLVVVVVVAFALCWTPYHLVVLVDILMDL GAL ARNCGRESRVDVAKSVTSGLGYMHCCLNPLLYAFVGVKFR ERMWMLLLRLGCPNQRGLQRQPSSSRDSSWSETSEASYSG L </p>
24	CCR8	<p> MDYTLDL SVTTVTDY YYPDIFSSPCDAELIQ TNGKLL LAVFY CLLFVFSLLGNSLVILV LVVCKKLR SITDVYLLNLALS DLLFV FSFPFQTY YLLDQWVFGTVMCKVVS GFYYIGFYSSMFFITLM SVDRYLAVVHAVYALKVRTIRMGTTLCLAVWLTAIMATIPL LVFYQVASEDGV LQCYSFY NQQTLKWKIFTNFKMNILG LLIP FTIFMFCYIKILHQLKRCQNHNKTKAIRLV LIVVIASLLFWVP FNVVFLTSLHSMHILDGCSISQQLTYATHVTEIISFTHCCVN PVIYAFVGEKFKKHLSEIFQKSCSQIFNYLGRQMPRESCEKSS SCQQHSSRSSSVDYIL </p>

25	Аденозиновый A2a	MPIMGSSVYITVELAIAVLAILGNVLCWAVWLNSNLQNVT NYFVVSLAAADIAVGVLAIPFAITISTGFCAACHGCLFIACFV LVLQSSIFSLLAIAIDRYIAIRIPLRYNGLVTGTRAKGIIAICW VLSFAIGLTPMLGWNNCGQPKEGKNHSQGC GEGQVACLFE DVVPMNYMVYFNFFACVLVPLLLMLGVYLRIFLAARRQLK QMESQPLPGERARSTLQKEVHAAKSLAIIVGLFALCWLP LHI NCFTFFCPDCSHAPLWLMYLAIVLSHTNSV VNPFIYAYRIREF RQTFRKIIRSHVLRQQEPFKAAGTSARVLA AHGSDGEQVSLR LNGHPPGVWANGSAPHPERRPNGYALGLVSGGSAQESQGN TGLPDVELLSHELKGVCPPEPGLDDPLA QDGAGVS
26	Орексиновый OX1	MEPSATPGAQMGVPPGSREPSVPPDY EDEFRLYLWRDYLY PKQYEWVLI AAYVAVFVVALVGNTLVCLAVWRNHMRTV TNYFIVNLSLADVLVTAICLPASLLVDITESWLF GHALCKVIP YLQAVSVSVAVLTLSFIALDRWYAICHPLLFKSTARRARGSI LGIWAVSLAIMVPQA AVMECSSVLP ELANRTRLFSVCDERW ADDLYPKIYHSCFFIVTYLAPLGLMAMAYFQIFRKLWGRQIP GTTSALVRNWKRPDQLGDLEQGLSGEPQPRARAFLAEVKQ MRARRKTAKMLMVLLVFALCYLPISVLNVLKRVFGMFRQ ASDREAVYACFTFSHWLVYANSAANPIIYNFLSGKFREQFKA AFSCCLPGLGPCGLKAPSPRSSASHKSLSLQSRC SISKISEHV VLTSVTTVLP
27	Орексиновый OX2	MSGTKLEDSPPCRNWSSASELNETQEPFLNPTDYDDEEFLRY LWREYLHPKEYEWVLIAGYIIVFVVALIGNVLCVAVWKNH HMRTVTNYFIVNLSLADVLVTITCLPATLVVDITETWFFGQS LCKVIPYLQTVSVSVSVLTLSCIALDRWYAICHPLMFKSTAK RARN SIVIIWIVSCIIMIPQAIVMECSTVF PGLANKTTLFTVCD ERWGGEIYPKMYHICFFLVTYMAPLCLMV LAYLQIFRKLWC RQIPGTSSVVQRKWKPLQPVSQPRGPGOPTKSRMSAVAAEI KQIRARRKTARMLMIVLLVFAICYLPISILNVLKRVFGMFAH TEDRETVYAWFTFSHWLVYANSAANPIIYNFLSGKFREEFKA AFSCCCLGVHHRQEDRLTRGRTSTESRKSLTTQISNFDNISKL SEQVVLTSISTLPAANGAGPLQNW
28	PAR-2	IQGTNRSSKGRSLIGKVDGTSHVTGKGVTVETVFSVDEFSAS VLTGKLTTVFLPIVYTIVFVVG LPSNGMALWVFLFRTKKKHP AVIYMANLALADLLSVIWFPLKIA YHIHGNNWIYGEALCNV

		LIGFFYGNMYCSILFMTCLSVQRYWVIVNPMGHSRKKANIAI GISLAIWLLILLVTIPLYVVKQTIFIPALNITTCHDVLPEQLLV GDMFNYFLSLAIGVFLFPAFLTASAYVLMIRMLRSSAMDENS EKKRKRAIKLIVTVLAMYLICFTPSNLLL VVHYFLIKSQGQSH VYALYIVALCLSTLNSCIDPFVYFVSHDFRDHAKNALLCRS VRTVKQMQVSLTSKKHSRKSSSYSSSSTTVKTSY
29	C3aR	MASFSAETNSTDLLSQPWNEPPVILSMVILSLTFLLGLPGNGL VLWVAGLKMQRVTNTIWFHLTLADLLCCLSLPFLAHLAL QGQWPYGRFLCKLIPSIIVLNMFA SVFLLTAISLDRCLVVKP IWCQNHRNVGMACSICGIWVVA FVMCIPVFVYREIFTTDN HNRCGYKFGLSSSLDYPDFYGDPLENRSLENIVQPPGEMND RLDPSSFQTN DHPWTVPTVFQPQTFQRPSADSLPRGSARLTS QNLYSNVFKPADVVSPKIPSGFPIEDHETSPLDNSDAFLSTHL KLFPSASSNSFYESELPQGFQDYYNLQGF TDDDQVPTPLVAI TITRLVVGFLLPVIMIACYSFIVFRMQRGRFAKSQSKTRVA VVVVAVFLVCWTPYHIFGVLSLLTDPETPLGKTLMSWDHVC IALASANSCFNPLYALLGKDFRKKARQSIQGILEAAFSEELT RSTHCPSNNVISERNSTTV
30	LGR5	GSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSELPSN LSVFTSYLDLSMNNISQLLPNPLPSLRFLEELRLAGNALTYP KGAFTGLYSLKVLMLQNNQLRHVPTEALQNLRSLSLRLDA NHISYVPPSCFSGLHSLRHLWLDDNALTEIPVQAFRSLSALQ AMTLALNKIHHIPDYAFGNLSSLVVLHLHNNRIHSLGKKCFD GLHSLETLDLNYNNLDEFPTAIRTL SNLKELG FHSNNIRSIPE KAFVGNPSLITIHFYDNPIQFVGRSAFQHLPELRTLTLNGASQ ITEFPDLTGTANLESLTLTGAQISSLPQTVCNQLPNLQVLDLS YNLLEDLPSFSVCQKLQKIDLRHNEIYEIKVDTFQQLLSLRS NLAWNKIAIHPNAFSTLPSLIKLDLSSNLLSSFPITGLHGLTH LKL TGNHALQSLISSENFPELKVIEMPYAYQCCAFGVCENAY KISNQWNKGDNSSMDDLHKKDAGMFQAQDERDLEDLDF EEDLKALHSVQCSPSPGPFKPEHLLDGWLIRIGVWTIAVLA LTCNALVTSTVFRSPLYISPIKLLIGVIAAVNMLTGVS SA VLA GVDAFTFGSFARHGAWWENGVGCHVIGFLSIFASESSVFLLT LAALERGF SVKYSAKFETKAPFSSLKVIILLCALLALTMAAV PLLGGSKYGASPLCLPLPFGEPTMGYMVAILLNSLCFLMM

		<p>TIAYTKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNCPV AFLSFSSLINLTFISPEVIKIFILLVVVPLPACLNPLLYILFNPHFK EDLVSLRKQTYVWTRSKHPSLMSINSDDVEKQSCDSTQALV TFTSSSITYDLPPSSVPSPAYPVTESCHLSSVAFVPCL</p>
31	GPR101	<p>MTSTCTNSTRESNSSHTCMPLSKMPISLAHGIIIRSTVLVIFLA ASFVGNIVLALVLQRKPQLLQVTNRFIFNLLVTDLLQISLVAP WVAVTSVPLFWPLNSHFCTALVSLTHLFAFASVNTIVVVS DRYLSIIHPLSYPSKMTQRRGYLLLYGTWIVAILQSTPPLYG WGQAAFDERNALCSMIWGASPSYILSVVSFIVIPLIVMIACY SVVFCARRQHALLYNVKRHSLEVRVKDCVENEDEEGAEEK KEEFQDESEFRQHEGEVKAKEGRMEAKDGLKAKEGSTG TSESSVEARGSEEVRESSTVSDGSMEGKEGSTKVEENSMK ADKGRTEVNQCSIDLGEDDMEFGEDDINFSEDDVEAVNIPES LPPSRNSNSNPPLPRCYQCKAAKVIIFSYVLSLGPYCFLA VLAVWVDVETQVPQWVITIIIWLFFLQCCIHPIVYGYMHKTI KKEIQDMLKFFCKEKPPEKEDSHPDLPGTEGGTEGKIVPSYD SATFP</p>
32	GPR151	<p>MLAAAFADSNSSSMNVSF AHLHFAGGYLPSDSQDWR TIIPA LLVAVCLVGFVGNL CVIGILLHNAWKGKPSMIHSLILNLSLA DLSLLLFSAPIRATAYSKSVWDLGW FVCKSSDWF IHTCMAA KSLTIVVVAKVCFMYASDP AKQVSIHNYTIWSVLVAIWTVA SLLPLPEWFFSTIRHHEGVEMCLVDVPAVAEEFMSMFGKLY PLLA FGLPLFFASFYFWRAYDQCKKRGTKTQNLRNQIRSKQ VTVM LLSIAIISALLWLPEWVAWLWVWHLKAAGPAPPQGFI ALSQVLMFSISSANPLIFLVMSEEFREGLKGVWKWMITKKPP TVSESQETPAGNSEGLPDKVPSPE SPASIP EKEKPSPPSSGK GK TEKAEIPILPDVEQFWHERDTVPSVQDNDPIPWEHEDQETGE GVK</p>
33	GPR161	<p>MSLNSSLSCRKELSNL TEEEGGEGGVIITQFIAIIVITIFVCLGN LVIVVTLYKKS YLLT LSNKFVFSLTLSN FLLSVLVLPFVVTSSI RREWIFGVVWCNFSALLYLLISSASMLTLGVIAIDRYYAVLY PMVYPMKITGNRAVMALVYIWLHSLIGCLPPLFGWSSVEFD EFKWMCVAAWHREPGYTAFWQIWCALFPFLVMLVCYGFIF RVARVKARKVHCGTVVIVEEDAQRTGRKNSSTSTSSSGSRR NAFQGVVYSANQCKALITILVVLGAFMVTWGPYMVVIASE</p>

		ALWKGKSSVSPSLETWATWLSFASAVCHPLIYGLWNKTVRKE LLGMCFGDRYYREPFVQRQRTSRLFSISNRITDLGLSPHLTAL MAGGQPLGHSSSTGDTGFSCSQDSGTDMMLEDYTSDDNPP SHCTCPPKRRSSVTFEDEVQIKEAAKNSILHVKADEVHKS SLD SYAASLAKAIEAEAKINLFGEEALPGVLVTARTVPGGGFGGR RGSRTLVSQRLQLQSIEEGDVLAAEQR
34	GPR17	MSKRSWWAGSRKPPREMLKLSGSDSSQSMNGLEVAPPGLIT NFSLATAEQCGQETPLENMLFASFYLLDFILALVGNTLALWL FIRDHKS GTPANVFLMHLAVADLSCVLVLPTRLVYHFSGNH WPFGEIACRLTGFLFYLNMYASIFYLTCISADRFLAIVHPVKS LKLRRPLYAHLACAFLWVVAVAMAPLLVSPQTVQTNHTV VCLQLYREKASHHALVSLAVAFTFPFITTVTCYLLIIRSLRQG LRVEKRLKTKAVRMIAIVLAIFLVCFVPYHVNRSVYVLHYRS HGASCATQRILALANRITSCLTSLNGALDPIMYFFVAEKFRH ALCNLLCGKRLKGPPPSFEGKTNESSLSAKSEL
35	GPR183	MDIQMANNFTPPSATPQGNDCLDYAHHSTARIVMPLHYSLV FIIGLVGNLLALVVIVQNRKKINSTTLYSTNLVISDILFTTALP TRIAYYAMGFDWRIGDALCRITALVFYINTYAGVNFMTCLSI DRFIAVVHPLRYNKIKRIEHAKGVCIFVWILVFAQTLPLLINP MSKQEAERITCMEYPNFEETKSLPWILLGACFIGYVLPLIILI CYSQICCKLFRATAKQNPLTEKSGVNKKALNTIILIVFVLCF TPYHVAVIIQHMIKKLRFNFLECSQRHSFQISLHFTVCLMNFN CCMDPFIYFFACKGYKRKVMRMLKRQVSVSISAVKSAPEE NSREMTETQMMIHSKSSNGK
36	CRTH2	MSANATLKPLCPILQMSRLQSHSNTSIRYIDHAAVLLHGLA SLLGLVENGVLFFVVGCRMQRQTVVTTWVLHLALS DLLASAS LPFFTYFLAVGHSWELGTTFCCLHSSIFFLNMFASGFLLSAIS LDRCLQVVRPVWAQNHRTVAAAHKVCLVLWALAVLNTVP YFVFRDTISRDLGRIMCYYNVLLLNPGPDRDATCNSRQVAL AVSKFLLAFLVPLAIIASSHAAVSLRLQHRGRRRPGRFVRLV AAVVAAFALCWGPYHVFSLEARAHANPGLRPLVWRGLPF VTSLAFFNSVANPVLVLTCPDMLRKLRRSLRTVLESVLVD DSELGGAGSSRRRRTSSTARSASPLALCSRPEEPRGPALLG WLLGSCAASPQTGPLNRALSSTSS

37	5-HT4	<p>MDKLDANVSSEEGFGSVEKVVLLTFLSTVILMAILGNLLVM VAVCWDRQLRKIKTNYFIVSLAFADLLVSVLVMPFGAIELV QDIWIYGEVFCLVRTSLDVLLTTASIFHLCCISLDRYYAICCCQ PLVYRNKMTPLRIALMLGGCWVIPTFISFLPIMQGWNNIGIID LIEKRKFNQNSNSTYCVFMVNKPYAITCSVVAFYIPFLLMVL AYYRIYVTAKEHAHQIQMLQRAGASSESRPQSADQHSTHRM RTETKAAKTLCIIMGCFCLCWAPFFVTNIVDPFIDYTVPGQV WTAFLWLGYINSGLNPFLYAFLNKSFRRAFLIILCCDDERYR RPSILGQTVPCSTTTINGSTHVLRDAVECGGQWESQCHPPAT SPLVAAQPSDT</p>
38	5-HT6	<p>MVPEPGPTANSTPAWGAGPPSAPGGSGWVAAALCVVIALT AAANSLIALICTQPALRNTSNFFLVSLFTSDLMVGLVVMPP AMLNALYGRWVLARGLCLLWTAFDVMCCSASILNLCLISLD RYLLILSPLRYKLRMTPLRALALVLGAWSLAALASFLPLLLG WHELGHARPPVPGQCRLLASLPFVLVASGLTFFLPSGAICFT YCRILLAARKQAVQVASLTTGMASQASETLQVPRTPRPGVE SADSRRLATKHSRKALKASLTGILLGMFFVTWLPFFVANIV QAVCDCISPGLFDVLTWLGVCNSTMNPIIYPLFMRDFKRALG RFLPCPRCPRERQASLASPSLRTSHSGPRPGLSLQQVLPPLPP DSDSDSDAGSGGSSGLRLTAQLLLPGPEATQDPPLPTRAAAA VNFFNIDPAEPELRPHPLGIPTN</p>
39	CB2	<p>MEECWVTEIANGSKDGLDSNPMKDYMILSGPQKTAVAVLC TLLGLLSALENVAVLYLILSSHQLRRKPSYLFIGSLAGADFLA SVVFACSFVNFHVFHGVDSKAVFLLKIGSVTMTFTASVGSLL LTAIDRYLCLRYPPSYKALLTRGRALVTLGIMWVLSALVSYL PLMGWTCCPRPCSELFPLIPNDYLLSWLLFIAFLFSGIITYYGH VLWKAHQHVASLSGHQDRQVPGMARMRLDVRLAKTLGLV LAVLLICWFPVLALMAHSLATTLSDQVKKAFAFCSMLCLINS MVNPVIYALRSGEIRSSAHHCLAHWKCVRGLGSEAKEEAP RSSVTETEADGKITPWPDSRDLDLSDC</p>
40	Гистаминовый 3	<p>MERAPPDGPLNASGALAGEAAAAGGARGFSAAWTAVLAAL MALLIVATVLGNALVMLAFVADSSLRTQNNFFLLNLAISDFL VGAFCIPLYVPYVLTGRWTFGRGLCKLWLVDYLLCTSSAF NIVLISYDRFLSVTRAVSYRAQQGDTRRAVRKMLLVWVLAFL LLYGPAILSWEYLSGGSSIPEGHCYAEFFYNWYFLITASTLEF</p>

		FTPFLSVTFFNLSIYLNQRRTRLRLDGAREAAGPEPPPEAQPS PPPPPGCWGCWQKGHGGEAMPLHRYGVGEAAVGAEAGEAT LGGGGGGGSSVASPTSSSGSSSRGTERPRSLKRGSKPSASSASL EKRMKMVSQSFTQRFRLSRDRKVAKSLAVIVSIFGLCWAPY TLLMIIRAACHGHCVPDYWYETSFWLLWANSVNPVLYPL CHHSFRRAFTKLLCPQKLLKIQPHSSLEHCWK
41	VPAC-1 или VIPR1	ARLQEECDYVQMIEVQHKQCLEEAQLENETIGCSKMWDNL TCWPATPRGQVVVLACPLIFKLFSSIQGRNVSRSC TDEGWTH LEPGPYPIACGLDDKAASLDEQQTMFYGSVKTGTYTIGYGLSL ATLLVATAILSLFRKLHCTRNYIHMHLFISFILRAAAVFIKDL ALFDSGESDQCSEGSVGCKAAMVFFQYCVMANFFWLLVEG LYLYTLLAVSFFSERKYFWGYILIGWGV PSTFTMVWTIARIH FEDYGCWDTINSSLWWIIGPILTSILVNFILFICIRILLQKLRP PDIRKSDSSPY SRLARSTLLL IPLFGVHYIMFAFFPDNFKPEVK MVFELVVG SFQGFVVAILYCFLNGEVQAELRRKWRRWHLQ GVLGWNPKYRHPSSGGSNGATCSTQVSMLTRVSPGARRSSSF QAEVSLV
42	GIPR	RAETGSKGQTAGELYQRWERYRRECQETLAAAEPPSGLAC NGSFDMYVCWDYAAPNATARASCPWYLPWHHHVAAGFVL RQCGSDGQWGLWRDHTQCENPEKNEAFLDQRLILERLQVM YTVGYSLSLATLLLALLILSLFRRLHCTRNYIHINLFTSFMLR AAAILSRDRLLP RP GPY LGDQALALWNQALAACRTAQIVTQ YCVGAN YTWLLVEGVYLHSLLVLVGGSEEGHFRYYLLLGW GAPALFVIPWVIVRYLYENTQCWERNEVKAIWWIIRTPILMT ILINFLIFIRILGILLSKLRTRQMRCRDYRLRLARSTLTLVPLLG VHEVVFAPVTEEQARGALRFAKLGFEIFLSSFGFLVSVLYC FINKEVQSEIRRGWHHCRLRRSLGEEQRQLPERAFRALPSGS GPGEVPTSRGLSSGTLPGPGNEASRELESYC
43	5-HT1B GPCR	MEEPGAQCAPPPPAGSETWVPQANLSSAPSQNCSAKDYIYQ DSISLPWKVLLVMLLALITLATTLSNAFVIATVYRTRKLHTP ANYLIASLAVTDLLVSILVMPISTMYTVTGRWTLGQVVCDF WLSSDITCCTASILHLCVIALDRYWAITDAVEYSAKRTPKRA AVMIALVWVFSISISLPPFFWRQAKAEVEVSECVVNTDHILY TVYSTVGAFYFPTLLLIALLYGRIYVEARSILKQTPNRTGKRL TRAQLITDSPGSTSSVTSINSRVPDVPSESGSPVYVNQVKVRV

		SDALLEKKKLMAARERKATKTLGILGAFIVCWLPFFIISLVM PICKDACWFHLAIFDFFTWLGYLNSLINPIIYTMSNEDFKQAF HKLIRFKCTS
44	CCR7	QDEVTDDYIGDNTTVDYTLFESLCSKKDVRNFKAWFLPIMY SIICFVGLLGNGLVVLTYYIFKRLKTMDDTYLLNLAVADILFL LTLPFWAYSAAKSWVFGVHFCKLIFAIYKMSFFSGMLLLLCI SIDRYVAIVQAVSAHRHRARVLLISKLSCVGIWILATVLSIPE LLYSDLQRSSSEQAMRCSLITEHVEAFITIQVAQMVIGFLVPL LAMSFCYLVIIRTLLQARNFERNKAIKVIIAVVVVFIVFQLPY NGVVLAQTVANFNITSSSTCELSKQLNIAVDVTYSLACVRCC VNPFLYAFIGVKFRNDLFLFKDLGCLSQEQLRQWSSCRHIR RSSMSVEAETTTTFSP
45	CXCR5	MNYPLTLEMDLENLEDFWELDRLDNYNDTSLVENHLCPA TEGPLMASFKAVFVPVAYSILFLLGVIGNVLVLVILERHRQT RSSTETFLFHLAVADLLL VFILPFAVAEGSVGWVLGTFLCKT VIALHKVNFYCSSLACIAVDRYLAIVHAVHAYRHRLLSI HITCGTIWLVGFLALPEILFAKVSQGHNNLSLPRCTFSQENQ AETHAWFTSRFLYHVAGFLLPMLVMGWCYVGVVHRLRQA QRRPQRQKAVRVAILVTSIFFLCWSPYHIVIFLDTLARLKAV DNTCKLNGSLPVAITMCEFLGLAHCCLNPMLYTFAGVKFRS DLSRLLTKLGCTGPASLCQLFPSWRRSSLSESENATSLTTF
46	GPR119	MESSFSFGVILAVLASLIATNTLVAVAVLLLHKNDGVS LCF TLNLAVADTLIGVAISGLLTDQLSSPSRPTQKTLCSLRMAFVT SSAAASVLTVM LITFDRYLAIKQPFRYLKIMSGFVAGACIAG LWLVSYLIGFLPLGIPMFQQTAYKGQCSFFAVFHPHFVLTLS CVGFFPAMLLFVFFYCDMLKIASMHSQQIRKMEHAGAMAG GYRSPRTPSDFKALRTVSVLIGSFALSWTPFLITGIVQVACQE CHLYLVLERYLWLLGVGNSLLNPLIYAYWQKEVRLQLYHM ALGVKKVLT SFLFLSARNCGPERPRESSCHIVTISSEFDG
47	GPR55	MSQQNTSGDCLFDGVNELMKTLQFAVHIPTFVLGLLLNLAI HGFSTFLKNRWPDYAATSIMINLAVFDLLL VLSLPFKMVLS QVQSPFSLCTLVECLYFVSMYGSVFTICFISMDRFLAIRYPL LVSHLRSRPRKIFGICCTIWVLVWTGSIPIYSFHGKVEKYMCFH NMSDDTWSAKVFFPLEVFGFLLPMGIMGFCCSRSIHILLGRR DHTQDWVQQKACIYSIAASLAVFVVSFLPVHLGFFLQFLVR

		NSFIVECRAKQISISFFLQLSMCFNSVNCCLDVFCYYFVIKEFR MNIRAHRPSRVQLVLQDTTISRG
--	--	---

[0072] Настоящее изобретение относится к остовам, содержащим связывающие GPCR домены, причем последовательности связывающих GPCR доменов служат основой для взаимодействия по меньшей мере с одним GPCR. Последовательность может быть гомологичной или идентичной последовательности лиганда GPCR. В соответствии с некоторыми вариантами, последовательность связывающего GPCR домена характеризуется по меньшей мере или приблизительно 70%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% или 100% идентичности последовательности с SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46 или 47. В соответствии с некоторыми вариантами, последовательность связывающего GPCR домена характеризуется по меньшей мере или приблизительно 95% гомологии с SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46 или 47. В соответствии с некоторыми вариантами, последовательность связывающего GPCR домена характеризуется по меньшей мере или приблизительно 97% гомологии с SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46 или 47. В соответствии с некоторыми вариантами, последовательность связывающего GPCR домена характеризуется по меньшей мере или приблизительно 99% гомологии с SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46 или 47. В соответствии с некоторыми вариантами, последовательность связывающего GPCR домена характеризуется по меньшей мере или приблизительно 100% гомологии с SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46 или 47. В соответствии с некоторыми вариантами, последовательность связывающего GPCR домена содержит по меньшей мере часть, имеющую по меньшей мере или приблизительно 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400 или более 400 аминокислот из SEQ ID NO: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46 или 47.

[0073] Библиотеки, содержащие нуклеиновые кислоты, кодирующие остовы, содержащие связывающие GPCR домены, могут связываться с одним или несколькими GPCR. В соответствии с некоторыми вариантами, остовы, содержащие связывающие GPCR домены, связываются с отдельным GPCR. В соответствии с некоторыми вариантами, остовы, содержащие связывающие GPCR домены, связываются с GPCR в одном и том же семействе или классе. В соответствии с некоторыми вариантами, остовы, содержащие связывающие GPCR домены, связываются с множеством GPCR. Например, остовы являются мультимерными и содержат по меньшей мере 2 остова. В соответствии с некоторыми вариантами, мультимерные остовы содержат по меньшей мере или приблизительно 3, 4, 5, 6, 7, 8 или более 8 остовов. В соответствии с некоторыми вариантами, мультимерные остовы содержат по меньшей мере 2 остова, соединенные с помощью, например, домена димеризации, аминокислотного линкера, дисульфидной связи, химической сшивки или любого другого известного из уровня техники линкера. В соответствии с некоторыми вариантами, мультимерные остовы связываются с одинаковыми GPCR или различными GPCR.

[0074] Настоящее изобретение относится к библиотекам связывающих GPCR доменов, содержащим нуклеиновые кислоты, кодирующие остовы, содержащие связывающие GPCR домены, при этом библиотеки содержат варианты по типу домена, длине домена или вариантам остатков. В соответствии с некоторыми вариантами, домен является участком в остове, содержащем связывающие GPCR домены. Например, участок представляет собой домен VH, CDR-H3 или VL. В соответствии с некоторыми вариантами, домен является связывающим GPCR доменом.

[0075] Описанные в настоящем документе способы предусматривают синтез библиотеки связывающих GPCR доменов, которая содержит нуклеиновые кислоты, при этом каждая из них кодирует заданный вариант по меньшей мере одной заданной эталонной последовательности нуклеиновой кислоты. В некоторых случаях заданная эталонная последовательность представляет собой последовательность нуклеиновой кислоты, кодирующую белок, а библиотека вариантов содержит последовательности, кодирующие вариант по меньшей мере отдельного кодона, так чтобы при стандартных процессах трансляции создавалось множество различных вариантов одного остатка в последующем белке, кодируемом синтезированной нуклеиновой кислотой. В соответствии с некоторыми вариантами, библиотека связывающих GPCR молекул содержит различные нуклеиновые кислоты, совместно кодирующие варианты в нескольких положениях. В соответствии с некоторыми вариантами, библиотека вариантов содержит последовательности, кодирующие вариант по меньшей мере отдельного кодона домена VH, CDR-H3 или VL. В

соответствии с некоторыми вариантами, библиотека вариантов содержит последовательности, кодирующие вариант по меньшей мере отдельного кодона домена в связывающем GPCR домене. Например, варьирует по меньшей мере один отдельный кодон связывающего GPCR домена, как указано в **таблице 2**. В соответствии с некоторыми вариантами, библиотека вариантов содержит последовательности, кодирующие варианты множества кодонов домена VH, CDR-H3 или VL. В соответствии с некоторыми вариантами, библиотека вариантов содержит последовательности, кодирующие варианты множества кодонов в связывающем GPCR домене. Иллюстративное количество кодонов у варианта включает без ограничения по меньшей мере или приблизительно 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300 или более 300 кодонов.

[0076] Описанные в настоящем документе способы предусматривают синтез библиотеки связывающих GPCR доменов, которая содержит нуклеиновые кислоты, при этом каждая из них кодирует заданный вариант по меньшей мере одной заданной эталонной последовательности нуклеиновой кислоты, причем библиотека связывающих GPCR доменов содержит последовательности, кодирующие различные варианты по длине домена. В соответствии с некоторыми вариантами, домен представляет собой домен VH, CDR-H3 или VL. В соответствии с некоторыми вариантами, домен является связывающим GPCR доменом. В соответствии с некоторыми вариантами, библиотека содержит последовательности, кодирующие вариант с длиной, которая по меньшей мере или приблизительно на 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 225, 250, 275, 300 или более 300 кодонов меньше по сравнению с заданной эталонной последовательностью. В соответствии с некоторыми вариантами, библиотека содержит последовательности, кодирующие вариант с длиной, которая по меньшей мере или приблизительно на 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300 или более 300 кодонов больше по сравнению с заданной эталонной последовательностью.

[0077] Настоящее изобретение относится к библиотекам связывающих GPCR доменов, содержащим нуклеиновые кислоты, кодирующие остовы, содержащие связывающие GPCR домены, причем библиотеки связывающих GPCR доменов синтезируют с различным количеством фрагментов. В соответствии с некоторыми вариантами, фрагменты составляют домен VH, CDR-H3 или VL. В соответствии с некоторыми вариантами, библиотеки связывающих GPCR доменов синтезируют по меньшей мере или приблизительно с 2 фрагментами, с 3 фрагментами, с 4 фрагментами, с 5 фрагментами или более чем с 5 фрагментами. Длина каждого из фрагментов нуклеиновой

кислоты или средняя длина синтезированных нуклеиновых кислот может составлять по меньшей мере или приблизительно 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600 или более 600 пар оснований. В соответствии с некоторыми вариантами, длина составляет приблизительно 50-600, 75-575, 100-550, 125-525, 150-500, 175-475, 200-450, 225-425, 250-400, 275-375 или 300-350 пар оснований.

[0078] Библиотеки связывающих GPCR доменов, содержащие нуклеиновые кислоты, кодирующие остовы, содержащие связывающие GPCR домены, которые описаны в настоящем документе, при трансляции содержат различные по длине цепи аминокислот. В соответствии с некоторыми вариантами, длина каждого из аминокислотных фрагментов или средняя длина синтезированных цепей из аминокислот может составлять по меньшей мере или приблизительно 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150 или более 150 аминокислот. В соответствии с некоторыми вариантами, длина цепи из аминокислот составляет приблизительно 15-150, 20-145, 25-140, 30-135, 35-130, 40-125, 45-120, 50-115, 55-110, 60-110, 65-105, 70-100 или 75-95 аминокислот. В соответствии с некоторыми вариантами, длина цепи из аминокислот составляет от приблизительно 22 до приблизительно 75 аминокислот.

[0079] Библиотеки связывающих GPCR доменов, содержащие *de novo* синтезированные варианты последовательностей, кодирующих остовы, содержащие связывающие GPCR домены, содержат определенное количество вариантов последовательностей. В соответствии с некоторыми вариантами, определенное количество вариантов последовательностей синтезируют *de novo* для CDR-H1, CDR-H2, CDR-H3, CDR-L1, CDR-L2, CDR-L3, VL, VH или их комбинации. В соответствии с некоторыми вариантами, определенное количество вариантов последовательностей синтезируют *de novo* для каркасного элемента 1 (FW1), каркасного элемента 2 (FW2), каркасного элемента 3 (FW3) или каркасного элемента 4 (FW4). В соответствии с некоторыми вариантами, определенное количество вариантов последовательностей *de novo* синтезируют для связывающего GPCR домена. Например, определенное количество вариантов последовательностей составляет от приблизительно 1 до приблизительно 10 последовательностей для домена VH, приблизительно 10^8 последовательностей для связывающего GPCR домена и от приблизительно 1 до приблизительно 44 последовательностей для домена VK. См. **фиг. 2**. Определенное количество вариантов последовательностей может составлять по меньшей мере или приблизительно 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500 или более 500 последовательностей. В

соответствии с некоторыми вариантами, определенное количество вариантов последовательностей составляет приблизительно 10-300, 25-275, 50-250, 75-225, 100-200 или 125-150 последовательностей.

[0080] Библиотеки связывающих GPCR доменов, содержащие de novo синтезированные варианты последовательностей, кодирующих остовы, содержащие связывающие GPCR домены, характеризуются повышенным разнообразием. Например, варианты создают путем размещения вариантов связывающих GPCR доменов в вариантах остовов иммуноглобулина, содержащих N-концевые варианты CDR-H3 и C-концевые варианты CDR-H3. В соответствии с некоторыми вариантами, варианты включают варианты с различным созреванием аффинности. Альтернативно или в сочетании, варианты включают варианты в других участках иммуноглобулина, в том числе без ограничения CDR-H1, CDR-H2, CDR-L1, CDR-L2 и CDR-L3. В соответствии с некоторыми вариантами, определенное количество вариантов у библиотек связывающих GPCR доменов составляет по меньшей мере или приблизительно 10^4 , 10^5 , 10^6 , 10^7 , 10^8 , 10^9 , 10^{10} или более 10^{10} неидентичных последовательностей. Например, библиотека, содержащая приблизительно 10 вариантов последовательностей для участка VH, приблизительно 237 вариантов последовательностей для участка CDR-H3 и приблизительно 43 вариантов последовательностей для участка VL и CDR-L3 содержит 10^5 неидентичных последовательностей ($10 \times 237 \times 43$). См. **фиг. 4А-4В**.

[0081] После синтеза библиотек связывающих GPCR доменов, содержащих нуклеиновые кислоты, кодирующие остовы, содержащие связывающие GPCR домены, библиотеки можно применять для скрининга и анализа. Например, библиотеки анализируют на способность экспонирования и пэннинг. В соответствии с некоторыми вариантами, способность экспонирования оценивают с помощью селективируемой метки. Иллюстративные метки включают без ограничения радиоактивную метку, флуоресцентную метку, фермент, хемилюминесцентную метку, колориметрическую метку, аффинную метку или другие меченные элементы или метки, которые известны в настоящей области техники. В соответствии с некоторыми вариантами, меткой является гистидин, полигистидин, тус, гемагглютинин (НА) или FLAG. Например, как можно видеть на **фиг. 3**, библиотеки связывающих GPCR доменов содержат нуклеиновые кислоты, кодирующие остовы, содержащие связывающие GPCR домены с множественными метками, такими как GFP, FLAG и Lucy, а также штрихкодированная ДНК. В соответствии с некоторыми вариантами, библиотеки анализируют путем секвенирования с помощью различных способов, включая без ограничения секвенирование в режиме реального времени (SMRT) одиночной молекулы, секвенирование методом молекулярных колоний, секвенирование с

помощью лигирования, секвенирование с обратимым терминатором, секвенирование с обнаружением протонов, секвенирование с ионными полупроводниками, секвенирование с помощью нанопор, электронное секвенирование, пиросеквенирование, секвенирование Максама-Гилберта, секвенирование с разрывом цепи (например, Сэнгера), секвенирование +S или секвенирование путем синтеза.

[0082] *Системы экспрессии*

[0083] Настоящее изобретение относится к библиотекам, содержащим нуклеиновые кислоты, кодирующие остовы, содержащие связывающие GPCR домены, причем последовательности библиотек характеризуются улучшенной специфичностью, стабильностью, экспрессией, фолдингом или последующей активностью. В соответствии с некоторыми вариантами, описываемые в настоящем документе библиотеки применяют для скрининга и анализа.

[0084] Настоящее изобретение относится к библиотекам, содержащим нуклеиновые кислоты, кодирующие остовы, содержащие связывающие GPCR домены, причем библиотеки нуклеиновых кислот применяют для скрининга и анализа. В соответствии с некоторыми вариантами, скрининг и анализ предусматривают анализ *in vitro*, *in vivo* или *ex vivo*. Клетки для скрининга включают первичные клетки, взятые у живых субъектов или из клеточных линий. Клетки могут иметь прокариотическую (например, бактерии и грибы) или эукариотическую (например, животные и растения) природу. Иллюстративные клетки животных включают без ограничения клетки мыши, кролика, примата и насекомого. В соответствии с некоторыми вариантами, клетки для скрининга подразумевают клеточную линию, включая без ограничения линию клеток яичника китайского хомячка (СНО), линию клеток эмбриональной почки человека (НЕК) или линию клеток почки новорожденного хомячка (ВНК). В соответствии с некоторыми вариантами, описываемые в настоящем документе библиотеки нуклеиновых кислот также можно доставить в многоклеточный организм. Иллюстративные многоклеточные организмы включают без ограничения растение, мышь, кролика, примата и насекомое.

[0085] Описываемые в настоящем документе библиотеки нуклеиновых кислот можно подвергнуть скринингу на различные фармакологические или фармакокинетические свойства. В соответствии с некоторыми вариантами, библиотеки подвергают скринингу с помощью анализов *in vitro*, анализов *in vivo* или анализов *ex vivo*. Например, к фармакологическим или фармакокинетическим свойствам *in vitro*, которые подвергают скринингу, относятся без ограничения аффинность связывания, специфичность связывания и avidность связывания. К иллюстративным фармакологическим или

фармакокинетическим свойствам описываемых в настоящем документе библиотек *in vivo*, которые подвергают скринингу, относятся без ограничения терапевтическая эффективность, активность, свойства по результатам доклинического изучения токсичности, свойства клинической эффективности, свойства клинической токсичности, иммуногенность, активность и свойства клинической безопасности.

[0086] Настоящее изобретение относится к библиотекам нуклеиновых кислот, причем библиотеки нуклеиновых кислот могут экспрессироваться в векторе. Векторы экспрессии для вставки раскрываемых в настоящем документе библиотек нуклеиновых кислот могут включать векторы экспрессии у эукариот или прокариот. К иллюстративным векторам экспрессии относятся без ограничения векторы экспрессии млекопитающего: pSF-CMV-NEO-NH2-PPT-3XFLAG, pSF-CMV-NEO-COOH-3XFLAG, pSF-CMV-PURO-NH2-GST-TEV, pSF-OXB20-COOH-TEV-FLAG(R)-6His, pCEP4 pDEST27, pSF-CMV-Ub-KrYFP, pSF-CMV-FMDV-daGFP, вектор pEF1a-mCherry-N1, вектор pEF1a-tdTomato, pSF-CMV-FMDV-Hygro, pSF-CMV-PGK-Puro, pMCP-tag(m), и pSF-CMV-PURO-NH2-CMYC; векторы экспрессии у бактерий: pSF-OXB20-BetaGal, pSF-OXB20-Fluc, pSF-OXB20 и pSF-Tac; векторы экспрессии у растений: pRI 101-AN DNA и pCambia2301; и векторы экспрессии у дрожжей: pTYB21 и pKLAC2, и векторы экспрессии у насекомых: pAc5.1/V5-His A и pDEST8. В соответствии с некоторыми вариантами, вектор представляет собой pсDNA3 или pсDNA3.1.

[0087] В настоящем документе описаны библиотеки нуклеиновых кислот, которые экспрессируются в векторе с образованием конструкции, содержащей остов, содержащий последовательности связывающих GPCR доменов. В соответствии с некоторыми вариантами, размер конструкции варьирует. В соответствии с некоторыми вариантами, конструкция содержит по меньшей мере или приблизительно 500, 600, 700, 800, 900, 1000, 1100, 1300, 1400, 1500, 1600, 1700, 1800, 2000, 2400, 2600, 2800, 3000, 3200, 3400, 3600, 3800, 4000, 4200, 4400, 4600, 4800, 5000, 6000, 7000, 8000, 9000, 10000 или более 10000 оснований. В соответствии с некоторыми вариантами, конструкция содержит количество оснований, варьирующее в диапазоне 300-1000, 300-2000, 300-3000, 300-4000, 300-5000, 300-6000, 300-7000, 300-8000, 300-9000, 300-10000, 1000-2000, 1000-3000, 1000-4000, 1000-5000, 1000-6000, 1000-7000, 1000-8000, 1000-9000, 1000-10000, 2000-3000, 2000-4000, 2000-5000, 2000-6000, 2000-7000, 2000-8000, 2000-9000, 2000-10000, 3000-4000, 3000-5000, 3000-6000, 3000-7000, 3000-8000, 3000-9000, 3000-10000, 4000-5000, 4000-6000, 4000-7000, 4000-8000, 4000-9000, 4000-10000, 5000-6000, 5000-7000, 5000-8000, 5000-9000, 5000-10000, 6000-7000, 6000-8000, 6000-9000, 6000-10000, 7000-8000, 7000-9000, 7000-10000, 8000-9000, 8000-10000 или 9000-10000 оснований.

[0088] Настоящее изобретение относится к библиотекам, содержащим нуклеиновые кислоты, кодирующие остовы, содержащие связывающие GPCR домены, причем библиотеки нуклеиновых кислот экспрессируют в клетке. В соответствии с некоторыми вариантами, библиотеки синтезируют такими, чтобы у них присутствовала экспрессия репортерного гена. К иллюстративным репортерным генам относятся без ограничения гены синтазы ацетогидроксикислот (AHAS), щелочной фосфатазы (AP), бета-галактозидазы (LacZ), бета-глюкоронидазы (GUS), хлорамфениколацетилтрансферазы (CAT), зеленого флуоресцентного белка (GFP), красного флуоресцентного белка (RFP), желтого флуоресцентного белка (YFP), голубого флуоресцентного белка (CFP), лазурного флуоресцентного белка, цитринового флуоресцентного белка, оранжевого флуоресцентного белка, вишневого флуоресцентного белка, бирюзового флуоресцентного белка, синего флуоресцентного белка, пероксидазы хрена (HRP), люциферазы (Luc), нопалинсинтазы (NOS), октопинсинтазы (OCS), люциферазы и их производных. Способы определения модуляции репортерного гена хорошо известны в настоящей области техники и включают без ограничения флуорометрические способы (например, флуоресцентную спектроскопию, сортировку флуоресцентно-активированных клеток (FACS), флуоресцентную микроскопию) и определение устойчивости к антибиотикам.

[0089] Заболевания и нарушения

[0090] Настоящее изобретение относится к библиотекам связывающих GPCR доменов, содержащим нуклеиновые кислоты, кодирующие остовы, содержащие связывающие GPCR домены, которые могут оказывать терапевтические эффекты. В соответствии с некоторыми вариантами, библиотеки связывающих GPCR доменов при трансляции приводят к получению белка, который применяют для лечения заболевания или нарушения. В соответствии с некоторыми вариантами, белок представляет собой иммуноглобулин. В соответствии с некоторыми вариантами, белок представляет собой пептидомиметик. К иллюстративным заболеваниям относятся без ограничения рак, воспалительные заболевания или нарушения, метаболическое заболевание или нарушение, заболевание или нарушение сердечнососудистой системы, заболевание или нарушение респираторной системы, боль, заболевание или нарушение пищеварительной системы, заболевание или нарушение репродуктивной системы, заболевание или нарушение эндокринной системы или неврологическое заболевание или нарушение. В соответствии с некоторыми вариантами, рак является солидным раком или гематологическим раком. В соответствии с некоторыми вариантами, ингибитор рецептора глюкагоноподобного пептида 1 (GLP1R), GPCR, который описан в настоящем документе, применяют для

лечения метаболического нарушения. В соответствии с некоторыми вариантами, описываемый в настоящем документе ингибитор GPCR, т. е. GLP1R, применяют для лечения увеличения массы тела (или для индукции потери массы), лечения ожирения или лечения сахарного диабета II типа. В соответствии с некоторыми вариантами, субъектом является млекопитающее. В соответствии с некоторыми вариантами, субъектом является мышь, кролик, собака или человек. Субъектами, подвергаемыми лечению с помощью описываемых в настоящем документе способов, могут быть младенцы, взрослые или дети. Фармацевтические композиции, содержащие описываемые в настоящем документе антитела или фрагменты антител, можно вводить внутривенно или подкожно. В соответствии с некоторыми вариантами, фармацевтическая композиция содержит описываемое в настоящем документе антитело или фрагмент антитела, содержащее CDR-НЗ, содержащий последовательность под любым из SEQ ID NO: 2420-2436. В соответствии с дополнительными вариантами, фармацевтическую композицию применяют для лечения метаболического нарушения.

[0091] Библиотеки вариантов

[0092] *Вариант кодона*

[0093] Описываемые в настоящем документе библиотеки вариантов нуклеиновых кислот могут содержать множество нуклеиновых кислот, причем каждая нуклеиновая кислота кодирует последовательность варианта кодона по сравнению с эталонной последовательностью нуклеиновой кислоты. В соответствии с некоторыми вариантами, каждая нуклеиновая кислота первой совокупности нуклеиновых кислот содержит вариант в отдельном сайте варианта. В соответствии с некоторыми вариантами, первая совокупность нуклеиновых кислот содержит множество вариантов отдельного сайта варианта, так чтобы первая совокупность нуклеиновых кислот содержала более одного варианта в одном и том же сайте варианта. Первая совокупность нуклеиновых кислот может содержать нуклеиновые кислоты, совместно кодирующие множество вариантов кодона в одном и том же сайте варианта. Первая совокупность нуклеиновых кислот может содержать нуклеиновые кислоты, совместно кодирующие до 19 или более кодонов в одном и том же положении. Первая совокупность нуклеиновых кислот может содержать нуклеиновые кислоты, совместно кодирующие до 60 вариантов триплетов в одном и том же положении, или первая совокупность нуклеиновых кислот может содержать нуклеиновые кислоты, совместно кодирующие до 61 различного триплета кодонов в одном и том же положении. Каждый вариант может кодировать кодон, который в процессе трансляции дает иную

аминокислоту. В **таблице 3** приведен перечень каждого возможного кодона (и типичной аминокислоты) для сайта варианта.

Таблица 3. Перечень кодонов и аминокислот

Аминокислоты	Однобуквенный код	Трехбуквенный код	Кодоны					
Аланин	A	Ala	GC A	GC C	GC G	GCT		
Цистеин	C	Cys	TG C	TGT				
Аспарагиновая кислота	D	Asp	GA C	GAT				
Глутаминовая кислота	E	Glu	GA A	GAG				
Фенилаланин	F	Phe	TTC	TTT				
Глицин	G	Gly	GG A	GG C	GG G	GGT		
Гистидин	H	His	CA C	CAT				
Изолейцин	I	Iso	AT A	AT C	ATT			
Лизин	K	Lys	AA A	AAG				
Лейцин	L	Leu	TT A	TT G	CT A	CT C	CT G	CT T
Метионин	M	Met	ATG					
Аспарагин	N	Asn	AA C	AAT				
Пролин	P	Pro	CC A	CC C	CC G	CCT		
Глутамин	Q	Gln	CA A	CAG				
Аргинин	R	Arg	AG A	AG G	CG A	CG C	CG G	CG T

Серин	S	Ser	AG C	AG T	TC A	TC C	TC G	TC T
Треонин	T	Thr	AC A	AC C	AC G	ACT		
Валин	V	Val	GT A	GT C	GT G	GTT		
Триптофан	W	Trp	TGG					
Тирозин	Y	Tyr	TA C	TAT				

[0094] Совокупность нуклеиновых кислот может содержать различные нуклеиновые кислоты, ковалентно кодирующие вплоть до 20 вариантов кодонов в нескольких положениях. В таких случаях каждая нуклеиновая кислота в данной совокупности содержит варианты для кодонов в нескольких положениях в одной и той же нуклеиновой кислоте. В соответствии с некоторыми вариантами, каждая нуклеиновая кислота в совокупности содержит варианты кодонов в 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 или более кодонов в отдельной нуклеиновой кислоте. В соответствии с некоторыми вариантами, каждый вариант длинной нуклеиновой кислоты содержит вариант кодонов в 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 или более кодонов в отдельной длинной нуклеиновой кислоте. В соответствии с некоторыми вариантами, совокупность вариантов нуклеиновых кислот содержит вариант кодонов в 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 или более кодонов в отдельной нуклеиновой кислоте. В соответствии с некоторыми вариантами, совокупность вариантов нуклеиновых кислот содержит вариант кодонов по меньшей мере приблизительно в 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 или более кодонов в отдельной длинной нуклеиновой кислоте.

[0095] Высокопараллельный синтез нуклеиновых кислот

[0096] В настоящем документе представлен платформенный подход с использованием миниатюризации, параллелизации и вертикальной интеграции непрерывного процесса от синтеза полинуклеотидов до сборки генов в нанолунках на силиконе для создания революционной платформы для синтеза. Описываемые в настоящем документе устройства обеспечивают на такой же площадке, что и 96-луночный планшет, силиконовую платформу для синтеза, способную увеличивать пропускную способность до 1000 или более раз по сравнению с традиционными способами синтеза, с получением до

приблизительно 1000000 или более полинуклеотидов или 10000 или более генов за отдельный высокопараллелизованный цикл.

[0097] С появлением секвенирования следующего поколения геномные данные высокого разрешения стали важным фактором для исследований, которые углубляются в биологическую роль различных генов как при нормальной биологии, так и при патогенезе заболевания. В основе этого исследования лежит центральная догма молекулярной биологии и концепция «последовательной передачи информации последовательности». Геномная информация, закодированная в ДНК, транскрибируется в матричную РНК, которая затем транслируется в белок, который является активным продуктом в заданном биологическом пути.

[0098] Другой интересной областью исследования является поиск, разработка и производство терапевтических молекул, ориентированных на высокоспецифичную клеточную мишень. В основе процесса разработки целенаправленной терапии лежат высокоразнообразные библиотеки последовательностей ДНК. Генных мутантов применяют для экспрессии белков в цикле создания, построения и конструирования тестируемого белка, результатом чего в идеале является оптимизированный ген для высокоэффективной экспрессии белка с высокой аффинностью к его терапевтической мишени. В качестве примера рассмотрим связывающий карман рецептора. Возможность одновременного тестирования всех пермутаций последовательностей всех остатков в связывающем кармане позволит провести тщательное исследование, что увеличивает шансы на успех. Одним из подходов к такой проблеме разработки является насыщающий мутагенез, при котором исследователь пытается создать все возможные мутации в определенном сайте в рецепторе. Несмотря на то, что это затратно по финансам и времени, а также трудоемко, данный подход позволяет внедрять каждый вариант в каждое положение. В отличие от этого, комбинаторный мутагенез, при котором несколько выбранных положений или короткая цепь ДНК могут быть значительно изменены, позволяет получить лишь неполный репертуар вариантов со смещенным представлением.

[0099] Для ускорения процесса разработки лекарственных средств библиотека с требуемыми вариантами, доступными с требуемой частотой в правильном положении, доступном для тестирования, иными словами, точная библиотека позволяет снизить затраты, а также сократить время скрининга. Настоящее изобретение относится к способам синтеза библиотек синтетических вариантов нуклеиновых кислот, которые обеспечивают точное введение каждого предполагаемого варианта с требуемой частотой. Для конечного пользователя это означает не только возможность тщательно выделить пространство последовательности, но и возможность эффективно проверять эти гипотезы со снижением

стоимости и времени проверки. Общегеномное редактирование может прояснить важные пути, библиотеки, в которых можно протестировать каждый вариант и пермутацию последовательности на оптимальную функциональность и можно использовать тысячи генов для реконструкции целых путей и геномов с целью повторного конструирования биологических систем для поиска лекарственных средств.

[00100] В первом примере с помощью описываемых в настоящем документе способов можно оптимизировать само лекарственное средство. Например, для улучшения указанной функции антитела конструируют и синтезируют библиотеку вариантов полинуклеотидов, кодирующих часть антитела. Затем с помощью описываемых в настоящем документе способов (например, ПЦР-мутагенезом с последующей вставкой в вектор) можно создать библиотеку вариантов нуклеиновых кислот для антитела. Затем экспрессируют антитело в линии клеток-продуцентов и подвергают его скринингу на повышенную активность. Примеры скрининга включают изучение модуляции аффинности связывания с антигеном, стабильности или эффекторной функции (например, ADCC, функцию комплемента или апоптоза). К иллюстративным участкам для оптимизации антитела относятся без ограничения Fc-участок, Fab-участок, переменный участок Fab-участка, константный участок Fab-участка, переменный домен тяжелой цепи или легкой цепи (V_H или V_L) и специфические определяющие комплементарность участки (CDR) V_H или V_L .

[00101] Библиотеки нуклеиновых кислот, синтезированные с помощью описываемых в настоящем документе способов, могут экспрессироваться в различных клетках, связанных с болезненным состоянием. К связанным с болезненным состоянием клеткам относятся клеточные линии, образцы тканей, первичные клетки от субъекта, культивируемые клетки, выделенные из субъекта, или клетки в модельной системе. К иллюстративным модельным системам относятся без ограничения растительные и животные модели болезненного состояния.

[00102] Для выявления варианта молекулы, связанного с предупреждением, снижением или лечением болезненного состояния, описываемую в настоящем документе библиотеку вариантов нуклеиновых кислот экспрессируют в клетке, связанной с болезненным состоянием, или в клетке, в которой может быть индуцировано болезненное состояние. В соответствии с некоторыми вариантами, для индукции болезненного состояния в клетках применяют определенное средство. К иллюстративным инструментам для индукции болезненного состояния относятся без ограничения система рекомбинации Cre/Lox, индукция воспаления LPS и стрептозотоцин для индукции гипогликемии. Связанные с болезненным состоянием клетки могут представлять собой клетки из

модельной системы или культивируемые клетки, а также клетки от имеющего конкретное болезненное патологическое состояние субъекта. К иллюстративным болезненным патологическим состояниям относится бактериальное, грибковое, вирусное, аутоиммунное или пролиферативное нарушение (например, рак). В соответствии с некоторыми вариантами, библиотеку вариантов нуклеиновых кислот экспрессируют в модельной системе, клеточной линии или первичных клетках, полученных от субъекта, и подвергают скринингу на наличие изменений по меньшей мере в одной клеточной активности. К иллюстративным клеточным активностям относятся без ограничения пролиферация, прохождение клеточного цикла, гибель клеток, адгезия, миграция, репродукция, передача клеточных сигналов, выработка энергии, использование кислорода, метаболическая активность и старение, ответ на повреждение свободными радикалами или любая их комбинация.

[00103] Подложки

[00104] Устройства, применяемые в качестве поверхности для синтеза полинуклеотидов, могут быть представлены в форме подложек, которые включают без ограничения однородные матричные поверхности, структурированные матричные поверхности, каналы, микроносители, гели и тому подобное. Настоящее изобретение относится к подложкам, содержащим множество кластеров, причем каждый кластер содержит множество локусов, которые служат основой для присоединения и синтеза полинуклеотидов. В соответствии с некоторыми вариантами, подложки имеют однородную матричную поверхность. Например, однородная матричная поверхность представляет собой однородный планшет. Применяемый в контексте настоящего описания термин «локус» относится к дискретному участку на структуре, который служит основой для полинуклеотидов, кодирующих отдельную заданную последовательность, достраивая их от поверхности. В соответствии с некоторыми вариантами, локус находится на двумерной поверхности, например, практически плоской поверхности. В соответствии с некоторыми вариантами, локус находится на трехмерной поверхности, например, в лунке, микролунке, канале или стержне. В соответствии с некоторыми вариантами, поверхность локуса содержит материал, который имеет активную функциональную группу для присоединения по меньшей мере одного нуклеотида для синтеза полинуклеотидов или, предпочтительно, совокупность идентичных нуклеотидов для синтеза совокупности полинуклеотидов. В соответствии с некоторыми вариантами, полинуклеотид относится к совокупности полинуклеотидов, кодирующих одну и ту же последовательность нуклеиновой кислоты. В некоторых случаях поверхность подложки включает одну поверхность или совокупность

поверхностей подложки. Средние уровни ошибок по полинуклеотидам, синтезируемым в описанной в настоящем документе библиотеке с помощью представленных систем и способов, зачастую составляют менее 1 на 1000, менее приблизительно 1 на 2000, менее приблизительно 1 на 3000 или менее, зачастую вообще без исправления ошибок.

[00105] Настоящее изобретение относится к поверхностям, которые служат основой для параллельного синтеза множества полинуклеотидов, имеющих разные заданные последовательности в имеющихся адрес местоположениях на общей основе. В соответствии с некоторыми вариантами, подложка служит основой для синтеза более 50, 100, 200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000, 5000, 10000, 20000, 50000, 100000, 200000, 300000, 400000, 500000, 600000, 700000, 800000, 900000, 1000000, 1200000, 1400000, 1600000, 1800000, 2000000, 2500000, 3000000, 3500000, 4000000, 4500000, 5000000, 10000000 или более неидентичных полинуклеотидов. В некоторых случаях поверхности служат основой для синтеза более 50, 100, 200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000, 5000, 10000, 20000, 50000, 100000, 200000, 300000, 400000, 500000, 600000, 700000, 800000, 900000, 1000000, 1200000, 1400000, 1600000, 1800000, 2000000, 2500000, 3000000, 3500000, 4000000, 4500000, 5000000, 10000000 или более полинуклеотидов, кодирующих различные последовательности. В соответствии с некоторыми вариантами, по меньшей мере часть полинуклеотидов имеет идентичную последовательность или их можно синтезировать с идентичной последовательностью. В соответствии с некоторыми вариантами, подложка обеспечивает поверхностную среду для роста полинуклеотидов, имеющих по меньшей мере 80, 90, 100, 120, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500 оснований или более.

[00106] Настоящее изобретение относится к способам синтеза полинуклеотидов на различных локусах основы, причем каждый локус служит основой для синтеза совокупности полинуклеотидов. В некоторых случаях каждый локус служит основой для синтеза совокупности полинуклеотидов, имеющих последовательность, отличную от совокупности выращенных в другом локусе полинуклеотидов. В соответствии с некоторыми вариантами, каждая полинуклеотидная последовательность синтезируется с избыточностью 1, 2, 3, 4, 5, 6, 7, 8, 9 или более в разных локусах внутри одного и того же кластера локусов на поверхности для синтеза полинуклеотидов. В соответствии с некоторыми вариантами, локусы подложки расположены в пределах множества кластеров. В соответствии с некоторыми вариантами, подложка содержит по меньшей мере 10, 500, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 11000, 12000, 13000, 14000, 15000, 20000, 30000, 40000, 50000 или более кластеров. В соответствии с некоторыми вариантами, подложка содержит более 2000, 5000, 10000, 100000, 200000, 300000, 400000,

500000, 600000, 700000, 800000, 900000, 1000000, 1100000, 1200000, 1300000, 1400000, 1500000, 1600000, 1700000, 1800000, 1900000, 2000000, 300000, 400000, 500000, 600000, 700000, 800000, 900000, 1000000, 1200000, 1400000, 1600000, 1800000, 2000000, 2500000, 3000000, 3500000, 4000000, 4500000, 5000000 или 10000000 или более различных локусов. В соответствии с некоторыми вариантами, подложка содержит приблизительно 10000 различных локусов. Количество локусов в пределах отдельного кластера отличается в различных вариантах. В некоторых случаях каждый кластер включает 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 130, 150, 200, 300, 400, 500 или более локусов. В соответствии с некоторыми вариантами, каждый кластер включает приблизительно 50-500 локусов. В соответствии с некоторыми вариантами, каждый кластер включает приблизительно 100-200 локусов. В соответствии с некоторыми вариантами, каждый кластер включает приблизительно 100-150 локусов. В соответствии с некоторыми вариантами, каждый кластер включает приблизительно 109, 121, 130 или 137 локусов. В соответствии с некоторыми вариантами, каждый кластер включает приблизительно 19, 20, 61, 64 или более локусов. В качестве альтернативы или в сочетании, синтез полинуклеотидов происходит на однородной матричной поверхности.

[00107] В соответствии с некоторыми вариантами, количество различных полинуклеотидов, синтезируемых на подложке, зависит от количества различных локусов, доступных у подложки. В соответствии с некоторыми вариантами, плотность локусов в пределах кластера или поверхности подложки составляет по меньшей мере или приблизительно 1, 10, 25, 50, 65, 75, 100, 130, 150, 175, 200, 300, 400, 500, 1000 или более локусов на мм². В некоторых случаях подложка составляет 10-500, 25-400, 50-500, 100-500, 150-500, 10-250, 50-250, 10-200 или 50-200 мм². В соответствии с некоторыми вариантами, расстояние между центрами двух смежных локусов в пределах кластера или поверхности составляет приблизительно 10-500, приблизительно 10-200 или приблизительно 10-100 мкм. В соответствии с некоторыми вариантами, расстояние между центрами двух смежных локусов составляет более приблизительно 10, 20, 30, 40, 50, 60, 70, 80, 90 или 100 мкм. В соответствии с некоторыми вариантами, расстояние между центрами двух смежных локусов составляет менее приблизительно 200, 150, 100, 80, 70, 60, 50, 40, 30, 20 или 10 мкм. В соответствии с некоторыми вариантами, каждый локус имеет ширину приблизительно 0,5, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90 или 100 мкм. В некоторых случаях каждый локус имеет ширину приблизительно 0,5-100, 0,5-50, 10-75 или 0,5-50 мкм.

[00108] В соответствии с некоторыми вариантами, плотность кластеров в подложке составляет по меньшей мере или приблизительно 1 кластер на 100 мм², 1 кластер на 10 мм², 1 кластер на 5 мм², 1 кластер на 4 мм², 1 кластер на 3 мм², 1 кластер на 2 мм², 1

кластер на 1 мм², 2 кластера на 1 мм², 3 кластера на 1 мм², 4 кластера на 1 мм², 5 кластеров на 1 мм², 10 кластеров на 1 мм², 50 кластеров на 1 мм² или более. В соответствии с некоторыми вариантами, подложка содержит от приблизительно 1 кластера на 10 мм² до приблизительно 10 кластеров на 1 мм². В соответствии с некоторыми вариантами, расстояние между центрами двух смежных кластеров составляет по меньшей мере или приблизительно 50, 100, 200, 500, 1000, 2000 или 5000 мкм. В некоторых случаях расстояние между центрами двух смежных кластеров составляет приблизительно 50-100, 50-200, 50-300, 50-500 и 100-2000 мкм. В некоторых случаях расстояние между центрами двух смежных кластеров составляет приблизительно 0,05-50, 0,05-10, 0,05-5, 0,05-4, 0,05-3, 0,05-2, 0,1-10, 0,2-10, 0,3-10, 0,4-10, 0,5-10, 0,5-5 или 0,5-2 мм. В некоторых случаях каждый кластер имеет поперечное сечение, которое составляет от приблизительно 0,5 до приблизительно 2, от приблизительно 0,5 до приблизительно 1 или от приблизительно 1 до приблизительно 2 мм. В некоторых случаях каждый кластер имеет поперечное сечение, которое составляет приблизительно 0,5, 0,6, 0,7, 0,8, 0,9, 1, 1,1, 1,2, 1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9 или 2 мм. В некоторых случаях каждый кластер имеет внутреннее поперечное сечение, которое составляет приблизительно 0,5, 0,6, 0,7, 0,8, 0,9, 1, 1,1, 1,15, 1,2, 1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9 или 2 мм.

[00109] В соответствии с некоторыми вариантами, подложка имеет приблизительно размер стандартного 96-луночного планшета, например, приблизительно 100 - приблизительно 200 мм на приблизительно 50 - приблизительно 150 мм. В соответствии с некоторыми вариантами, подложка имеет диаметр менее или равный приблизительно 1000, 500, 450, 400, 300, 250, 200, 150, 100 или 50 мм. В соответствии с некоторыми вариантами, диаметр подложки составляет приблизительно 25-1000, 25-800, 25-600, 25-500, 25-400, 25-300 или 25-200 мм. В соответствии с некоторыми вариантами, подложка имеет область плоской поверхности, которая составляет по меньшей мере приблизительно 100, 200, 500, 1000, 2000, 5000, 10000, 12000, 15000, 20000, 30000, 40000, 50000 мм² или более. В соответствии с некоторыми вариантами, толщина подложки составляет приблизительно 50-2000, 50-1000, 100-1000, 200-1000 или 250-1000 мм.

[00110] *Материалы поверхности*

[00111] Представленные в настоящем документе подложки, устройства и реакторы изготовлены из любого разнообразия материалов, подходящих для описываемых в настоящем документе способов, композиций и систем. В соответствии с определенными вариантами, материалы подложки изготавливают такими, чтобы они характеризовались низким уровнем связывания нуклеотидов. В соответствии с некоторыми вариантами,

материалы для подложки модифицируют для получения различных поверхностей, которые характеризуются высоким уровнем связывания нуклеотидов. В соответствии с некоторыми вариантами, материалы для подложки являются прозрачными для видимого света и/или УФ света. В соответствии с некоторыми вариантами, материалы для подложки являются достаточно проводящими, например, способными формировать однородные электрические поля по всей или части подложки. В соответствии с некоторыми вариантами, проводящие материалы соединены с электрическим заземлением. В соответствии с некоторыми вариантами, подложка является теплопроводной или изолированной. В соответствии с некоторыми вариантами, материалы являются химически стойкими и жаростойкими для поддержания химических или биохимических реакций, например, процессов реакции синтеза полинуклеотидов. В соответствии с некоторыми вариантами, подложка содержит гибкие материалы. В случае гибких материалов материалы могут включать без ограничения нейлон, как модифицированную, так и немодифицированную нитроцеллюлозу, полипропилен и тому подобное. В соответствии с некоторыми вариантами, подложка содержит жесткие материалы. В случае жестких материалов материалы могут включать без ограничения стекло, плавленный кварц, силикон, пластмассы (например, политетрафторэтилен, полипропилен, полистирол, поликарбонат, а также их смеси и тому подобное), металлы (например, золото, платину и тому подобное). Основа, твердая основа или реакторы могут быть изготовлены из материала, выбранного из группы, состоящей из силикона, полистирола, агарозы, декстрана, целлюлозных полимеров, полиакриламидов, полидиметилсилоксана (PDMS) и стекла. Основы/твердые основы или микроструктуры и находящиеся в них реакторы могут быть изготовлены путем объединения перечисленных в настоящем документе материалов или из любого другого подходящего материала, известного в настоящей области техники.

[00112] Архитектура поверхности

[00113] Настоящее изобретение относится к подложкам для описываемых в настоящем документе способов, композиций и систем, причем подложки имеют архитектуру поверхности, подходящую для описываемых в настоящем документе способов, композиций и систем. В соответствии с некоторыми вариантами, подложка содержит возвышающиеся и/или углубленные элементы. Одним из преимуществ наличия таких признаков является увеличение площади поверхности, служащей основой для синтеза полинуклеотидов. В соответствии с некоторыми вариантами, подложку имеющую возвышающиеся и/или углубленные элементы, называют трехмерной подложкой. В некоторых случаях трехмерная подложка содержит один или несколько каналов. В

некоторых случаях канал содержат один или несколько локусов. В некоторых случаях каналы доступны для внесения реагентов с помощью устройства для внесения, такого как устройства для внесения материала. В некоторых случаях реагенты и/или жидкости собирают в более крупную лунку, которая находится в жидкостном сообщении с одним или несколькими каналами. Например, подложка содержит множество каналов, соответствующих множеству локусов с кластером, и множество каналов находятся в жидкостном сообщении с одной лункой кластера. Согласно некоторым способам, библиотеку полинуклеотидов синтезируют в множестве локусов кластера.

[00114] Настоящее изобретение относится к подложкам для описываемых в настоящем документе способов, композиций и систем, причем подложки выполнены с возможностью проведения синтеза полинуклеотидов. В соответствии с некоторыми вариантами, структура выполнена так, чтобы она обеспечивала пути для контролируемых потоков и переноса масс для синтеза полинуклеотидов на поверхности. В соответствии с некоторыми вариантами, конфигурация подложки обеспечивает пути для контролируемого и равномерного распределения масс, периоды пребывания химического реактива и/или эффективность промывания в ходе синтеза полинуклеотидов. В соответствии с некоторыми вариантами, конфигурация подложки обеспечивает повышенную эффективность вытеснения, например, путем обеспечения достаточного объема для растущего полинуклеотида, так чтобы объем, исключаемый растущим полинуклеотидом, не занимал более 50, 45, 40, 35, 30, 25, 20, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1% или менее от первоначально доступного объема, который доступен или подходит для выращивания полинуклеотида. В соответствии с некоторыми вариантами, трехмерная структура обеспечивает управляемый поток жидкости для обеспечения быстрой замены химического реактива.

[00115] Настоящее изобретение относится к подложкам для описываемых в настоящем документе способов, композиций и систем, причем подложки содержат структуры, подходящие для описываемых в настоящем документе способов, композиций и систем. В соответствии с некоторыми вариантами, разделение осуществляют с помощью физической структуры. В соответствии с некоторыми вариантами, разделение осуществляют с помощью дифференциальной функционализации поверхности, создавая активные и пассивные участки для синтеза полинуклеотидов. В соответствии с некоторыми вариантами, дифференциальную функционализацию осуществляют путем изменения гидрофобности на поверхности подложки, создавая тем самым эффекты краевого угла с водой, которые вызывают образование капель или смачивание внесенных реагентов. Использование более крупных структур может уменьшить разбрызгивание и перекрестное

загрязнение отдельных местоположений синтеза полинуклеотидов реагентами из соседних зон. В некоторых случаях устройство, такое как устройство для внесения материала, применяют для внесения реагентов в различные местоположения синтеза полинуклеотидов. Подложки, имеющие трехмерные элементы, выполнены таким образом, чтобы обеспечивался синтез большого количества полинуклеотидов (например, более приблизительно 10000) с низким уровнем ошибок (например, менее приблизительно 1:500, 1:1000, 1:1500, 1:2000, 1:3000, 1:5000 или 1:10000). В некоторых случаях подложка содержит элементы с плотностью приблизительно или более чем приблизительно 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 300, 400 или 500 элементов на мм².

[00116] Лунка подложки может иметь такую же или другую ширину, высоту и/или объем, что и другая лунка подложки. Канал подложки может иметь такую же или другую ширину, высоту и/или объем, что и другой канал подложки. В соответствии с некоторыми вариантами, диаметр кластера, или диаметр лунки, содержащей кластер, или как первого, так и второй составляет приблизительно 0,05-50, 0,05-10, 0,05-5, 0,05-4, 0,05-3, 0,05-2, 0,05-1, 0,05-0,5, 0,05-0,1, 0,1-10, 0,2-10, 0,3-10, 0,4-10, 0,5-10, 0,5-5 или 0,5-2 мм. В соответствии с некоторыми вариантами, диаметр кластера или лунки или как первого, так и второй составляет менее или приблизительно 5, 4, 3, 2, 1, 0,5, 0,1, 0,09, 0,08, 0,07, 0,06 или 0,05 мм. В соответствии с некоторыми вариантами, диаметр кластера или лунки или как первого, так и второй составляет от приблизительно 1,0 до 1,3 мм. В соответствии с некоторыми вариантами, диаметр кластера или лунки или как первого, так и второй составляет приблизительно 1,150 мм. В соответствии с некоторыми вариантами, диаметр кластера или лунки или как первого, так и второй составляет приблизительно 0,08 мм. Диаметр кластера относится к кластерам в двухмерной или трехмерной подложке.

[00117] В соответствии с некоторыми вариантами, высота лунки составляет приблизительно 20-1000, 50-1000, 100-1000, 200-1000, 300-1000, 400-1000 или 500-1000 мкм. В некоторых случаях высота лунки составляет менее 1000, 900, 800, 700 или 600 мкм.

[00118] В соответствии с некоторыми вариантами, подложка содержит множество каналов, соответствующих множеству локусов в пределах кластера, причем высота и глубина канала составляет 5-500, 5-400, 5-300, 5-200, 5-100, 5-50 или 10-50 мкм. В некоторых случаях высота канала составляет менее 100, 80, 60, 40 или 20 мкм.

[00119] В соответствии с некоторыми вариантами, диаметр канала, локуса (например, у практически плоской подложки) или как канала, так и локуса (например, в трехмерной подложке, в которой локус соответствует каналу) составляет приблизительно 1-1000, 1-500, 1-200, 1-100, 5-100 или 10-100 мкм, например, приблизительно 90, 80, 70, 60,

50, 40, 30, 20 или 10 мкм. В соответствии с некоторыми вариантами, диаметр канала, локуса или как канала, так и локуса составляет менее приблизительно 100, 90, 80, 70, 60, 50, 40, 30, 20 или 10 мкм. В соответствии с некоторыми вариантами, расстояние между центрами двух смежных каналов, локусов или каналов и локусов составляет приблизительно 1-500, 1-200, 1-100, 5-200, 5-100, 5-50 или 5-30, например, приблизительно 20 мкм.

[00120] Модификации поверхности

[00121] Настоящее изобретение относится к способам синтеза полинуклеотидов на поверхности, причем поверхность имеет различные модификации поверхности. В соответствии с некоторыми вариантами, модификации поверхности используют для химического и/или физического изменения поверхности путем аддитивного или субтрактивного процесса для изменения одного или нескольких химических и/или физических свойств поверхности подложки или выбранного сайта или участка поверхности подложки. Например, модификации поверхности включают без ограничения (1) изменение свойств смачивания поверхности, (2) функционализацию поверхности, т. е. внесение, модификацию или замену поверхностных функциональных групп, (3) дефункционализацию поверхности, т. е. удаление с поверхности функциональных групп, (4) иное изменение химического состава поверхности, например, посредством травления, (5) увеличение или уменьшение шероховатости поверхности, (6) нанесение покрытия на поверхность, например, покрытия, которое обладает свойствами смачивания, которые отличаются от свойств смачивания поверхности, и/или (7) нанесение частиц на поверхность.

[00122] В некоторых случаях добавление химического слоя на поверхность (называемого усилителем адгезии) облегчает формирование структурированных рельефов локусов на поверхности подложки. Иллюстративные поверхности для нанесения усилителя адгезии включают без ограничения стекло, силикон, диоксид кремния и нитрид кремния. В некоторых случаях усилитель адгезии представляет собой химическое вещество с высокой поверхностной энергией. В соответствии с некоторыми вариантами, на поверхности подложки размещают второй химический слой. В некоторых случаях второй химический слой имеет низкую поверхностную энергию. В некоторых случаях поверхностная энергия химического слоя, нанесенного на поверхность, поддерживает локализацию капель на поверхности. В зависимости от выбранной структуры рельефа можно изменять близость локусов и/или область контакта с жидкостью в локусах.

[00123] В соответствии с некоторыми вариантами, поверхность подложки или разделенные локусы, на которые вносят нуклеиновые кислоты или другие молекулы,

например, для синтеза полинуклеотидов, являются гладкими или практически плоскими (например, двумерными) или имеют неровности, такие как возвышающиеся или углубленные элементы (например, трехмерные элементы). В соответствии с некоторыми вариантами, поверхность подложки модифицируют одним или несколькими различными слоями соединений. Такие представляющие интерес модифицирующие слои включают без ограничения неорганические и органические слои, такие как металлы, оксиды металлов, полимеры, малые органические молекулы и тому подобное.

[00124] В соответствии с некоторыми вариантами, разделенные локусы подложки функционализируют одной или несколькими группами, которые увеличивают и/или уменьшают поверхностную энергию. В некоторых случаях группа является химически инертной. В некоторых случаях группа может поддерживать требуемую химическую реакцию, например, один или нескольких процессов в реакции синтеза полинуклеотидов. Поверхностная энергия, или гидрофобность, поверхности является фактором, определяющим склонность нуклеотида прикрепиться к поверхности. В соответствии с некоторыми вариантами, способ функционализации подложки предусматривает: (а) обеспечение наличия подложки, имеющей поверхность, которая содержит диоксид кремния; и (b) силанизацию поверхности с помощью подходящего описанного в настоящем документе силанизирующего средства или иного известного из уровня техники средства, например, молекулы органофункционального алкоксисилана. Способы и функционализирующие средства описаны в патенте США № 5474796, который настоящим включен посредством ссылки в полном его объеме.

[00125] В соответствии с некоторыми вариантами, поверхность подложки функционализируют путем контакта с дериватизирующей композицией, которая содержит смесь силанов, в условиях протекания реакции, эффективных для связывания силанов с поверхностью подложки, обычно через реакционноспособные гидрофильные группы, присутствующие на поверхности подложки. Силанизация обычно охватывает поверхность путем самосборки с молекулами органофункционального алкоксисилана. Дополнительно можно применять различные силоксановые функционализирующие реагенты, на данный момент известные в настоящей области техники, например, для снижения или увеличения поверхностной энергии. Органофункциональные алкоксисиланы классифицируют в соответствии с их органическими функциональными группами.

[00126] Синтез полинуклеотидов

[00127] Способы синтеза полинуклеотидов по настоящему раскрытию могут предусматривать способы, предусматривающие использование фосфорамидитной химии.

В соответствии с некоторыми вариантами, синтез полинуклеотидов предусматривает связывание основания с фосфорамидитом. Синтез полинуклеотидов может предусматривать связывание основания путем внесения фосфорамидита в условиях связывания, причем одно и то же основание необязательно вносится с фосфорамидитом более одного раза, т. е. двойное связывание. Синтез полинуклеотидов может предусматривать кэппирование непрореагировавших сайтов. В соответствии с некоторыми вариантами, кэппирование является необязательным. Синтез полинуклеотидов также может предусматривать окисление, или стадию окисления, или стадии окисления. Синтез полинуклеотидов может предусматривать деблокирование, детритилирование и сульфуризацию. В соответствии с некоторыми вариантами, синтез полинуклеотидов предусматривает окисление или сульфуризацию. В соответствии с некоторыми вариантами, между одной или каждой стадией во время реакции синтеза полинуклеотидов устройство промывают, например, с помощью тетразола или ацетонитрила. Временные рамки для любой стадии в способе синтеза с применением фосфорамидита могут составлять менее приблизительно 2 мин, 1 мин, 50 с, 40 с, 30 с, 20 с и 10 с.

[00128] Синтез полинуклеотидов с помощью фосфорамидитного способа может предусматривать последующее добавление фосфорамидитного строительного блока (например, нуклеозидного фосфорамидита) к растущей полинуклеотидной цепи для образования фосфитнотриэфирной связи. Фосфорамидитный синтез полинуклеотидов происходит в 3' - 5' направлении. Фосфорамидитный синтез полинуклеотидов предусматривает контролируемое добавление одного нуклеотида к растущей цепи нуклеиновой кислоты за цикл синтеза. В соответствии с некоторыми вариантами, каждый цикл синтеза предусматривает стадию связывания. Фосфорамидитное связывание включает образование фосфитнотриэфирной связи между активированным нуклеозидным фосфорамидитом и нуклеозидом, связанным с подложкой, например, с помощью линкера. В соответствии с некоторыми вариантами, нуклеозидный фосфорамидит подают в устройство активированным. В соответствии с некоторыми вариантами, нуклеозидный фосфорамидит подают в устройство с активатором. В соответствии с некоторыми вариантами, нуклеозидные фосфорамидиты подают в устройство с 1,5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100-кратным избытком или более по сравнению со связанными с подложкой нуклеозидами. В соответствии с некоторыми вариантами, добавление нуклеозидного фосфорамидита проводят в безводной среде, например, в безводном ацетонитриле. После добавления нуклеозидного фосфорамидита устройство необязательно промывают. В соответствии с некоторыми вариантами, стадию связывания повторяют один или несколько дополнительных раз,

необязательно со стадией промывки между добавлениями нуклеозидного фосфорамидита на подложку. В соответствии с некоторыми вариантами, применяемый в настоящем документе способ синтеза полинуклеотидов предусматривает 1, 2, 3 или более последовательных стадий связывания. Перед связыванием во многих случаях у связанного с устройством нуклеозида удаляют защитную группу, причем функция защитной группы заключается в предотвращении полимеризации. Стандартной защитной группой является 4,4'-диметокситритил (DMT).

[00129] После связывания фосфорамидитные способы синтеза полинуклеотидов необязательно предусматривают стадию кэппирования. На стадии кэппирования растущий полинуклеотид обрабатывают кэппирующим средством. Стадия кэппирования полезна для блокировки непрореагировавших связанных с подложкой групп 5'-ОН после связывания с дальнейшим удлинением цепи, предупреждая образование полинуклеотидов с внутренними делециями оснований. Кроме того, фосфорамидиты, активированные 1Н-тетразолом, могут в незначительной степени вступать в реакцию с Об-положением у гуанозина. Без привязки к какой-либо теории, при окислении посредством I₂/воды этот побочный продукт, возможно путем миграции с Об на N7, может подвергаться депуринизации. Апуриновые сайты могут, в конечном итоге, расщепляться в ходе окончательного удаления защитной группы с полинуклеотида, таким образом снижая выход полноразмерного продукта. Модификации на Об можно удалить обработкой кэппирующим реагентом перед окислением посредством I₂/воды. В соответствии с некоторыми вариантами, включение стадии кэппирования во время синтеза полинуклеотидов снижает уровень ошибок по сравнению с синтезом без кэппирования. В качестве примера, стадия кэппирования предусматривает обработку связанного с подложкой полинуклеотида смесью уксусного ангидрида и 1-метилимидазола. После стадии кэппирования устройство необязательно промывают.

[00130] В соответствии с некоторыми вариантами, после добавления нуклеозидного фосфорамидита и, необязательно, после кэппирования и одной или нескольких стадий промывки, связанную с устройством растущую нуклеиновую кислоту подвергают окислению. Стадия окисления предусматривает, что триэфир фосфинистой кислоты окисляют до тетракоординированного триэфира фосфорной кислоты, защищенного предшественника природной фосфодизэфирной межнуклеозидной связи. В соответствии с некоторыми вариантами, окисление растущего полинуклеотида осуществляют обработкой йодом и водой, необязательно в присутствии слабого основания (например, пиридина, лутидина, коллидина). Окисление можно проводить в безводных условиях с применением, например, трет-бутилгидропероксида или (1S)-(+)-(10-

камфорсульфонил)-оксазиридина (CSO). Согласно некоторым способам после окисления проводят стадию кэппирования. Вторая стадия кэппирования предусматривает сушку устройства, поскольку полученная в ходе окисления остаточная вода, которая может сохраняться, может подавлять последующее связывание. После окисления устройство и растущий полинуклеотид необязательно промывают. В соответствии с некоторыми вариантами, стадию окисления заменяют стадией сульфуризации для получения полинуклеотидных фосфотиоатов, причем после сульфуризации можно проводить какие-либо стадии кэппирования. Многие реагенты способны эффективно переносить серу, включая без ограничения 3-(диметиламинометилен)амино)-3Н-1,2,4-дитиазол-3-тион, DDTT, 3Н-1,2-бензодитиол-3-он 1,1-диоксид, также известный как реагент Бокажа, и N,N,N'N'-тетраэтилтиурамдисульфид (TETD).

[00131] Для прохождения последующего цикла включения нуклеозидов путем связывания связанный с защитной группой 5'-конец связанного с устройством растущего полинуклеотида удаляли, так чтобы первичная гидроксильная группа могла прореагировать со следующим нуклеозидным фосфорамидитом. В соответствии с некоторыми вариантами, защитной группой является DMT, а удаление защитной группы происходит с трихлоруксусной кислотой в дихлорметане. Проведение детритилирования в течение продолжительного времени или с более концентрированными, чем рекомендуемые, растворами кислот может привести к усилению депуринизации связанного с твердой основой полинуклеотида и, как следствие, снижает выход целевого полноразмерного продукта. Описываемые в настоящем документе способы и композиции по настоящему раскрытию обеспечивают условия контролируемого деблокирования, ограничивающие нежелательные реакции депуринизации. В соответствии с некоторыми вариантами, связанный с устройством полинуклеотид после деблокирования промывают. В соответствии с некоторыми вариантами, эффективное промывание после деблокирования способствует достижению низкого уровня ошибок у синтезированных полинуклеотидов.

[00132] Способы синтеза полинуклеотидов обычно предусматривают повторяющуюся последовательность следующих стадий: внесение защищенного мономера на активно функционализированную поверхность (например, локус) для связывания либо с активированной поверхностью, либо с линкером, либо с мономером, у которого ранее была удалена защитная группа; удаление защитной группы с внесенного мономера с тем, чтобы он вступил в реакцию с последующим вносимым защищенным мономером; и внесение другого защищенного мономера для связывания. Одна или несколько промежуточных стадий предусматривают окисление или сульфуризацию. В соответствии с некоторыми

вариантами, одной или всем стадиям предшествуют одна или несколько стадий промывания или идут после них.

[00133] Способы фосфорамидитного синтеза полинуклеотидов предусматривают ряд химических стадий. В соответствии с некоторыми вариантами, одна или несколько стадий способа синтеза предусматривают циркуляцию реагентов, при этом одна или несколько стадий способа предусматривают внесение в устройство реагента, пригодного для данной стадии. Например, реагенты циклически используют путем серии стадий внесения в виде жидкости и вакуумной сушки. Для подложек, содержащих трехмерные элементы, такие как лунки, микролунки, каналы и тому подобное, реагенты необязательно пропускают через один или несколько участков устройства через лунки и/или каналы.

[00134] Описываемые в настоящем документе способы и системы относятся к устройствам для синтеза полинуклеотидов с целью синтеза полинуклеотидов. Синтез может быть параллельным. Например, параллельно можно синтезировать по меньшей мере или приблизительно по меньшей мере 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 1000, 10000, 50000, 75000, 100000 или более полинуклеотидов. Общее количество полинуклеотидов, которые можно параллельно синтезировать, может составлять 2-100000, 3-50000, 4-10000, 5-1000, 6-900, 7-850, 8-800, 9-750, 10-700, 11-650, 12-600, 13-550, 14-500, 15-450, 16-400, 17-350, 18-300, 19-250, 20-200, 21-150, 22-100, 23-50, 24-45, 25-40, 30-35. Специалистам в настоящей области техники будет понятно, что общее количество параллельно синтезируемых полинуклеотидов может попадать в любой диапазон, ограниченный любым из этих значений, например, 25-100. Общее количество параллельно синтезируемых полинуклеотидов может попадать в любой диапазон, определяемый любым из значений, служащих конечными точками такого диапазона. Общая молярная масса полинуклеотидов, синтезированных в устройстве, или молярная масса каждого из полинуклеотидов может составлять по меньшей мере или приблизительно 10, 20, 30, 40, 50, 100, 250, 500, 750, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 25000, 50000, 75000, 100000 пикомоль или более. Длина каждого из полинуклеотидов или средняя длина полинуклеотидов в устройстве может составлять по меньшей мере или приблизительно по меньшей мере 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, 150, 200, 300, 400, 500 нуклеотидов или более. Длина каждого из полинуклеотидов или средняя длина полинуклеотидов в устройстве может составлять не более или приблизительно не более 500, 400, 300, 200, 150, 100, 50, 45, 35, 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10 нуклеотидов или менее. Длина каждого из полинуклеотидов или средняя

длина полинуклеотидов в устройстве может попадать в диапазон 10-500, 9-400, 11-300, 12-200, 13-150, 14-100, 15-50, 16-45, 17-40, 18-35, 19-25. Специалистам в настоящей области будет понятно, что длина каждого из полинуклеотидов или средняя длина полинуклеотидов в устройстве может попадать в любой диапазон, ограниченной любым из этих значений, например, 100-300. Длина каждого из полинуклеотидов или средняя длина полинуклеотидов в устройстве может попадать в любой диапазон, определяемый любым из значений, служащих в качестве конечных точек диапазона.

[00135] Представленные в настоящем документе способы синтеза полинуклеотидов на поверхности позволяют производить синтез с высокой скоростью. В качестве примера, за час синтезируют по меньшей мере 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 70, 80, 90, 100, 125, 150, 175, 200 нуклеотидов или более. К нуклеотидам относятся адениновые, гуаниновые, тиминные, цитозинные, уридиновые строительные блоки или их аналоги/модифицированные варианты. В соответствии с некоторыми вариантами, библиотеки полинуклеотидов параллельно синтезируют на подложке. Например, устройство, содержащее приблизительно или по меньшей мере приблизительно 100, 1000, 10000, 30000, 75000, 100000, 1000000, 2000000, 3000000, 4000000 или 5000000 отдельных локусов, может поддерживать синтез по меньшей мере того же числа различных полинуклеотидов, причем полинуклеотид, кодирующий отличную последовательность, синтезируют на отдельном локусе. В соответствии с некоторыми вариантами, библиотеку полинуклеотидов синтезируют на устройстве с описанным в настоящем документе низким уровнем ошибок за менее чем за три месяца, два месяца, один месяц, три недели, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 дня, 24 часа или менее. В соответствии с некоторыми вариантами, более крупные нуклеиновые кислоты, собираемые из библиотеки полинуклеотидов, синтезируемой с низкой частотой ошибок с применением описываемых в настоящем документе субстратов и способов, получают менее чем за три месяца, два месяца, один месяц, три недели, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2 дня, 24 часа или менее.

[00136] В соответствии с некоторыми вариантами, описанные в настоящем документе способы предусматривают создание библиотеки нуклеиновых кислот, содержащих варианты нуклеиновых кислот, различающиеся по множеству сайтов кодонов. В соответствии с некоторыми вариантами, нуклеиновая кислота может иметь 1 сайт, 2 сайта, 3 сайта, 4 сайта, 5 сайтов, 6 сайтов, 7 сайтов, 8 сайтов, 9 сайтов, 10 сайтов, 11 сайтов, 12 сайтов, 13 сайтов, 14 сайтов, 15 сайтов, 16 сайтов, 17 сайтов, 18 сайтов, 19 сайтов, 20 сайтов, 30 сайтов, 40 сайтов, 50 сайтов или более вариантов сайтов кодонов.

[00137] В соответствии с некоторыми вариантами, один или несколько сайтов вариантов сайтов кодонов могут быть смежными. В соответствии с некоторыми вариантами, один или несколько сайтов вариантов кодонов могут не быть смежными и могут быть разделены 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 или более кодонами.

[00138] В соответствии с некоторыми вариантами, нуклеиновая кислота может содержать множество вариантов сайтов кодонов, причем все варианты сайтов кодонов являются смежными друг с другом, образуя пространство вариантов сайтов кодонов. В соответствии с некоторыми вариантами, нуклеиновая кислота может содержать множество вариантов сайтов кодонов, причем ни один из вариантов сайтов кодонов не является смежными друг с другом. В соответствии с некоторыми вариантами, нуклеиновая кислота может содержать множество вариантов сайтов кодонов, причем некоторые варианты сайтов кодонов являются смежными друг с другом, образуя пространство вариантов сайтов кодонов, а некоторые варианты сайтов кодонов не являются смежными друг с другом.

[00139] Обратимся теперь к фигурам, на **фиг. 7** проиллюстрирован пример технологического процесса синтеза нуклеиновых кислот (например, генов) из более коротких нуклеиновых кислот. Технологический процесс обычно разделен на фазы: (1) синтез de novo библиотеки одноцепочечных нуклеиновых кислот, (2) соединение нуклеиновых кислот с образованием более крупных фрагментов, (3) исправление ошибок, (4) контроль качества и (5) отгрузка. Перед синтезом de novo предварительно отбирают предполагаемую последовательность нуклеиновой кислоты или группу последовательностей нуклеиновой кислоты. Например, для создания предварительно отбирают группу генов.

[00140] После того, как были отобраны для создания большие нуклеиновые кислоты, для синтеза de novo конструируют заданную библиотеку нуклеиновых кислот. Известны различные подходящие способы создания полинуклеотидных матриц высокой плотности. В примере технологического процесса представлен поверхностный слой устройства. В этом примере для улучшения процесса синтеза полинуклеотидов изменен химический состав поверхности. Для отталкивания жидкости создают области с низкой поверхностной энергией, тогда как для притягивания жидкостей создают области с высокой поверхностной энергией. Сама поверхность может иметь форму плоской поверхности или иметь отклонения по форме, такие как выступы или микролунки, которые увеличивают площадь поверхности. В примере технологического процесса выбранные молекулы с высокой поверхностной энергией выполняют двойную функцию поддержания химической структуры ДНК, как раскрыто в публикации международной патентной заявки

WO/2015/021080, которая полностью включена в настоящий документ посредством ссылки.

[00141] Получение полинуклеотидных матриц *in situ* производят на твердой основе и в нем используют процесс наращивания по одному нуклеотиду для параллельного удлинения нескольких олигомеров. Устройство внесения, такое как устройство внесения материала, выполнено с возможностью поэтапного высвобождения реагентов таким образом, чтобы множественные полинуклеотиды параллельно удлинялись на один остаток за раз, образуя олигомеры с заданной последовательностью **702** нуклеиновой кислоты. В соответствии с некоторыми вариантами, на этой стадии полинуклеотиды отщепляют от поверхности. Отщепление предусматривает отщепление газом, например, аммиаком или метиламином.

[00142] Созданные библиотеки полинуклеотидов помещают в реакционную камеру. В таком иллюстративном технологическом процессе реакционная камера (также называемая «нанореактором») представляет собой лунку с силиконовым покрытием, содержащую реагенты для ПЦР и опущенную на библиотеку **703** полинуклеотидов. До или после герметизации **704** полинуклеотидов добавляют реагент для высвобождения полинуклеотидов из подложки. В иллюстративном технологическом процессе полинуклеотиды высвобождают после герметизации нанореактора **705**. После высвобождения фрагменты одноцепочечных полинуклеотидов гибридизируются для охвата всей перспективной последовательности ДНК. Возможна частичная гибридизация **705**, поскольку каждый синтезируемый полинуклеотид сконструирован так, чтобы у него была небольшая часть, перекрывающаяся по меньшей мере с одним другим полинуклеотидом в пуле.

[00143] После гибридизации начинают реакцию РСА. Во время циклов с полимеразой полинуклеотиды гибридизируются с комплементарными фрагментами, а пробелы, заполняются полимеразой. Каждый цикл случайным образом увеличивает длину различных фрагментов в зависимости от того, какие полинуклеотиды находят друг друга. Комплементарность между фрагментами позволяет сформировать полный большой диапазон двухцепочечных ДНК **706**.

[00144] После завершения РСА нанореактор отделяют от устройства **707** и размещают для взаимодействия с устройством, имеющим праймеры для ПЦР **708**. После герметизации нанореактор подвергают ПЦР **709**, и амплифицируются более крупные нуклеиновые кислоты. После ПЦР **710** нанокамеру открывают **711**, добавляют **712** реагенты для исправления ошибок, герметизируют **713** камеру, и происходит реакция исправления ошибок для удаления несовпадающих пар оснований и/или цепей с плохой

комплементарностью из двухцепочечных продуктов 714 ПЦР-амплификации. Нанореактор открывают и отделяют 715. Продукт с исправленными ошибками затем подвергают дополнительным стадиям обработки, таким как ПЦР и молекулярное штрихкодирование, и затем упаковывают 722 для отгрузки 723.

[00145] В соответствии с некоторыми вариантами, принимают меры по контролю качества. После исправления ошибок стадии контроля качества предусматривают, например, взаимодействие с пластиной, имеющей праймеры для секвенирования, для амплификации продукта 716 с исправленными ошибками, герметизацию пластиной камеры, содержащей продукт 717 амплификации с исправленными ошибками, и проведение дополнительного раунда амплификации 718. Нанореактор открывают 719, а продукты объединяют 720 в пул и секвенируют 721. После определения приемлемого контроля качества упакованный продукт 722 одобрен для отгрузки 723.

[00146] В соответствии с некоторыми вариантами, нуклеиновую кислоту, созданную с помощью такого технологического процесса, как на **фиг. 7**, подвергают мутагенезу с помощью раскрытых в настоящем документе перекрывающихся праймеров. В соответствии с некоторыми вариантами, библиотеку праймеров создают путем получения *in situ* на твердой основе и используют процесс наращивания по одному нуклеотиду для параллельного удлинения нескольких олигомеров. Устройство внесения, такое как устройство внесения материала, выполнено с возможностью поэтапного высвобождения реагентов таким образом, чтобы множественные полинуклеотиды параллельно удлинялись на один остаток за раз, образуя олигомеры с заданной последовательностью 702 нуклеиновой кислоты.

[00147] Компьютерные системы

[00148] Любая из описываемых в настоящем документе систем может быть функционально связана с компьютером и может быть либо локально, либо удаленно автоматизирована с помощью компьютера. В различных вариантах способы и системы по настоящему раскрытию могут дополнительно предусматривать программное обеспечение для компьютерных систем и их применение. Соответственно, под объем настоящего раскрытия подпадает компьютеризированное управление по синхронизации функций внесения дозы/вакуумного забора/повторного наполнения, таких как организация и синхронизация перемещения устройства для внесения материала, действие по внесению дозы и приведение в действие вакуумного насоса. Компьютерные системы можно запрограммировать на взаимодействие между заданной пользователем

последовательностью оснований и положением устройства для внесения материала для доставки соответствующих реагентов в указанные участки подложки.

[00149] Компьютерную систему **800**, проиллюстрированную на **фиг. 8**, можно представить как логическое устройство, которое может считывать инструкции с носителя **811** и/или из сетевого порта **805**, которое необязательно может быть подключено к серверу **809**, имеющему несъемный носитель **812**. Система, такая как представленная на **фиг. 8**, может включать CPU **801**, дисководы **803**, дополнительные устройства ввода, такие как клавиатура **815** и/или мышь **816**, и необязательный монитор **807**. Передачу данных можно осуществлять через указанную среду передачи данных на сервер в локальном или удаленном местоположении. К среде передачи данных можно отнести любые средства передачи и/или приема данных. Например, среда передачи данных может представлять собой сетевое соединение, беспроводное соединение или интернет-соединение. Такое соединение может предусматривать связь через интернет-пространство. Предполагается, что данные, относящиеся к настоящему раскрытию, можно передавать по таким сетям или соединениям для приема и/или просмотра участником **822**, как показано на **фиг. 8**.

[00150] **Фиг. 14** представляет собой блок-схему, иллюстрирующую первый пример архитектуры компьютерной системы **1400**, которую можно применять в сочетании с примерами по настоящему изобретению. На **фиг. 14** показано, что примерная компьютерная система может включать процессор **1402** для обработки инструкций. Неограничивающие примеры процессоров включают: процессор Intel Xeon™, процессор AMD Opteron™, 32-разрядный процессор Samsung RISC ARM 1176JZ (F)-S v1.0™, процессор ARM Cortex-A8 Samsung S5PC100™, процессор ARM Cortex-A8 Apple A4™, процессор Marvell PXA 930™ или функционально эквивалентный процессор. Для параллельной обработки можно использовать многопоточность выполнения. В соответствии с некоторыми вариантами, также можно применять несколько процессоров или процессоры с несколькими ядрами, будь то в одной компьютерной системе, в кластере или распределенных среди систем в сети, содержащей множество компьютеров, мобильных телефонов и/или карманных компьютеров.

[00151] На **фиг. 9** проиллюстрировано, что к процессору **902** может быть подключен высокоскоростной кэш **904**, или он может быть включен в него, для создания высокоскоростной памяти для инструкций или данных, которые недавно или часто используются процессором **902**. Процессор **902** соединен с северным мостом **906** с помощью процессорной шины **908**. Северный мост **906** подключен к оперативной памяти (RAM) **910** с помощью шины **912** памяти и управляет доступом процессора **902** к RAM **910**. Северный мост **906** также соединен с южным мостом **914** с помощью шины **916** чипсета.

Южный мост **914**, в свою очередь, соединен с периферийной шиной **918**. Периферийной шиной может быть, например, PCI, PCI-X, PCI Express или другая периферийная шина. Северный мост и южный мост часто называют чипсетом процессора, и они управляют передачей данных между процессором, RAM и периферийными компонентами на периферийной шине **918**. В некоторых альтернативных архитектурах функциональность северного моста может быть включена в процессор вместо использования отдельного чипа северного моста. В соответствии с некоторыми вариантами, система **900** может включать графический ускоритель **922**, присоединенный к периферийной шине **918**. Ускоритель может включать программируемую пользователем логическую матрицу (FPGA) или другое оборудование для ускорения определенной обработки. Например, ускоритель можно использовать для адаптивной реструктуризации данных или для оценки алгебраических выражений, применяемых при обработке расширенного множества.

[00152] Программное обеспечение и данные хранятся во внешнем хранилище **924** и могут быть загружены в RAM **910** и/или в кэш **904** для использования процессором. Система **900** включает операционную систему для управления системными ресурсами; неограничивающие примеры операционных систем включают: Linux, WindowsTM, MACOSTM, BlackBerry OSTTM, iOSTM и другие функционально эквивалентные операционные системы, а также прикладное программное обеспечение, работающее поверх операционной системы, для управления хранением и оптимизацией данных в соответствии с примерами по настоящему раскрытию. В этом примере система **900** также включает сетевые карты (NIC) **920** и **921**, подключенные к периферийной шине для обеспечения сетевых интерфейсов для внешнего хранилища, такого как сетевое хранилище (NAS), и других компьютерных систем, которые можно применять для распределенной параллельной обработки.

[00153] **Фиг. 10** представляет собой схему, на которой показана сеть **1000** с множеством компьютерных систем **1002a** и **1002b**, множеством сотовых телефонов и карманных компьютеров **1002c** и сетевым хранилищем (NAS) **1004a** и **1004b**. В вариантах примеров системы **1002a**, **1002b** и **1002c** могут управлять хранением данных и оптимизировать доступ к данным в случае данных, хранящихся в сетевом хранилище (NAS) **1004a** и **1004b**. Для данных можно использовать математическую модель и оценить их с помощью распределенной параллельной обработки среди компьютерных систем **1002a**, **1002b** и карманных компьютеров **1002c**. Компьютерные системы **1002a** и **1002b** и сотовые телефоны, а также карманные компьютеры **1002c** также могут обеспечивать параллельную обработку для адаптивной реструктуризации данных, хранящихся в сетевом хранилище (NAS) **1004a** и **1004b**. На **фиг. 10** проиллюстрирован лишь пример, и совместно с

различными вариантами по настоящему раскрытию можно использовать широкий спектр других компьютерных архитектур и систем. Например, для обеспечения параллельной обработки можно использовать сверхкомпактный сервер. Для обеспечения параллельной обработки платы с процессорами могут быть соединены с помощью задней панели. Хранилище также может быть подключено к задней панели или в виде сетевого хранилища (NAS) через отдельный сетевой интерфейс. В соответствии с некоторыми вариантами, примеров процессоры могут поддерживать отдельные пространства адресов памяти и передавать данные через сетевые интерфейсы, заднюю панель или другие коннекторы для параллельной обработки другими процессорами. В других вариантах некоторые или все процессоры могут использовать общее пространство виртуальных адресов памяти.

[00154] На **фиг. 11** представлена блок-схема многопроцессорной компьютерной системы **1100**, использующей общее пространство виртуальных адресов памяти в соответствии с вариантом примера. Система включает множество процессоров **1102a-f**, которые могут обращаться к подсистеме **1104** общей памяти. Система включает множество программируемых аппаратных процессоров с алгоритмами доступа к памяти (MAP) **1106a-f** в подсистеме **1104** памяти. Каждый MAP **1106a-f** может содержать память **1108a-f** и одну или несколько программируемых пользователем логических матриц (FPGA) **1110a-f**. MAP представляет собой конфигурируемый функциональный блок, и для обработки в тесной координации с соответствующим процессором в FPGA **1110a-f** можно ввести конкретные алгоритмы или части алгоритмов. Например, в вариантах примеров MAP можно использовать для оценки алгебраических выражений, касающихся модели данных, и для выполнения адаптивной реструктуризации данных. В этом примере для этих целей всем процессорам глобально доступен каждый MAP. В одной конфигурации каждый MAP может использовать прямой доступ к памяти (DMA) для доступа к ассоциированной памяти **1108a-f**, что позволяет ему выполнять задачи независимо и асинхронно от соответствующего микропроцессора **1102a-f**. В этой конфигурации MAP может передавать результаты непосредственно в другой MAP для конвейерной обработки и параллельного выполнения алгоритмов.

[00155] Вышеприведенные компьютерные архитектуры и системы являются лишь примерами, и в сочетании с вариантами примеров можно использовать широкий спектр других архитектур и систем компьютеров, сотовых телефонов и карманных компьютеров, включая системы, в которых использована любая комбинация общих процессоров, сопроцессоров, FPGA и других программируемых логических устройств, систем-на-кристаллах (SOC), специализированных интегральных схем (ASIC) и других логических и обрабатывающих элементов. В соответствии с некоторыми вариантами, вся

компьютерная система или ее часть может быть реализована программно или аппаратно. В сочетании с вариантами примеров можно использовать любой спектр носителей данных, в том числе оперативную память, жесткие диски, флэш-память, ленточные накопители, дисковые массивы, сетевое хранилище данных (NAS) и другие локальные или распределенные устройства и системы хранения данных.

[00156] В вариантах примеров компьютерная система может быть реализована с использованием программных модулей, выполняющихся на любой из вышеперечисленных или других компьютерных архитектурах и системах. В других вариантах функции системы могут быть частично или полностью реализованы в программно-аппаратных средствах, программируемых логических устройствах, таких как программируемые пользователем логические матрицы (FPGA), как показано на **фиг. 9**, системах-на-кристаллах (SOC), специализированных интегральных микросхемах (ASIC) или других обрабатывающих и логических элементах. Например, Set Processor and Optimizer можно реализовать с помощью аппаратного ускорения посредством использования карты аппаратного ускорения, такой как карта **922** ускорителя, проиллюстрированная на **фиг. 9**.

[00157] Последующие примеры приведены для более ясной иллюстрации специалистам в настоящей области техники идеи и практической реализации раскрываемых в настоящем документе вариантов осуществления, и их не следует истолковывать как ограничивающие объем каких-либо заявляемых вариантов осуществления. Если не указано иное, все части и проценты представлены в пересчете на массу.

Примеры

[00158] Приведенные далее примеры представлены с целью иллюстрации различных вариантов осуществления по настоящему раскрытию и не предназначены для какого-либо ограничения настоящего раскрытия. Настоящие примеры, наряду с описанными в настоящем документе способами, в настоящее время являются типичными предпочтительными вариантами осуществления, являются иллюстративными и не предназначены для ограничения объема настоящего раскрытия. Специалистам в настоящей области техники будут очевидны изменения в них и другие применения, которые охватываются сутью настоящего раскрытия, которая определена объемом формулы изобретения.

[00159] **Пример 1. Функционализация поверхности устройства**

[00160] Устройство функционализировали для того, чтобы оно служило основой для присоединения и синтеза библиотеки полинуклеотидов. Поверхность устройства сначала подвергали влажной очистке в течение 20 минут с помощью раствора «пиранья», содержащего 90% H_2SO_4 и 10% H_2O_2 . Устройство промывали несколькими стаканами деионизированной (DI) воды, выдерживали в течение 5 минут под краном с DI водой и сушили посредством N_2 . Затем устройство замачивали в NH_4OH (1:100; 3 мл:300 мл) в течение 5 минут, промывали DI водой из пистолета, замачивали тремя последовательными стаканами DI воды в течение 1 мин каждый, а затем снова промывали DI водой с помощью пистолета. Затем устройство подвергали плазменной очистке, подвергая поверхность устройства воздействию O_2 . Для плазменного травления O_2 использовали устройство SAMCO PC-300 на 250 Вт в течение 1 минуты в нисходящем режиме.

[00161] Очищенную поверхность устройства активно функционализировали с применением раствора, содержащего N-(3-триэтоксисилпропил)-4-гидроксибутирамид, с использованием системы YES-1224P печей для осаждения в паровой фазе со следующими параметрами: 0,5-1 торр, 60 мин, 70°C, испаритель на 135°C. Поверхность устройства покрывали защитным слоем с помощью центрифуги для нанесения покрытий Brewer Science 200X. Наносили фотоустойчивый слой SPR™ 3612 методом центрифугирования на устройстве со скоростью 2500 об/мин в течение 40 с. Устройство предварительно прокаливало в течение 30 мин при 90°C на плитке Брюера. Устройство подвергали фотолитографии с помощью прибора для литографии Karl Suss MA6. Устройство экспонировали на 2,2 с и проявляли в течение 1 мин в MSF 26A. Оставшийся проявитель смывали пистолетом и устройство замачивали водой на 5 минут. Устройство прокаливало в течение 30 мин при 100°C в термостате с последующим визуальным осмотром дефектов литографии с помощью Nikon L200. Процесс удаления окалина применяли для удаления остаточного защитного слоя с использованием прибора SAMCO PC-300 для плазменного травления O_2 при 250 Вт в течение 1 мин.

[00162] Поверхность устройства пассивно функционализировали 100 мкл раствора перфтороктилтрихлорсилана, смешанного с 10 мкл легкого минерального масла. Устройство помещали в камеру, прокачивали в течение 10 мин, а затем закрывали клапан к насосу и оставляли стоять в течение 10 мин. Из камеры удаляли воздух. У устройства удаляли защитный слой путем проведения двух замачиваний в течение 5 минут в 500 мл NMP при 70°C с ультразвуковой обработкой на максимальной мощности (9 по системе Crest). Затем устройство замачивали в течение 5 минут в 500 мл изопропанола при комнатной температуре с ультразвуковой обработкой на максимальной мощности. Устройство погружали в 300 мл этанола с крепостью 200 и сушили продувкой N_2 .

Функционализированную поверхность активировали с тем, чтобы она служила основой для синтеза полинуклеотидов.

[00163] Пример 2. Синтез 50-мерной последовательности на устройстве для синтеза олигонуклеотидов

[00164] Двухмерное устройство для синтеза олигонуклеотидов собирали с проточной ячейкой путем присоединения к проточной ячейке (Applied Biosystems (ДНК-синтезатор ABI394)). Двухмерное устройство для синтеза олигонуклеотидов равномерно функционализировали с помощью N-(3-триэтоксисилпропил)-4-гидроксибутирамида (Gelest) и использовали для синтеза иллюстративного полинуклеотида из 50 п. о. («50-мерного полинуклеотида») с помощью описываемых в настоящем документе способов синтеза полинуклеотидов.

[00165] Последовательность 50-мера олигонуклеотида была такой, как описано под SEQ ID NO.: 48.

5'AGACAATCAACCATTTGGGGTGGACAGCCTTGACCTCTAGACTTCGGCAT##TTTTT TTTTT3' (SEQ ID NO.: 48), где # обозначает тимидин-сукцинилгексамидный CED-фосфорамидит (CLP-2244 от ChemGenes), который является расщепляемым линкером, делающим возможным высвобождение олигонуклеотидов с поверхности в процессе удаления защитной группы.

[00166] Синтез проводили с помощью стандартной химической процедуры синтеза ДНК (связывание, кэппирование, окисление и деблокирование) в соответствии с протоколом, представленным в **таблице 4**, и синтезатора ABI.

Таблица 4. Протоколы синтеза

Общее название процесса синтеза ДНК	Таблица 4	
	Стадия процесса	Время (с)
ПРОМЫВКА (промывка в потоке ацетонитрила)	Промывание системы струей ацетонитрила	4
	Ацетонитрил в проточную ячейку	23
	Продувка системы посредством N ₂	4
	Промывание системы струей ацетонитрила	4

Таблица 4		
Общее название процесса синтеза ДНК	Стадия процесса	Время (с)
ДОБАВЛЕНИЕ ОСНОВАНИЯ ДНК (поток фосфорамидита + активатора)	Промывание системы труб активатором	2
	Активатор в проточную ячейку	6
	Активатор + фосфорамидит в проточную ячейку	6
	Активатор в проточную ячейку	0,5
	Активатор + фосфорамидит в проточную ячейку	5
	Активатор в проточную ячейку	0,5
	Активатор + фосфорамидит в проточную ячейку	5
	Активатор в проточную ячейку	0,5
	Активатор + фосфорамидит в проточную ячейку	5
	Инкубировать в течение 25 с	25
ПРОМЫВКА (промывка в потоке ацетонитрила)	Промывание системы струей ацетонитрила	4
	Ацетонитрил в проточную ячейку	15
	Продувка системы посредством N ₂	4

Таблица 4		
Общее название процесса синтеза ДНК	Стадия процесса	Время (с)
	Промывание системы струей ацетонитрила	4
ДОБАВЛЕНИЕ ОСНОВАНИЯ ДНК (поток фосфорамидита + активатора)	Промывание системы труб активатором	2
	Активатор в проточную ячейку	5
	Активатор + фосфорамидит в проточную ячейку	18
	Инкубировать в течение 25 с	25
ПРОМЫВКА (промывка в потоке ацетонитрила)	Промывание системы струей ацетонитрила	4
	Ацетонитрил в проточную ячейку	15
	Продувка системы посредством N ₂	4
	Промывание системы струей ацетонитрила	4
КЭППИРОВАНИЕ (поток СаpA+B, 1:1)	СаpA+B в проточную ячейку	15
ПРОМЫВКА (промывка в потоке ацетонитрила)	Промывание системы струей ацетонитрила	4
	Ацетонитрил в проточную ячейку	15
	Промывание системы струей ацетонитрила	4
ОКИСЛЕНИЕ (поток окислителя)	Окислитель в проточную ячейку	18
ПРОМЫВКА (промывка в потоке ацетонитрила)	Промывание системы струей ацетонитрила	4

Таблица 4		
Общее название процесса синтеза ДНК	Стадия процесса	Время (с)
	Продувка системы посредством N ₂	4
	Промывание системы струей ацетонитрила	4
	Ацетонитрил в проточную ячейку	15
	Промывание системы струей ацетонитрила	4
	Ацетонитрил в проточную ячейку	15
	Продувка системы посредством N ₂	4
	Промывание системы струей ацетонитрила	4
	Ацетонитрил в проточную ячейку	23
	Продувка системы посредством N ₂	4
	Промывание системы струей ацетонитрила	4
ДЕБЛОКИРОВКА (поток деблокировщика)	Деблокировщик в проточную ячейку	36
ПРОМЫВКА (промывка в потоке ацетонитрила)	Промывание системы струей ацетонитрила	4
	Продувка системы посредством N ₂	4
	Промывание системы струей ацетонитрила	4
	Ацетонитрил в проточную ячейку	18

Таблица 4		
Общее название процесса синтеза ДНК	Стадия процесса	Время (с)
	Продувка системы посредством N ₂	4,13
	Промывание системы струей ацетонитрила	4,13
	Ацетонитрил в проточную ячейку	15

[00167] Сочетание фосфорамидит/активатор доставляли аналогично объемным реагентам через проточную ячейку. Стадии сушки не проводили, поскольку среда все время оставалась «влажной» с реагентом.

[00168] Из синтезатора ABI 394 удаляли ограничитель потока для обеспечения более быстрого потока. Без ограничителя потока скорости потока для амидитов (0,1 М в ACN), активатора (0,25 М бензоилтиотетразола («ВТТ»; 30-3070-xx от GlenResearch) в ACN) и Ох (0,02 М I₂ в 20% пиридина, 10% воды и 70% THF) составляли примерно ~100 мкл/с, для ацетонитрила («ACN») и кэппирующих реагентов (1:1 смеси CapA и CapB, где CapA представлял собой уксусный ангидрид в THF/пиридине, а CapB представлял собой 16% 1-метилимидизол в THF) — примерно ~200 мкл/с, а для деблокировщика (3% дихлоруксусной кислоты в толуоле) — примерно ~300 мкл/сек (по сравнению с ~50 мкл/с для всех реагентов с ограничителем потока). Отслеживали время для полного вытеснения окислителя, длительность для потоков химических реагентов соответствующим образом корректировали и между различными химическими реагентами включали дополнительную промывку посредством ACN. После синтеза полинуклеотидов у чипа удаляли защитные группы в газообразном аммиаке в течение ночи при 75 фунт/кв. дюйм. Для восстановления полинуклеотидов на поверхность наносили пять капель воды. Восстановленные полинуклеотиды затем анализировали на чипе для анализа небольших молекул РНК от BioAnalyzer.

[00169] Пример 3. Синтез 100-мерной последовательности на устройстве для синтеза олигонуклеотидов

[00170] Тот же самый процесс, который описан в примере 2 для синтеза 50-мерной последовательности, использовали для синтеза 100-мерного полинуклеотида («100-мерный полинуклеотид»);

CGGGATCCTTATCGTCATCGTCGTACAGATCCCGACCCATTTGCTGTCCACCAGTCA TGCTAGCCATACCATGATGATGATGATGATGAGAACCCCGCAT##TTTTTTTTTTT3', где # обозначает тимидин-сукцинилгексамидный CED-фосфорамидит (CLP-2244 от ChemGenes); SEQ ID NO: 49) на двух разных силиконовых чипах, первый из которых был равномерно функционализирован N-(3-триэтоксисилпропил)-4-гидроксибутирамидом, а второй был функционализирован 5/95 смесью 11-ацетоксиундецилтриэтоксисилана и н-децилтриэтоксисилана, и полинуклеотиды, выделенные с поверхности, анализировали на приборе BioAnalyzer.

[00171] Все десять образцов с двух чипов дополнительно амплифицировали с помощью ПЦР с применением прямого (5'ATGCGGGGTTCTCATCATC3'; SEQ ID NO.: 50) и обратного (5'CGGGATCCTTATCGTCATCG3'; SEQ ID NO.: 551) праймера в 50 мкл смеси для ПЦР (25 мкл мастермикса NEB Q5, 2,5 мкл 10 мкМ прямого праймера, 2,5 мкл 10 мкМ обратного праймера, 1 мкл полинуклеотида, экстрагированного с поверхности, и воды до 50 мкл) с применением следующей программы термоциклирования:

98°C, 30 с

98°C, 10 с; 63°C, 10 с; 72°C, 10 с; повторяли 12 циклов

72°C, 2 мин.

[00172] ПЦР-продукты также прогоняли на BioAnalyzer, наблюдая резкие пики в 100-мерном положении. Затем амплифицированные с помощью ПЦР образцы подвергали клонированию и секвенированию по Сэнгеру. В **таблице 5** подытожены результаты секвенирования по Сэнгеру для образцов, взятых из зон 1-5 из чипа 1, и для образцов, взятых из зон 6-10 из чипа 2.

Таблица 5. Результаты секвенирования

Зона	Уровень ошибок	Эффективность цикла
1	1/763 п. о.	99,87%
2	1/824 п. о.	99,88%
3	1/780 п. о.	99,87%
4	1/429 п. о.	99,77%
5	1/1525 п. о.	99,93%
6	1/1615 п. о.	99,94%
7	1/531 п. о.	99,81%
8	1/1769 п. о.	99,94%
9	1/854 п. о.	99,88%

Зона	Уровень ошибок	Эффективность цикла
10	1/1451 п. о.	99,93%

[00173] Таким образом, высокое качество и однородность синтезированных полинуклеотидов были воспроизведены на двух чипах с различными химическими свойствами поверхности. В целом, 89% из 100-мерных олигонуклеотидов, которые были секвенированы, представляли собой идеальные последовательности без ошибок, что соответствовало 233 из 262.

[00174] В **таблице 6** подытожены характеристики ошибок для последовательностей, полученных из образцов полинуклеотидов из зон 1-10.

Таблица 6. Характеристики ошибок

ID образца/зона №	SA_00 46/1	SA_00 47/2	SA_00 48/3	SA_00 49/4	SA_00 50/5	OSA_ 0051/6	OSA_ 0052/7	OSA_ 0053/8	OSA_ 0054/9	OSA_ 0055/10
Всего последовательностей	32	32	32	32	32	32	32	32	32	32
Качество секвенирования	25 из 28	27 из 27	26 из 30	21 из 23	25 из 26	29 из 30	27 из 31	29 из 31	28 из 29	25 из 28
Качество олигонуклеотида	23 из 25	25 из 27	22 из 26	18 из 21	24 из 25	25 из 29	22 из 27	28 из 29	26 из 28	20 из 25
Количество совпадений ROI	2500	2698	2561	2122	2499	2666	2625	2899	2798	2348
Мутация ROI	2	2	1	3	1	0	2	1	2	1
Делеция множества оснований ROI	0	0	0	0	0	0	0	0	0	0
Небольшая вставка ROI	1	0	0	0	0	0	0	0	0	0
Делеция отдельного основания ROI	0	0	0	0	0	0	0	0	0	0
Количество крупных делеций	0	0	1	0	0	1	1	0	0	0
Мутация: G>A	2	2	1	2	1	0	2	1	2	1
Мутация: T>C	0	0	0	1	0	0	0	0	0	0

Количество ошибок ROI	3	2	2	3	1	1	3	1	2	1
Уровень ошибок ROI	Ош.: ~1 в 834	Ош.: ~1 в 1350	Ош.: ~1 в 1282	Ош.: ~1 в 708	Ош.: ~1 в 2500	Ош.: ~1 в 2667	Ош.: ~1 в 876	Ош.: ~1 в 2900	Ош.: ~1 в 1400	Ош.: ~1 в 2349
Уровень ошибок ROI минус праймер	Ош. MP: ~1 в 763	Ош. MP: ~1 в 824	Ош. MP: ~1 в 780	Ош. MP: ~1 в 429	Ош. MP: ~1 в 1525	Ош. MP: ~1 в 1615	Ош. MP: ~1 в 531	Ош. MP: ~1 в 1769	Ош. MP: ~1 в 854	Ош. MP: ~1 в 1451

[00175] Пример 4. Конструирование связывающих сопряженный с G-белком рецептор доменов на основе конформационных взаимодействий с лигандом

[00176] Связывающие сопряженный с G-белком рецептор (GPCR) домены конструировали с помощью поверхностей взаимодействия между конформационными лигандами, которые взаимодействуют с молекулами GPCR. Из результатов анализа поверхностей взаимодействия между хемокинами и цитокинами и GPCR было видно, что N-концевой пептид до первого конформационного цистеина представляет собой активационный пептид, а центральная спиральная и бета-изгиб-бета топологии обеспечивают взаимодействие с внеклеточным доменом (ECD) у GPCR.

[00177] Дополнительные 254 лиганда GPCR конструировали на основе перекрестного поиска в базах данных Uniprot и IUPHAR. Лиганды представляли классы взаимодействий, полученные от 112 человек, 71 крысы, 4 свиней, 1 овцы и 1 коровы. Затем лиганды объединяли со 101 последовательностью лигандов от различных видов со следующими аннотациями: ADM, ADM2, родственный агути белок, ангиотензиноген, аннексин A1, апелин, ранний апелиновый рецептор, регулирующий аппетит гормон, бета-дефензин 4A, хемокин с C-C-мотивом, хемокин с C-X-C-мотивом, кальцитонин, связанный с геном кальцитонина пептид, катепсин G, катепсин G (фрагмент), холецистокинин, C3 компонент системы комплемента, C5 компонент системы комплемента, (фрагмент) C5 компонента системы комплемента, кортиколиберин, кортистатин, бета-цитокин SCM-1, эндотелин-2, эндотелин-3, эотаксин, фракталкин, галаниновые пептиды, галанин-подобный пептид, желудочный ингибирующий полипептид, гастрин, гастрин-высвобождающий пептид, глюкагон, регулируемый стадией роста альфа-белок, связывающий гем белок 1, гуманин, инсулиноподобный фактор 3, инсулиноподобный пептид INSL5, интерлейкин-8, островковый амилоидный полипептид, кининоген-1, лимфотактин, подавитель метастазирования KiSS-1, нейрокинин-B, нейромедин-B, нейромедин-S, нейромедин-U, нейропептид B, нейропептид S, нейропептид W, нейротензин/нейромедин N, орексигенный

нейропептид QRFP, орексин, окситоцин-нейрофизин 1, прогормон поджелудочной железы, паратиреоидный гормон, родственный паратиреоидному гормону белок, пептид YY, активирующий аденилатциклазу фактор гипофиза, основной белок тромбоцитов, фактор тромбоцитов 4, препроноцицептин, родственный про-FMRFamide нейропептид FF, родственный про-FMRFamide VF нейропептид, про-MCH, пронеуропептид Y, проопиомеланокортин, протиреотропин-высвобождающий гормон, проенкефалин-А, проенкефалин-В, прогонадолиберин-1, прогонадолиберин-2, прокинетицин-1, прокинетицин-2, пролактин-высвобождающий пептид, промотилин, протачикинин-1, белок Wnt-2, белок Wnt-3a, белок Wnt-4, белок Wnt-5a, белок Wnt-7b, протромбин, протоонкоген Wnt-1, протонкоген Wnt-3, предполагаемый неохарактеризованный белок, RCG55748, рецептор ретиноевой кислоты, секретин, соматолиберин, соматостатин, фактор стромальных клеток, Т-кининоген 2, тубероинфундибулярный пептид, урокортин, урокортин-2, урокортин-3, уротензин-2, уротензин-2В, хемокин, совместно регулируемый с VEGF, пептиды VIP и вазопрессин-нейрофизиновый 2-копептин.

[00178] Проводили структурный анализ лигандов и определяли, что большинство из них содержали N-концевой активационный пептид из приблизительно 11 аминокислот. Затем создавали варианты мотива путем обратной обрезки N-концевого активационного пептида. Как можно видеть из **таблицы 7**, иллюстративный набор вариантов создавали на основе N-концевого активационного пептида для стромального фактора-1. Варианты мотивов также комбинаторно размещали во множестве положений в CDR-H3. Всего для размещения в CDR-H3 извлекали 1016 мотивов. Кроме того, вносили варианты мотивов с переменным расположением границ и с 5-20 вариантами подцепочек, которые также размещали в CDR-H3.

Таблица 7. Варианты аминокислотных последовательностей для стромального фактора-1

SEQ ID NO.		Вариант	Аминокислотная последовательность
53		1	ggggSDYKPVLSYR
54		2	ggggDYKPVLSYR
55		3	ggggYKPVLSYR
56		4	ggggKPVLSYR
57		5	ggggPVLSYR

[00179] Как можно видеть из **таблицы 8**, иллюстративный набор вариантов создавали для интерлейкина-8 на основе следующей последовательности:

**MTSKLAVALLAAFLISAALCEGAVLPRSAKELRCQCIKTYSKPFHPKFIKELRVIESGPH
CANTEIIVKLSGDRELCLDPKENWVQRVVEKFLKRAENS (SEQ ID NO: 52).**

Таблица 8. Варианты аминокислотных последовательностей для интерлейкина-8

SEQ ID NO.	Вариант	Аминокислотная последовательность
58	1	ggggSAALCEGAVLPRSA
59	2	ggggAALCEGAVLPRSA
60	3	ggggALCEGAVLPRSA
61	4	ggggLCEGAVLPRSA
62	5	ggggCEGAVLPRSA
63	6	ggggSAALCEGAVLPRSAKE
64	7	ggggAALCEGAVLPRSAKE
65	8	ggggALCEGAVLPRSAKE
66	9	ggggLCEGAVLPRSAKE
67	10	ggggCEGAVLPRSAKE
68	11	ggggSAALCEGAVLPRSAKELR
69	12	ggggAALCEGAVLPRSAKELR
70	13	ggggALCEGAVLPRSAKELR
71	14	ggggLCEGAVLPRSAKELR
72	15	ggggCEGAVLPRSAKELR

[00180] Пример 5. Конструирование связывающих сопряженный с G-белком рецептор доменов на основе взаимодействий с пептидным лигандом

[00181] Связывающие GPCR домены конструировали на основе поверхностей взаимодействия между пептидными лигандами, которые взаимодействуют с GPCR класса В. Использовали приблизительно 66 различных лигандов, и к ним относились последовательности лигандов со следующими аннотациями: адренomedуллин, амилин, ангиотензин, ангиотензин I, ангиотензин II, ангиотензин III, апелин, апстатин, большой эндотелин, большой гастрин, брадикинин, церулин, кальцитонин, связанный с геном кальцитонина пептид, CGRP, холецистокинин, эндотелин, эндотелин 1, эндотелин 2, эндотелотин 3, GIP, различные GIP, GLP, галанин, гастрин, грелин, глюкагон, IAPP, кисспептин, Мса, метастин, нейромедин, нейромедин N, нейропептид, нейропептид F,

нейропептид Y, нейротензин, ноцицептин, орексин, орексин А, орфанин, окситоцин, галанин окситоцина, PACAP, различные PACAP, пептиды PAR (активируемый протеазой рецептор), агонист PAR-1, прамлинтид, сцилиоринин I, секретин, сенктид, соматостатин, соматостатин 14, соматостатин 28, вещество P, уротензин II, VIP, различные VIP, вазопрессин, ксенин, циннамоил, фуроил, гастрин, холецистокинин, фактор полового феромона α . Наблюдали, что пептиды формировали стабилизированное взаимодействие с внеклеточным доменом GPCR (ECD).

[00182] Варианты мотива создавали на основе поверхности взаимодействий пептидов с ECD, а также поверхности взаимодействий N-концевого лиганда с GPCR. Это осуществляли с помощью структурного моделирования. Иллюстративные варианты мотивов создавали на основе взаимодействия глюкагоноподобного пептида с его GPCR, как показано в **таблице 9**. Варианты последовательностей мотива создавали с помощью следующей последовательности из глюкагоноподобного пептида: HAEGTFTSDVSSYLEGQAAKEFIAWLKGRG (SEQ ID NO: 73).

Таблица 9. Варианты аминокислотных последовательностей для глюкагоноподобного пептида

SEQ ID NO.	Вариант	Аминокислотная последовательность
74	1	sggggsgggsggggHAEGTFTSDVSSYLEGQAAKEFIAWL
75	2	sggggsgggsggggAEGTFTSDVSSYLEGQAAKEFIAWL
76	3	sggggsgggsggggEGTFTSDVSSYLEGQAAKEFIAWL
77	4	sggggsgggsggggGTFTSDVSSYLEGQAAKEFIAWL
78	5	sggggsgggsggggTFTSDVSSYLEGQAAKEFIAWL
79	6	sggggsgggsggggFTSDVSSYLEGQAAKEFIAWL
80	7	sggggsgggsggggTSDVSSYLEGQAAKEFIAWL
81	8	sggggsgggsggggSDVSSYLEGQAAKEFIAWL
82	9	sggggsgggsggggDVSSYLEGQAAKEFIAWL

[00183] Пример 6. Конструирование связывающих сопряженный с G-белком рецептор доменов на основе взаимодействий с малыми молекулами

[00184] Связывающие GPCR домены сконструировали на основе поверхностей взаимодействий между низкомолекулярными лигандами, которые взаимодействуют с GPCR. Посредством анализа множества лигандов GPCR создавали аминокислотную библиотеку для Tyr, Pro, Phe, His и Gly, которая может воспроизводить

многие структурные контакты этих лигандов. Иллюстративный вариант мотива, который был создан на основе этих результатов наблюдений, содержал следующую последовательность:

sgggg(F,G,H,P,Y)(F,G,H,P,Y)(F,G,H,P,Y)(F,G,H,P,Y)(F,G,H,P,Y)(F,G,H,P,Y)(F,G,H,P,Y)(F,G,H,P,Y) (SEQ ID NO: 83).

[00185] Пример 7. Конструирование связывающих сопряженный с G-белком рецептор доменов на основе взаимодействий с внеклеточными доменами

[00186] Связывающие GPCR домены конструировали на основе поверхностей взаимодействия на внеклеточных доменах (ECD) и внеклеточных петлях (ECL) GPCR. Было проанализировано приблизительно 2257 GPCR от человека (356), мыши (369), крысы (259), коровы (102), свиньи (60), примата, рыбы, мухи и более 200 других организмов, и было обнаружено, что ECD характеризуются множественными комплементарными контактами с другими петлями и спиральями GPCR, которые в длину составляли 15 аминокислот. По результатам дальнейшего анализа ECL из приблизительно 2257 GPCR и всех разрешенных структур GPCR было видно, что N-концевые ECD1 и ECL2 содержали более длинные внеклеточные последовательности и обеспечивали внеклеточные контакты с GPCR.

[00187] Затем на основании этих последовательностей создавали варианты мотива. Создавали иллюстративные варианты на основании следующей последовательности из индуцируемого ретиноевой кислотой белка 3 (GPCR5A): EYIVLTMNRTNVNVFSELSAPRRNED (SEQ ID NO: 84). См. **таблицу 10**.

Таблица 10. Аминокислотные последовательности

SEQ ID NO.	Вариант	Аминокислотная последовательность
85	1	YIVLTMNRTNVNVFSELSAPRRNE
86	2	IVLTMNRTNVNVFSELSAPRRN
87	3	VLTMNRTNVNVFSELSAPRR

[00188] Пример 8. Конструирование остовов антител

[00189] Для создания остовов проводили структурный анализ, анализ тяжелой цепи по результатам секвенирования репертуара и специфический анализ наборов данных высокопроизводительного секвенирования гетеродимеров. Каждую тяжелую цепь связывали с каждым остовом легкой цепи. Каждый остов тяжелой цепи соотносили с 5 различными вариантами длинных петель CDR-H3. Каждый остов легкой цепи соотносили

с 5 различными остовами L3. Стебли CDR-H3 тяжелой цепи выбирали из часто встречающихся стеблей длинной петли H3 (10 аминокислот на N-конце и C-конце), встречающихся как у индивидуумов, так и среди сегментов V-гена. L3 остова легкой цепи выбирали из гетеродимеров, содержащих длинные H3. Использовали прямые гетеродимеры, основанные на информации из банка данных белков (PDB) и наборов данных, полученных в результате глубокого секвенирования, в которых были стебли CDR H1, H2, L1, L2, L3 и CDR-H3 были фиксированными. Затем для различных остовов подбирали формат для экспонирования на фаге для оценки экспрессии.

[00190] *Структурный анализ*

[00191] Анализировали приблизительно 2017 структур антител, из которых изучали 22 структуры с длинными CDR-H3 длиной по меньшей мере 25 аминокислот. К тяжелым цепям относились следующие: IGHV1-69, IGHV3-30, IGHV4-49 и IGHV3-21. К выявленным легким цепям относились следующие: IGLV3-21, IGKV3-11, IGKV2-28, IGKV1-5, IGLV1-51, IGLV1-44 и IGKV1-13. При анализе несколько раз наблюдали четыре комбинации гетеродимеров, в том числе: IGHV4-59/61-IGLV3-21, IGHV3-21-IGKV2-28, IGHV1-69-IGKV3-11 и IGHV1-69-IGKV1-5. Анализ последовательностей и структур позволил выявить дисульфидные связи внутри CDR-H3 в нескольких структурах с упаковкой объемных боковых цепей, таких как тирозин в стебле, обеспечивающих поддержание длительной стабильности H3. Также были обнаружены вторичные структуры, в том числе бета-изгиб-бета-листы и «молотообразный» субдомен.

[00192] *Анализ репертуара*

[00193] Анализ репертуара проводили на 1083875 последовательностях рецепторов IgM+/CD27 наивных В-клеток (BCR) и на 1433011 последовательностях CD27+, полученных с помощью несмещенной 5'RACE от 12 здоровых контролей. 12 здоровых контролей включали равное количество мужчин и женщин и были составлены из 4 европеоидных, 4 азиатских и 4 латиноамериканских индивидуумов. Из результатов анализа репертуара было видно, что менее 1% человеческого репертуара содержало BCR с CDR-H3, которые были длиннее 21 аминокислоты. Смещение V-гена наблюдали в субрепертуаре длинных CDR3, при этом у IGHV1-69, IGHV4-34, IGHV1-18 и IGHV1-8 наблюдали преимущественное увеличение количества BCR с длинными петлями H3. Смещение относительно длинных петель наблюдали у IGHV3-23, IGHV4-59/61, IGHV5-51, IGHV3-48, IGHV3-53/66, IGHV3-15, IGHV3-74, IGHV3-73, IGHV3-72 и IGHV2-70. Было показано, что остов IGHV4-34 был аутореактивным и имел короткий период полужизни.

[00194] На основе эталонного репертуара, полученного с помощью 5'RACE, также конструировали жизнеспособные варианты остовов с N-концевыми и C-концевыми CDR-H3 для длинных петель. Было обнаружено приблизительно 81065 CDR-H3 с длиной в 22 аминокислоты или более. Путем сравнения остовов V-гена удавалось избежать специфической для остова вариабельности стебля H3, что позволяло клонировать различные остовы во множественные эталонные для остова последовательности.

[00195] *Анализ гетеродимеров*

[00196] Анализ гетеродимеров проводили на остовах, имеющих последовательности, которые показаны на **фиг. 12А-12С**. Анализировали варианты последовательностей и длины остовов.

[00197] *Структурный анализ*

[00198] Структурный анализ проводили с применением остовов GPCR с различными последовательностями и длинами. См. **фиг. 13**.

[00199] **Пример 9. Создание библиотек антител к GPCR**

[00200] На основе поверхностей взаимодействия с лигандами GPCR и структур остовов de novo конструировали и синтезировали библиотеки. См. примеры 4-8. Обратимся к **фиг. 5**: 10 вариантов последовательностей конструировали для вариабельного домена тяжелой цепи **503**, 237 вариантов последовательностей конструировали для определяющего комплементарность участка 3 тяжелой цепи **507** и 44 варианта последовательностей конструировали для вариабельного домена легкой цепи **513**. Фрагменты синтезировали в виде трех фрагментов, которые можно видеть на **фиг. 6**, согласно способам, аналогичным описанным в примерах 1-3.

[00201] После синтеза de novo 10 вариантов последовательностей получали для вариабельного домена тяжелой цепи **602**, 236 вариантов последовательностей получали для определяющего комплементарность участка 3 тяжелой цепи **604** и 43 варианта последовательностей получали для участка, содержащего вариабельный домен **606** легкой цепи и CDR-L3, и из них конструировали 9 вариантов вариабельного домена легкой цепи. Это приводило к созданию библиотеки с приблизительным разнообразием 10^5 ($10 \times 236 \times 43$). Это подтверждали с помощью секвенирования следующего поколения (NGS) с 16 миллионами считываний. Как можно видеть на **фиг. 14**, нормализованные результаты считывания секвенирования для каждого из 10 вариантов для вариабельного домена тяжелой цепи были равны приблизительно 1. Как можно видеть на **фиг. 15**,

нормализованные результаты считывания секвенирования для каждого из 43 вариантов для переменного домена легкой цепи были равны приблизительно 1. Как можно видеть на **фиг. 16** видно, нормализованные результаты считывания секвенирования для 236 вариантов последовательностей для определяющего комплементарность участка 3 тяжелой цепи, были равны приблизительно 1.

[00202] Затем различные легкие и тяжелые цепи тестировали на экспрессию и фолдинг белка. Обратимся к **фиг. 17A-17D**: 10 вариантов последовательностей для переменного домена тяжелой цепи включали следующие: IGHV1-18, IGHV1-69, IGHV1-8 IGHV3-21, IGHV3-23, IGHV3-30/33m, IGHV3-28, IGHV3-74, IGHV4-39 и IGHV4-59/61. Из 10 вариантов последовательностей IGHV1-18, IGHV1-69 и IGHV3-30/33m имели улучшенные характеристики, такие как улучшенная термостабильность. Обратимся к **фиг. 18A-18F**: 9 вариантов последовательностей для переменного домена легкой цепи включали следующие: IGKV1-39, IGKV1-9, IGKV2-28, IGKV3-11, IGKV3-15, IGKV3-20, IGKV4-1, IGLV1-51 и IGLV2-14. Из 9 вариантов последовательностей IGKV1-39, IGKV3-15, IGLV1-51 и IGLV2-14 имели улучшенные характеристики, такие как улучшенная термостабильность.

[00203] **Пример 10. Экспрессия библиотек антител к GPCR в клетках HEK293**

[00204] После создания библиотек антител к GPCR, как показано в примере 13, для скрининга отбирали приблизительно 47 GPCR. Разрабатывали схему конструкций GPCR размером от приблизительно 1,8 до приблизительно 4,5 т. о. в векторе pCDNA3.1. Затем конструкции GPCR синтезировали согласно способам, аналогичным описанным в примерах 2-4, включая иерархическую сборку. Из 47 конструкций GPCR синтезировали 46 конструкций GPCR.

[00205] Синтезированные конструкции GPCR трансфицировали в HEK293 и анализировали на экспрессию с помощью иммунофлуоресценции. Обратимся к **фиг. 19A-19C**: клетки HEK293 трансфицировали конструкциями GPCR, содержащими меченный гемагглютинином (НА) на N-конце человеческий рецептор Y_1 . Спустя 24-48 часов после трансфекции клетки промывали фосфатно-солевым буферным раствором (PBS) и фиксировали 4% параформальдегидом. Клетки окрашивали с помощью флуоресцентного первичного антитела к НА-метке или вторичного антитела, содержащего флуорофор и DAPI, для визуализации ядер в синем свете. Обратимся к **фиг. 19A-19C**: человеческий рецептор Y_1 визуализировали на клеточной поверхности непрмеабилитированных клеток, а также на клеточной поверхности и внутриклеточно у прмеабилитированных клеток.

[00206] Конструкции GPCR также визуализировали путем разработки конструкций GPCR, содержащих аутофлуоресцентные белки. Обратимся к **фиг. 20А-20С**: человеческий рецептор Y₁ содержал EYFP, слитый с его С-концом, а человеческий рецептор Y₅ содержал ECFP, слитый с его С-концом. Клетки НЕК293 трансфицировали человеческим рецептором Y₁ или совместно трансфицировали человеческим рецептором Y₁ и человеческим рецептором Y₅. После трансфекции клетки промывали и фиксировали 4% параформальдегидом. Клетки окрашивали посредством DAPI. Локализацию человеческого рецептора Y₁ и человеческого рецептора Y₅ визуализировали с помощью флуоресцентной микроскопии.

[00207] Пример 11. Конструирование библиотеки иммуноглобулинов

[00208] Конструировали библиотеку остовов иммуноглобулина для размещения связывающих GPCR доменов и для улучшения стабильности для ряда последовательностей, кодирующих связывающий GPCR домен. Остов иммуноглобулина включал домен VH, связанный с доменом VL при помощи линкера. Создавали варианты последовательностей нуклеиновых кислот для каркасных элементов и CDR-элементов домена VH и домена VL. Структура конструкции показана на **фиг. 21А**. Полная архитектура домена показана на **фиг. 21В**. Последовательности для лидерной последовательности, линкера и рIII приведены в **таблице 11**.

Таблица 11. Нуклеотидные последовательности

SE Q ID NO	Домен	Последовательность
88	Лидерна я	GCAGCCGCTGGCTTGCTGCTGCTGGCAGCTCAGCCGGCCATGGC C
89	Линкер	GCTAGCGGTGGAGGCGGTTTCAGGCGGAGGTGGCTCTGGCGGTG GC GGATCGCATGCATCC
90	рIII	CGCGCGGCCGCTGGAAGCGGCTCCCACCATCACCATCACCAT

[00209] Конструировали домены VL, к которым относятся IGKV1-39, IGKV3-15, IGLV1-51 и IGLV2-14. Каждый из четырех доменов VL собирали с их соответствующими инвариантными четырьмя каркасными элементами (FW1, FW2, FW3, FW4) и 3 переменными CDR (L1, L2, L3) элементами. В случае IGKV1-39 было 490

вариантов, сконструированных для L1, 420 вариантов, сконструированных для L2, и 824 варианта, сконструированных для L3, что давало разнообразие в $1,7 \times 10^8$ ($490 \times 420 \times 824$). В случае IGKV3-15 было 490 вариантов, сконструированных для L1, 265 вариантов, сконструированных для L2, и 907 вариантов, сконструированных для L3, что давало разнообразие в $1,2 \times 10^8$ ($490 \times 265 \times 907$). В случае IGLV1-51 было 184 варианта, сконструированных для L1, 151 вариант, сконструированный для L2, и 824 варианта, сконструированных для L3, что давало разнообразие в $2,3 \times 10^7$ ($184 \times 151 \times 824$). В случае IGLV2-14 было 967 вариантов, сконструированных для L1, 535 вариантов, сконструированных для L2, и 922 варианта, сконструированных для L3, что давало разнообразие в $4,8 \times 10^8$ ($967 \times 535 \times 922$). В **таблице 12** приведены аминокислотные последовательности и нуклеотидные последовательности для четырех каркасных элементов (FW1, FW2, FW3, FW4) для IGLV1-51. В **таблице 13** приведены 3 переменных элемента CDR (L1, L2, L3) для IGLV1-51. Также для IGKV1-39, IGKV3-15 и IGLV2-14 конструировали варианты аминокислотных последовательностей и нуклеотидных последовательностей для четырех каркасных элементов (FW1, FW2, FW3, FW4) и 3 переменных CDR (L1, L2, L3) элементов.

Таблица 12. Последовательности для каркасных элементов IGLV1-51

Элемент	SEQ ID NO	Аминокислотная последовательность	SEQ ID NO	Нуклеотидная последовательность
IGLV1-51				
FW1	91	QSVLTQPPSVSAAPGQKV TISC	92	CAGTCTGTGTTGACGCAGCCGCC TCAGTGTCTGCGGCCCCAGGACAG AAGGTCACCATCTCCTGC
FW2	93	WYQQLPGTAPKLLIY	94	TGGTATCAGCAGCTCCCAGGAACA GCCCCAAACTCCTCATTTAT
FW3	95	GIPDRFSGSKSGTSATLGI TGLQTGDEADYY	96	GGGATTCCTGACCGATTCTCTGGC TCCAAGTCTGGCACGTCAGCCACC CTGGGCATCACCGGACTCCAGACT GGGGACGAGGCCGATTATTAC
FW4	97	GGGTKLTVL	98	GGCGGAGGGACCAAGCTGACCGTC CTA

Таблица 13. Последовательности для CDR элементов IGLV1-51

SEQ ID NO	Аминокислотная последовательность	SEQ ID NO	Нуклеотидная последовательность
IGLV1-51-L1			
99	SGSSSNIGS NHVS	282	TCTGGAAGCAGCTCCAACATTGGGAGTAATCATG TATCC
100	SGSSSNIGN NYLS	283	TCTGGAAGCAGCTCCAACATTGGGAATAATTATC TATCC
101	SGSSSNIAN NYVS	284	TCTGGAAGCAGCTCCAACATTGCGAATAATTATG TATCC
102	SGSSPNIGN NYVS	285	TCTGGAAGCAGCCCCAACATTGGGAATAATTATG TATCG
103	SGRSNIGS NYVS	286	TCTGGAAGCAGATCCAATATTGGGAGTAATTATG TTTCG
104	SGSSSNVG DNYVS	287	TCTGGAAGCAGCTCCAACGTTGGCGATAATTATG TTTCC
105	SGSSSNIGI QYVS	288	TCTGGAAGCAGCTCCAACATTGGGATTCAATATG TATCC
106	SGSSSNVG NNFVS	289	TCTGGAAGCAGCTCCAATGTTGGTAACAATTTTGT CTCC
107	SGSASNIG NNYVS	290	TCTGGAAGCGCCTCCAACATTGGGAATAATTATG TATCC
108	SGSGSNIG NNDVS	291	TCTGGAAGCGGCTCCAATATTGGGAATAATGATG TGTC
109	SGSISNIGN NYVS	292	TCTGGAAGCATCTCCAACATTGGTAATAATTATGT ATCC
110	SGSISNIGK NYVS	293	TCTGGAAGCATCTCCAACATTGGGAAAAATTATG TGTCG
111	SGSSSNIGH NYVS	294	TCTGGAAGCAGCTCCAACATTGGGCATAATTATG TATCG
112	PGSSSNIGN NYVS	295	CCTGGAAGCAGCTCCAACATTGGGAATAATTATG TATCC

113	SGSTSNIGI HYVS	296	TCTGGAAGCACCTCCAACATTGGAATTCATTATGT ATCC
114	SGSSSNIGS HYVS	297	TCTGGAAGCAGCTCCAACATTGGCAGTCATTATG TTTCC
115	SGSSSNIGN EYVS	298	TCCGGAAGCAGCTCCAACATTGGAAATGAATATG TATCC
116	SGSTSNIG NNYIS	299	TCTGGAAGCACCTCCAACATTGGAAATAATTATA TATCG
117	SGSSSNIGN HFVS	300	TCTGGAAGCAGCTCCAATATTGGGAATCATTTTGT ATCG
118	SGSSSNIGN NYVA	301	TCTGGAAGCAGCTCCAACATTGGGAATAATTATG TGGCC
119	SGSSSNIGS YYVS	302	TCTGGAAGCAGCTCCAACATTGGAAGTTATTATG TATCC
120	SGSGFNIG NNYVS	303	TCTGGAAGTGGTTTCAACATTGGGAATAATTATG TCTCT
121	SGSTSNIG NNYVS	304	TCTGGAAGCACCTCCAACATTGGGAATAATTATG TGTCC
122	SGSSSDIGN NYVS	305	TCTGGAAGCAGCTCCGACATTGGCAATAATTATG TATCC
123	SGSSSNIGN NVVS	306	TCTGGAAGCAGCTCCAACATTGGGAATAATGTTG TATCC
124	SGSKSNIG KNYVS	307	TCTGGAAGCAAGTCTAACATTGGGAAAATTATG TATCC
125	SGSSTNIG NNYVS	308	TCTGGAAGCAGCACCAACATTGGGAATAATTATG TATCC
126	SGSISNIGD NYVS	309	TCTGGAAGCATCTCCAACATTGGGGATAATTATG TATCC
127	SGSSSNIGS KDVS	310	TCTGGAAGCAGCTCCAACATTGGGAGTAAGGATG TATCA
128	SGSSSNIEN NDVS	311	TCTGGAAGCAGCTCCAACATTGAGAATAATGATG TATCG
129	SGSSSNIGN HYVS	312	TCTGGAAGCAGCTCCAACATTGGGAATCATTATG TATCC

130	SGSSSNIGK DFVS	313	TCTGGAAGCAGCTCCAACATTGGGAAGGATTTG TCTCC
131	SGSTSNIGS NFVS	314	TCTGGCAGTACTTCCAACATCGGAAGTAATTTGT TTCC
132	SGSTSNIG HNYVS	315	TCTGGAAGCACCTCCAACATTGGGCATAATTATG TATCC
133	SASSSNIGN NYVS	316	TCTGCAAGCAGCTCCAACATTGGGAATAATTATG TATCC
134	SGSSSIGN NYVS	317	TCTGGAAGCAGCTCCAGCATTGGCAATAATTATG TATCC
135	SGSSSTIGN NYVS	318	TCTGGAAGCAGCTCCACCATTGGGAATAATTATG TATCC
136	SGSSNIEN NYVS	319	TCTGGAAGCAGCTCCAACATTGAAAATAATTATG TATCC
137	SGSSSNIGN QYVS	320	TCTGGAAGCAGCTCCAACATTGGGAATCAGTATG TATCC
138	SGSSSNIGN NYVF	321	TCTGGAAGCAGCTCCAACATTGGGAATAATTATG TATTC
139	SGSSSNIGR NYVS	322	TCTGGAAGCAGCTCCAACATTGGGAGGAATTATG TCTCC
140	SGGSSNIG NYYVS	323	TCTGGAGGCAGCTCCAACATTGGAAATTATTATG TATCG
141	SGSSSNIGD NYVS	324	TCTGGAAGCAGCTCCAACATTGGAGATAATTATG TCTCC
142	SGGSSNIGI NYVS	325	TCTGGAGGCAGCTCCAACATTGGAATTAATTATG TATCC
143	SGGSSNIG KNYVS	326	TCTGGAGGCAGCTCCAACATTGGGAAGAATTATG TATCC
144	SGSSSNIGK RSVS	327	TCTGGAAGCAGCTCCAACATTGGGAAGAGATCTG TATCG
145	SGRSNIG NNYVS	328	TCTGGAAGCAGATCCAACATTGGGAATAACTATG TATCC
146	SGSSSNIGN NLVS	329	TCGGGAAGCAGCTCCAACATTGGGAATAATCTTG TTCC

147	SGSSSNIGI NYVS	330	TCTGGAAGCAGCTCCAACATTGGGATCAATTATG TATCC
148	SGSSSNIGN NFVS	331	TCTGGAAGCAGCTCCAACATCGGGAATAATTTTG TATCC
149	SGTSSNIGR NFVS	332	TCTGGAACCAGCTCCAACATTGGCAGAAATTTTG TATCC
150	SGRRSNIG NNYVS	333	TCTGGAAGGAGGTCCAACATTGGAAATAATTATG TGTCC
151	SGGSFNIG NNYVS	334	TCTGGAGGCAGCTTCAATATTGGGAATAATTATG TATCC
152	SGSTSNIGE NYVS	335	TCTGGAAGCACTTCCAACATTGGGGAGAATTATG TGTCC
153	SGSSSNIGS DYVS	336	TCTGGAAGCAGCTCCAATATTGGGAGTGATTATG TATCC
154	SGTSSNIGS NYVS	337	TCTGGAACCAGCTCCAACATTGGGAGTAATTATG TATCC
155	SGSSSNIGT NFVS	338	TCTGGAAGCAGCTCCAACATTGGGACTAATTTTG TATCC
156	SGSSSNFG NNYVS	339	TCTGGAAGCAGCTCCAACCTTGGGAATAATTATG TATCC
157	SGSTSNIG NNHVS	340	TCTGGAAGCACCTCCAACATTGGGAATAATCATG TATCC
158	SGSSSNIGN DFVS	341	TCTGGAAGCAGCTCCAACATTGGGAATGATTTTG TATCC
159	SGSSSDIGD NYVS	342	TCTGGAAGCAGCTCCGACATTGGCGATAATTATG TGTCC
160	SGSSSNIGK YYVS	343	TCTGGAAGCAGCTCCAACATTGGGAATATTATG TATCC
161	SGSSSNIGG NYVS	344	TCTGGAAGCAGCTCCAACATTGGCGGTAATTATG TATCC
162	SGSSSNTG NNYVS	345	TCTGGAAGCAGCTCCAACACTGGGAATAATTATG TATCC
163	SGSSSNVG NNYVS	346	TCTGGAAGCAGCTCCAACGTTGGGAATAATTATG TGTCT

164	SGSSSNIAN NFVS	347	TCTGGAAGCAGCTCCAACATTGCGAATAATTTTG TATCC
165	SGSSSNIGN DYVS	348	TCTGGAAGCAGCTCCAACATTGGGAATGATTATG TATCC
166	SGSTSNIE NYVS	349	TCTGGAAGCACCTCCAATATTGAGAATAATTATG TTTCC
167	SGGSSNIG NNDVS	350	TCTGGAGGCAGCTCCAATATTGGCAATAATGATG TGTCC
168	SGSTSNIG NHYVS	351	TCTGGAAGCACCTCCAACATTGGGAATCATTATG TATCC
169	SGSSSNIGD NDVS	352	TCAGGAAGCAGCTCCAATATTGGGGATAATGATG TATCC
170	SGYSSNIG NNYVS	353	TCTGGATACAGCTCCAACATTGGGAATAATTATG TATCC
171	SGSGSNIG NNFVS	354	TCTGGAAGCGGCTCCAACATTGGAAATAATTTTG TATCC
172	SGSSSNIW NNYVS	355	TCTGGAAGCAGCTCCAACATTTGGAATAATTATG TATCC
173	FGSSSNIGN NYVS	356	TTTGGAAGCAGCTCCAACATTGGGAATAATTATG TATCC
174	SGSSSNIK NYVS	357	TCTGGAAGCAGCTCCAACATTGAGAAGAATTATG TATCC
175	SGSRSNIG NYYVS	358	TCTGGAAGTAGATCCAATATTGGAAATTATTATG TATCC
176	SGTKSNIG NNYVS	359	TCTGGAACCAAGTCAAACATTGGGAATAATTATG TATCT
177	SGSTSNIG NYYVS	360	TCTGGAAGCACCTCCAACATTGGGAATTATTATG TATCC
178	SGTSSNIG NNYVA	361	TCTGGAACCAGCTCCAACATTGGGAATAATTATG TGGCC
179	PGTSSNIG NNYVS	362	CCTGGAACCAGCTCCAACATTGGGAATAATTATG TATCC
180	SGSTSNIGI NYVS	363	TCCGGAAGCACCTCCAACATTGGGATTAATTATG TATCC

181	SGSSSNIGS NLVS	364	TCTGGAAGCAGCTCCAACATTGGGAGTAATCTGG TATCC
182	SGSSSNIEN NHVS	365	TCTGGAAGCAGCTCCAACATTGAGAATAATCATG TATCC
183	SGTRSNIG NNYVS	366	TCTGGAACCAGGTCCAACATCGGCAATAATTATG TTTCG
184	SGSTSNIG DNYVS	367	TCTGGAAGCACCTCCAACATTGGGGACAATTATG TTTCC
185	SGGSSNIG KNFVS	368	TCTGGAGGCAGTTCCAACATTGGGAAGAATTTTG TATCC
186	SGSRSDIG NNYVS	369	TCTGGAAGCAGGTCCGACATTGGGAATAATTATG TATCC
187	SGTSSNIG NNDVS	370	TCTGGAAGTAGCTCCAACATTGGGAATAATGATG TATCC
188	SGSSSNIGS KYVS	371	TCTGGAAGCAGCTCCAACATTGGGAGTAAATATG TATCA
189	SGSSFNIGN NYVS	372	TCTGGAAGCAGCTTCAACATTGGGAATAATTATG TATCC
190	SGSSSNIGN TYVS	373	TCTGGAAGCAGCTCCAACATTGGGAATACTTATG TATCC
191	SGSSSNIGD NHVS	374	TCTGGAAGCAGCTCCAATATTGGGGATAATCATG TATCC
192	SGSSSNIGN NHVS	375	TCTGGAAGCAGCTCCAACATTGGCAATAATCATG TTTCC
193	SGSTSNIG NNDVS	376	TCTGGAAGCACCTCCAACATTGGGAATAATGATG TATCC
194	SGSRSNVG NNYVS	377	TCTGGAAGCAGATCCAACGTTGGCAATAATTATG TTTCA
195	SGGTSNIG KNYVS	378	TCCGGAGGCACCTCCAACATTGGGAAGAATTATG TGTCT
196	SGSSSNIAD NYVS	379	TCTGGAAGCAGCTCCAACATTGCCGATAATTATG TTTCC
197	SGSSSNIGA NYVS	380	TCTGGAAGCAGCTCCAACATTGGCGCCAATTATG TATCC

198	SGSSSNIGS NYVA	381	TCTGGAAGCAGCTCCAACATTGGGAGTAATTATG TGGCC
199	SGSSSNIGN NFLS	382	TCTGGAAGCAGCTCCAACATTGGGAACAATTTTC TCTCC
200	SGRSSNIG KNYVS	383	TCTGGAAGAAGCTCCAACATTGGGAAGAATTATG TATCC
201	SGSSPNIGA NYVS	384	TCTGGAAGCAGCCCCAACATTGGGGCTAATTATG TATCC
202	SGSSSNIGP NYVS	385	TCCGGAAGCAGCTCCAACATTGGGCCTAATTATG TGTCC
203	SGSSSTIGN NYIS	386	TCTGGAAGCAGCTCCACCATTGGGAATAATTATA TATCC
204	SGSSSNIGN YFVS	387	TCTGGAAGCAGCTCCAACATTGGGAATTATTTTGT ATCC
205	SGRSNIG NNFVS	388	TCTGGAAGCCGCTCCAACATTGGTAATAATTTTGT ATCC
206	SGGSSNIGS NFVS	389	TCTGGAGGCAGCTCCAACATTGGGAGTAATTTTG TATCC
207	SGSSSNIGY NYVS	390	TCTGGAAGCAGCTCCAACATTGGGTATAATTATG TATCC
208	SGTSSNIEN NYVS	391	TCTGGAACCAGCTCGAACATTGAGAACAATTATG TATCC
209	SGSSSNIGN YYVS	392	TCTGGAAGTAGCTCCAACATTGGGAATTATTATG TATCC
210	SGSTSNIG KNYVS	393	TCTGGAAGCACCTCCAACATTGGGAAGAATTATG TATCC
211	SGSSSNIGT YYVS	394	TCTGGAAGCAGTTCCAACATTGGGACTTATTATGT CTCT
212	SGSSSNVG KNYVS	395	TCTGGAAGCAGCTCCAACGTTGGGAAAATTATG TATCT
213	SGSTSNIG DNFVS	396	TCTGGAAGCACCTCCAACATTGGGGATAATTTTG TATCC
214	SGSTSNIGT NYVS	397	TCTGGAAGCACCTCCAACATTGGAACATAATTATG TTTCC

215	SGGTSNIG NNYVS	398	TCTGGAGGTA TCTCC
216	SGSYSNIG NNYVS	399	TCTGGAAGCT TATCC
217	SGSSSNIED NYVS	400	TCTGGAAGCAGCT TATCC
218	SGSSSNIGK HYVS	401	TCTGGAAGCAGCT TATCC
219	SGSGSNIGS NYVS	402	TCCGGTTCCGGCT TCTCC
220	SGSSSNIGN NYIS	403	TCTGGAAGCAGCT TATCA
221	SGASSNIG NNYVS	404	TCTGGAGCCAGT TTTCC
222	SGRTSNIG NNYVS	405	TCTGGACGCACCT TATCC
223	SGGSSNIGS NYVS	406	TCTGGAGGCAGCT TATCC
224	SGSGSNIG NNYVS	407	TCTGGAAGCGGCT TATCC
225	SGSTSNIGS NYVS	408	TCTGGAAGCACCT TATCC
226	SGSSSSIGN NYVA	409	TCTGGAAGCAGCT TGGCG
227	SGSSSNLIG NNYVS	410	TCTGGAAGCAGT TATCC
228	SGTSSNIG KNYVS	411	TCTGGAACCAGCT TATCC
229	SGSSSDIGN KYIS	412	TCTGGAAGCAGCT TATCC
230	SGSSSNIGS NYIS	413	TCTGGAAGCAGCT TATCC
231	SGSTSNIG ANYVS	414	TCTGGAAGCACCT TGTCC

232	SGSSSNIGN KYVS	415	TCTGGAAGCAGCTCCAACATTGGGAATAAGTATG TATCC
233	SGSSSNIGN NYGS	416	TCTGGAAGCAGCTCCAACATTGGGAATAATTATG GATCC
234	SGSTSNIA NNYVS	417	TCTGGAAGCACCTCCAACATTGCGAATAATTATG TATCC
235	SGSYSNIGS NYVS	418	TCTGGAAGCTACTCCAATATTGGGAGTAATTATG TATCC
236	SGSSSNIGS NFVS	419	TCTGGAAGCAGCTCCAACATTGGGAGTAATTTTG TATCC
237	SGSSSNLE NNYVS	420	TCTGGAAGCAGCTCCAATCTTGAGAATAATTATG TATCC
238	SGSISNIGS NYVS	421	TCTGGAAGCATCTCCAATATTGGCAGTAATTATGT ATCC
239	SGSSSDIGS NYVS	422	TCTGGAAGCAGCTCCGACATTGGGAGTAATTATG TATCC
240	SGSSSNIGT NYVS	423	TCTGGAAGCAGCTCCAACATTGGGACTAATTATG TATCC
241	SGSSSNIGK NFVS	424	TCTGGAAGCAGCTCCAACATTGGGAAGAATTTTG TATCC
242	SGSSSNIGN NFIS	425	TCTGGAAGCAGCTCCAACATTGGGAATAATTTTA TATCC
243	SGGSSNIG NNYVS	426	TCTGGAGGCAGCTCCAACATTGGCAATAATTATG TTTCC
244	SGSSSNIGE NYVS	427	TCTGGAAGCAGCTCCAACATTGGGGAGAATTATG TATCC
245	SGSSSNIGN NFVA	428	TCTGGAAGCAGCTCCAATATTGGGAATAATTTTG TGGCC
246	SGGSSNIG NNYVA	429	TCTGGAGGCAGCTCCAACATTGGGAATAATTATG TAGCC
247	SGSSSHIGN NYVS	430	TCTGGAAGCAGCTCCCACATTGGAAATAATTATG TATCC
248	SGSSSNIGS NDVS	431	TCTGGAAGCAGCTCCAATATTGGAAGTAATGATG TATCG

249	SGSSSNIGN NYVT	432	TCTGGAAGCAGCTCCAACATTGGGAATAATTATG TAACC
250	SGSSSNIGN NPVS	433	TCTGGAAGCAGCTCCAACATTGGGAATAATCCTG TATCC
251	SGGSSNIG NHYVS	434	TCTGGAGGCAGCTCCAATATTGGGAATCATTATG TATCC
252	SGTSSNIG NNYVS	435	TCTGGAACCAGCTCCAACATTGGGAATAATTATG TATCC
253	SGSSSNIGS NYVS	436	TCTGGAAGCAGCTCCAACATTGGAAGTAATTATG TCTCG
254	SGGTSNIG SNYVS	437	TCTGGAGGCACCTCCAACATTGGAAGTAATTATG TATCC
255	SGSKSNIG NNYVS	438	TCTGGAAGCAAGTCCAACATTGGGAATAATTATG TATCC
256	SGRSSNIG NNYVS	439	TCTGGAAGAAGCTCCAACATTGGGAATAATTATG TATCG
257	SGSSSNVG SNYVS	440	TCTGGAAGCAGCTCCAACGTTGGGAGTAATTATG TTTCC
258	SGSTSNIG NNFVS	441	TCTGGAAGCACCTCCAATATTGGGAATAATTTTGT ATCC
259	SGSNFNIG NNYVS	442	TCTGGAAGCAACTTCAACATTGGGAATAATTATG TCTCC
260	SGSTSNIG YNYVS	443	TCTGGAAGCACCTCCAATATTGGATATAATTATGT ATCC
261	SGSSSNIVS NYVS	444	TCTGGAAGCAGCTCCAATATTGTAAGTAATTATG TATCC
262	SGTSSNIG NNFVS	445	TCTGGAACCAGCTCCAACATTGGGAATAATTTTG TATCC
263	SGSSSNIGR NFVS	446	TCTGGAAGCAGCTCCAACATTGGGAGGAATTTTG TGTCC
264	SGTTSNIG NNYVS	447	TCTGGAACGACCTCCAACATTGGGAATAATTATG TCTCC
265	SGSSSNIGN NDVS	448	TCTGGAAGCAGCTCCAACATTGGGAATAATGATG TATCC

266	SGSSSNIGN HDVS	449	TCTGGAAGCAGCTCCAACATTGGGAATCATGATG TATCC
267	SGSSSNIGS SHVS	450	TCTGGAAGCAGCTCCAACATTGGAAGTAGTCATG TATCC
268	SGSSSNIGI HYVS	451	TCTGGAAGCAGCTCCAACATTGGGATTCATTATG TATCC
269	SGGGSNIG YNYVS	452	TCTGGAGGCGGCTCCAACATTGGCTATAATTATG TCTCC
270	SGSSSNIGD HYVS	453	TCTGGAAGCAGCTCCAACATTGGGGATCATTATG TGTCG
271	SGSSSNLG KNYVS	454	TCTGGAAGCAGCTCCAACCTTGGGAAGAATTATG TATCT
272	SGSSSNIGD NFVS	455	TCTGGAAGCAGCTCCAACATTGGCGATAATTTTG TATCC
273	SGSTSNIK NYVS	456	TCTGGAAGCACCTCCAACATTGAGAAAACTATG TATCG
274	SGSSSNIGK DYVS	457	TCTGGAAGCAGCTCCAACATTGGGAAGGATTATG TATCC
275	SGSSSNIGK NYVS	458	TCTGGAAGCAGCTCCAACATTGGGAAGAATTATG TATCC
276	SGSSSNIGN NYVS	459	TCTGGAAGCAGCTCCAACATTGGGAATAATTATG TATCC
277	SGSSSNIGN NYAS	460	TCTGGAAGCAGCTCCAACATTGGGAATAATTATG CCTCC
278	SGISSNIGN NYVS	461	TCTGGAATCAGCTCCAACATTGGGAATAATTATG TATCC
279	TGSSSNIG NNYVS	462	ACTGGAAGCAGCTCCAACATTGGGAATAATTATG TATCC
280	SGTSSNIG NNHVS	463	TCTGGAACCAGCTCCAACATTGGGAATAATCATG TTTCC
281	SGSRSNIG KNYVS	464	TCTGGAAGTCGTTCCAACATTGGGAAAAATTATG TATCC
IGLV1-51-L2			
465	DNNKRPP	616	GACAATAATAAGCGACCCCA

466	ENNRPS	617	GAGAATAATAGGCGACCCTCA
467	DNNKQPS	618	GACAATAATAAGCAACCCTCA
468	DNNKRPL	619	GACAATAACAAGCGACCCTTG
469	DNDKRPA	620	GACAATGATAAGCGACCCGCA
470	DNHERPS	621	GACAATCATGAGCGACCCTCA
471	ENRKRPS	622	GAAAACCGTAAGCGACCCTCA
472	DNDQRPS	623	GACAATGATCAGCGACCCTCA
473	ENYKRPS	624	GAGAATTATAAGCGACCCTCA
474	ENTKRPS	625	GAAAATACTAAGCGACCCTCA
475	DTEKRPS	626	GACACTGAGAAGAGGCCCTCA
476	DNDKRPP	627	GACAATGATAAGCGACCCCCA
477	DHNRPS	628	GACCATAATAAGCGACCCTCA
478	GNNERPS	629	GGCAATAATGAGCGACCCTCA
479	DTSKRPS	630	GACACTAGTAAGCGACCCTCA
480	EYNKRPS	631	GAATATAATAAGCGCCCCTCA
481	ENIKRPS	632	GAAAATATTAAGCGACCCTCA
482	DNVKRPS	633	GACAATGTTAAGCGACCCTCA
483	ENDKRSS	634	GAAAACGATAAACGATCCTCA
484	ENNRHS	635	GAAAATAATAAGCGACACTCA
485	GNDQRPS	636	GGAAATGATCAGCGACCCTCA
486	DNDRRPS	637	GACAATGATAGGCGACCCTCA
487	DNHKRPS	638	GACAATCATAAGCGGCCCTCA
488	DNNDRPS	639	GACAATAATGACCGACCCTCA
489	ENNQRPS	640	GAGAATAATCAGCGACCCTCA
490	DNNQRPS	641	GACAATAATCAGCGACCCTCA
491	ENVKRPS	642	GAGAATGTTAAGCGACCCTCA
492	DTYKRPS	643	GCACTTATAAGAGACCCTCA
493	NNNNRPS	644	AACAATAATAACCGACCCTCA
494	GNNNRPS	645	GGCAATAATAATCGACCCTCA
495	ENDQRPS	646	GAAAATGATCAGCGACCCTCA
496	DNNKRAS	647	GACAATAATAAGCGAGCCTCA
497	DNDKRPL	648	GACAATGATAAGCGACCCTTA
498	DTDERPS	649	GCACTGATGAGCGACCTTCA
499	DNRKRPS	650	GACAATAGGAAGCGACCCTCA

500	DNDARPS	651	GACAATGATGCTCGACCCTCA
501	DNNKRLS	652	GACAATAATAAGCGACTCTCA
502	DNDKRAS	653	GACAATGATAAGCGAGCCTCA
503	DNTERPS	654	GACAATACTGAGCGACCCTCA
504	DNNIRPS	655	GACAATAATATTCGACCCTCA
505	DNKRRPS	656	GACAATAAGAGGCGACCCTCA
506	DDNNRPS	657	GACGATAATAACCGACCCTCA
507	ANNRRPS	658	GCGAATAATCGACGACCCTCA
508	DNDKRLS	659	GACAATGATAAGCGACTGTCA
509	DNNKRPA	660	GACAATAATAAGCGACCCGCA
510	DNYRRPS	661	GACAATTATAGACGTCCCTCA
511	ANDQRPS	662	GCCAATGATCAGCGACCCTCA
512	DNDKRRS	663	GACAATGATAAGCGACGCTCA
513	DKNERPS	664	GACAAGAATGAGCGACCCTCA
514	DNKERPS	665	GACAATAAGGAGCGACCCTCA
515	DNNKGPS	666	GACAATAATAAGGGACCCTCA
516	ENDRRPS	667	GAAAATGATAGACGACCCTCA
517	ENDERPS	668	GAAAATGATGAGCGACCCTCA
518	QNNKRPS	669	CAAATAATAAGCGACCCTCA
519	DNRERPS	670	GACAATCGTGAGCGACCCTCA
520	DNNRRPS	671	GACAATAATAGACGACCCTCA
521	GNNRRPS	672	GGAAATAATAGGCGACCCTCA
522	DNDNRPS	673	GACAATGATAACCGACCCTCA
523	EDNKRPS	674	GAAGATAATAAGCGACCCTCA
524	DDDERPS	675	GACGATGATGAGCGGCCCTCA
525	ASNKRPS	676	GCAAGTAATAAGCGACCCTCA
526	DNNKRSS	677	GACAATAATAAGCGATCCTCA
527	QNNERPS	678	CAAATAATGAGCGACCCTCA
528	DDDRRPS	679	GACGATGATAGGCGACCCTCA
529	NNDKRPS	680	AACAATGATAAGCGACCCTCA
530	DNNNRPS	681	GACAATAATAACCGACCCTCA
531	DNNVRPS	682	GACAATAATGTGCGACCCTCA
532	ENNERPS	683	GAAAATAATGAGCGACCCTCA
533	DNNHRPS	684	GACAATAATCACCGACCCTCA

534	DNDERPS	685	GACAATGATGAGCGCCCCTCG
535	DNIRRPS	686	GACAATATCCGGCGACCCTCA
536	DFNKRPS	687	GACTTTAATAAGCGACCCTCA
537	ETNKRPS	688	GAAACTAATAAGCGACCCTCA
538	NDNKRPS	689	AACGATAATAAGCGACCCTCA
539	DDNKRPS	690	GACGATAATAAGCGACCCTCA
540	DNYKRPS	691	GACAATTATAAGCGACCCTCA
541	HNNKRPS	692	CACAATAATAAGCGACCCTCA
542	DNHQRPS	693	GACAATCATCAGCGACCCTCA
543	DNYKRAS	694	GACAATTATAAGCGAGCCTCA
544	DNIKRPS	695	GACAATATTAAGCGACCCTCA
545	DTHKRPS	696	GACACTCATAAGCGACCCTCA
546	DTNRRPS	697	GACACTAATAGGCGACCCTCT
547	DTNQRPS	698	GACACTAATCAGCGACCCTCA
548	ESDKRPS	699	GAAAGTGATAAGCGACCCTCA
549	DNDKRSS	700	GACAATGATAAGCGATCTTCG
550	GSNKRPS	701	GGCAGTAATAAGCGACCCTCA
551	DNNKRVS	702	GACAATAACAAGCGAGTTTCA
552	NNNRRPS	703	AACAATAATAGGCGACCCTCA
553	DNFKRPS	704	GACAATTTTAAGCGACCCTCA
554	ENDKRPS	705	GAAAATGATAAACGACCCTCA
555	ENNKRLS	706	GAAAATAATAAGCGACTCTCA
556	ADNKRPS	707	GCAGATAATAAGCGACCCTCA
557	EDNERPS	708	GAAGATAATGAGCGCCCCTCA
558	DTDQRPS	709	GACACTGATCAGCGACCCTCA
559	DNYQRPS	710	GACAATTATCAGCGACCCTCA
560	DENKRPS	711	GACGAGAATAAGCGACCCTCA
561	DTNKRPS	712	GACACTAATAAGCGACCCTCA
562	DDYRRPS	713	GACGATTATCGGCGACCCTCA
563	DNDKRHS	714	GACAACGATAAGCGGCACTCA
564	ENDNRPS	715	GAAAATGATAATCGACCCTCA
565	DDNERPS	716	GACGATAATGAGCGCCCCTCA
566	DNKKRPS	717	GACAATAAGAAGCGACCCTCA
567	DVDKRPS	718	GACGTTGATAAGCGACCCTCA

568	ENKKRPS	719	GAAAATAAAAAACGACCCTCT
569	VNDKRPS	720	GTCAATGATAAGCGACCCTCA
570	DNDHRPS	721	GACAATGATCACCGACCCTCA
571	DINKRPS	722	GACATTAATAAGCGACCCTCA
572	ANNERPS	723	GCCAATAATGAGCGACCCTCA
573	DNENRPS	724	GACAATGAAAACCGACCGTCA
574	GDDKRPS	725	GGCGATGATAAGCGACCCTCA
575	ANNQRPS	726	GCCAATAATCAGCGACCTTCA
576	DDDKRPS	727	GACGATGATAAGCGACCCTCA
577	YNNKRPS	728	TACAATAATAAGCGGCCCTCA
578	EDDKRPS	729	GAAGATGATAAGCGACCCTCA
579	ENNNRPS	730	GAAAACAATAACCGACCCTCG
580	DNNLRPS	731	GACAATAATCTGCGACCCTCA
581	ESNKRPS	732	GAGAGTAACAAGCGACCCTCA
582	DTDKRPS	733	GACACTGATAAGCGGCCCTCA
583	DDDQRPS	734	GACGATGATCAGCGACCCTCA
584	VNNKRPS	735	GTGAATAATAAGAGACCCTCC
585	DDYKRPS	736	GACGATTATAAGCGACCCTCA
586	DNTKRPS	737	GACAATACTAAGCGACCCTCA
587	DDTERPS	738	GACGATACTGAGCGACCCTCA
588	GNDKRPS	739	GGCAATGATAAGCGACCCTCA
589	DNEKRPS	740	GACAATGAAAAGCGACCCTCA
590	DNDDRPS	741	GACAATGATGACCGACCCTCA
591	DDNRRPS	742	GACGATAATAGGCGTCCCTCA
592	GNNKRPS	743	GGCAATAATAAGCGACCCTCA
593	ANDKRPS	744	GCCAATGATAAGCGACCCTCA
594	DNNKRHS	745	GACAATAATAAGCGACACTCA
595	DDNQRPS	746	GACGACAATCAGCGACCCTCA
596	GNDRRPS	747	GGCAATGATAGGCGACCCTCA
597	DNHNRPS	748	GACAATCATAACCGACCCTCA
598	DNYERPS	749	GACAATTATGAGCGACCCTCA
599	ENNKRSS	750	GAAAATAATAAGCGATCCTCA
600	DDHKRPS	751	GACGATCATAAGCGGCCCTCA
601	DNNKRRS	752	GACAATAATAAACGACGTTCA

602	DNDKRPS	753	GACAATGATAAGCGACCGTCA
603	DKNKRPS	754	GACAAGAATAAGCGACCCTCA
604	DNNKRPS	755	GACAATAATAAGCGACCCTCA
605	DIDKRPS	756	GACATTGATAAGCGACCCTCA
606	DDKKRPS	757	GACGATAAGAAGCGACCCTCA
607	ANNKRPS	758	GCCAATAATAAGCGACCCTCA
608	DNDKGPS	759	GACAATGATAAGGGACCCTCA
609	EDNRRPS	760	GAAGATAATAGGCGACCCTCA
610	ENNKRPS	761	GAGAATAATAAGCGACCCTCA
611	NNNKRPS	762	AACAATAATAAGCGACCCTCA
612	DNNRPS	763	GACAATAATGAGCGACCCTCA
613	DNIQRPS	764	GACAATATTCAGCGACCCTCA
614	DNNYRPS	765	GACAATAATTACCGACCCTCA
615	DNYNRPS	766	GACAATTATAACCGACCCTCA
IGLV1-51-L3			
767	CGTWDTSL SAVVF	1591	TGCGGAACATGGGATACCAGCCTGAGTGCTGTGG TG TTC
768	CGTWDTSL SAGVF	1592	TGCGGAACATGGGATACCAGCCTGAGTGCTGGGG TG TTC
769	CGTWDTSL SAWVF	1593	TGCGGAACATGGGATACCAGCCTGAGTGCTTGGG TG TTC
770	CGTWDRS LSAGVF	1594	TGCGGAACATGGGATAGGAGCCTGAGTGCGGGG GTG TTC
771	CGTWDRS LSAWVF	1595	TGCGGAACATGGGATAGGAGCCTGAGTGCTTGGG TATTT
772	CGTWDTSL SGGVF	1596	TGCGGAACATGGGATACCAGCCTGAGTGGTGGGG TG TTC
773	CGTWDTSL RAGVF	1597	TGCGGAACATGGGATACTAGCCTGCGTGCTGGCG TCTTC
774	CGTWDRS LSVWVF	1598	TGCGGAACATGGGATAGGAGCCTGAGTGTTTGGG TG TTC
775	CGTWDTSL SVVVF	1599	TGCGGAACATGGGATACCAGTCTGAGTGTTGTGG TCTTC

776	CGTWDTSL SAAVF	1600	TGCGGAACGTGGGATAACCAGCCTGAGTGCTGCGG TGTTTC
777	CGAWDTS LSAGVF	1601	TGCGGAGCATGGGATAACCAGCCTGAGTGCTGGAG TGTTTC
778	CATWDTSL SAVVF	1602	TGCGCAACATGGGATAACCAGCCTGAGTGCTGTGG TATTC
779	CATWDTSL SAGVF	1603	TGCGCAACATGGGATAACCAGCCTGAGTGCTGGTG TGTTTC
780	CGTWESSL SAWVF	1604	TGTGGAACATGGGAGAGCAGCCTGAGTGCTTGGG TGTTTC
781	CGTWDTT LSAGVF	1605	TGCGGAACATGGGATAACCACCCTGAGTGCGGGTG TCTTC
782	CGTWDTSL SVWVF	1606	TGCGGAACATGGGATACTAGCCTGAGTGTGTGGG TGTTTC
783	CGTWDTSL SVGVF	1607	TGCGGAACATGGGATACTAGCCTGAGTGTTGGGG TGTTTC
784	CGTWDTSL STGVF	1608	TGCGGAACATGGGACACCAGTCTGAGCACTGGCG TCTTC
785	CGTWDTSL SGVVF	1609	TGCGGAACATGGGATAACCAGCCTGAGTGGTGTGG TCTTC
786	CGTWDTSL SAYVF	1610	TGCGGAACATGGGATAACCAGCCTGAGTGCTTATG TCTTC
787	CGTWDTSL SAEVF	1611	TGCGGAACATGGGATAACCAGCCTGAGTGCTGAGG TGTTTC
788	CGTWDTG LSAGVF	1612	TGCGGAACATGGGATAACCGCCTGAGTGCTGGGG TATTC
789	CGTWDRS LSAYVF	1613	TGCGGAACGTGGGATAGGAGCCTGAGTGCTTATG TCTTC
790	CGTWDRS LSAVVF	1614	TGCGGAACATGGGATAGGAGCCTCAGTGCCGTGG TATTC
791	CGTWDNT LSAWVF	1615	TGCGGAACATGGGATAACACCCTGAGTGCGTGGG TGTTTC
792	CGTWDNR LSAGVF	1616	TGCGGAACATGGGATAACAGGCTGAGTGCTGGGG TGTTTC

793	CGTWDISL SAWVF	1617	TGCGGAACATGGGACATCAGCCTGAGTGCTTGGG TG TTC
794	CGTWHSSL SAGVF	1618	TGCGGAACATGGCATAGCAGCCTGAGTGCTGGGG TATTC
795	CGTWGSSL SAWVF	1619	TGCGGAACATGGGGTAGCAGTTTGAGTGCTTGGG TG TTC
796	CGTWESSL SGWVF	1620	TGCGGAACATGGGAGAGCAGCCTGAGTG GTTGGG TG TTC
797	CGTWESSL SAVVF	1621	TGCGGAACATGGGAGAGCAGCCTGAGTGCTGTGG TTTTTC
798	CGTWDYS LSAVVF	1622	TGCGGAACATGGGATTACAGCCTGAGTGCTGTGG TATTC
799	CGTWDYS LSAGVF	1623	TGCGGAACATGGGATTACAGCCTGAGTGCTGGGG TATTC
800	CGTWDVS LSVGVF	1624	TGCGGAACATGGGATGTCAGCCTGAGTGTTGGAG TG TTC
801	CGTWDTT LSAVVF	1625	TGCGGAACATGGGATAACCACCCTGAGTGCTGTGG TTTTTC
802	CGTWDTT LNIGVF	1626	TGCGGAACATGGGATACTCTGAATATTGGGG TG TTC
803	CGTWDTSL TAVVF	1627	TGCGGAACATGGGATAACCAGCCTGACTGCTGTGG TATTC
804	CGTWDTSL TAAVF	1628	TGCGGAACCTGGGATAACCAGCCTGACTGCTGCTG TG TTC
805	CGTWDTSL SVGLF	1629	TGCGGCACATGGGATAACCAGCCTGAGTGTGGGGC TATTC
806	CGTWDTSL SGRVF	1630	TGCGGAACCTGGGATAACCAGCCTGAGTGGTAGGG TG TTC
807	CGTWDTSL SGAVF	1631	TGCGGAACATGGGATAACCAGCCTGAGTGGTGCAG TG TTC
808	CGTWDTSL SAGLF	1632	TGCGGAACATGGGATAACCAGCCTGAGTGCTGGCC TG TTC
809	CGTWDTSL SAGGVF	1633	TGCGGAACATGGGATAACCAGCCTGAGTGCTGGAG GGGTCTTC

810	CGTWDTSL RAYVF	1634	TGCGGAACATGGGATAACCAGCCTGCGTGCTTATG TCTTC
811	CGTWDTSL RAWVF	1635	TGCGGAACATGGGATACTAGTTTTCGCTGCTTGGG TATTC
812	CGTWDTSL NTGVF	1636	TGCGGAACATGGGATAACCAGCCTGAATACTGGGG TATTC
813	CGTWDTSL NIWVF	1637	TGCGGAACATGGGATAACCAGCCTGAATATTTGGG TGTTTC
814	CGTWDTSL NIGVF	1638	TGCGGAACATGGGATAACAAGCCTGAATATTGGGG TGTTTC
815	CGTWDTSL IAVVF	1639	TGCGGAACATGGGATAACCAGCCTGATTGCTGTGG TGTTTC
816	CGTWDRS LSGWVF	1640	TGCGGAACGTGGGATAGGAGCCTGAGTGGTTGGG TGTTTC
817	CGTWDNR LSGWVF	1641	TGCGGAACATGGGATAACAGGCTGAGTGGTTGGG TGTTTC
818	CGTWDKS LSAVVF	1642	TGCGGAACGTGGGATAAGAGCCTGAGTGCTGTGG TCTTC
819	CGTWDKG LSAWVF	1643	TGCGGAACATGGGATAAAGGCCTGAGTGCTTGGG TGTTTC
820	CGTWDISL SAGVF	1644	TGCGGAACATGGGATATCAGCCTGAGTGCTGGGG TGTTTC
821	CGTWDES� SGGEVVF	1645	TGCGGAACATGGGATGAGAGCCTGAGTGGTGGCG AGGTGGTCTTC
822	CGTWDAS LSAWVF	1646	TGCGGAACATGGGATGCCAGCCTGAGTGCCTGGG TGTTTC
823	CGTWDAG LSAWVF	1647	TGCGGAACTTGGGATGCCGGCCTGAGTGCTTGGG TGTTTC
824	CGAWDTS LSAWVF	1648	TGCGGAGCATGGGATAACCAGCCTGAGTGCTTGGG TGTTTC
825	CGAWDTS LSAVVF	1649	TGCGGAGCATGGGATAACCAGCCTGAGTGCTGTGG TGTTTC
826	CGAWDTS LRAGVF	1650	TGCGGAGCATGGGATAACCAGCCTGCGTGCTGGGG TTTTTC

827	CATWDTS VSAWVF	1651	TGCGCAACATGGGATAACCAGCGTGAGTGCTTGGG TGTTTC
828	CATWDTSL SAWVF	1652	TGCGCAACATGGGATAACCAGCCTGAGTGCGTGGG TGTTTC
829	CATWDNT LSAGVF	1653	TGCGCAACATGGGACAACACCCTGAGTGCTGGGG TGTTTC
830	CAAWDRS LSVWVF	1654	TGCGCAGCATGGGATAGGAGCCTGAGTGTTTGGG TGTTTC
831	CYTWHSSL RGGVF	1655	TGCTACACATGGCATTCCAGTCTGCGTGGTGGGG TGTTTC
832	CVTWTSSP SAWVF	1656	TGCGTAACGTGGACTAGTAGCCCGAGTGCTTGGG TGTTTC
833	CVTWRGG LVLF	1657	TGCGTGACATGGCGTGGTGGCCTTGTGTTGTTTC
834	CVTWDTSL TSVVL	1658	TGCGTAACATGGGATAACCAGCCTGACTTCTGTGG TACTC
835	CVTWDTSL SVYWVF	1659	TGCGTAACATGGGATAACCAGCCTGAGTGTTTATT GGGTGTTTC
836	CVTWDTSL SAWVF	1660	TGCGTTACATGGGATAACCAGCCTGAGTGCCTGGG TGTTTC
837	CVTWDTD LSVALF	1661	TGCGTCACATGGGATAACCGACCTCAGCGTTGCGC TCTTC
838	CVTWDRS LSGWVF	1662	TGCGTAACATGGGATAGGAGCCTGAGTGGTTGGG TGTTTC
839	CVTWDRS LREVLf	1663	TGCGTAACATGGGATCGCAGCCTGAGAGAGGTGT TATTC
840	CVTWDRS LRAVVF	1664	TGCGTAACATGGGATCGCAGCCTGAGAGCGGTGG TATTC
841	CVTWDRS LDAGVF	1665	TGCGTAACATGGGACAGGAGCCTCGATGCTGGGG TTTTTC
842	CVTWDNT LSAGVF	1666	TGCGTGACATGGGATAACACCCTGAGTGCTGGGG TCTTC
843	CVTWDNN LFGVVF	1667	TGCGTAACATGGGATAACAACCTGTTTGGTGTGG TCTTC

844	CVSWDTSL SGAVF	1668	TGCGTATCATGGGATAACCAGCCTGAGTGGTGCGG TATTC
845	CVSWDTSL SAGVF	1669	TGCGTCTCATGGGATAACCAGCCTGAGTGCTGGGG TATTC
846	CTTWFRTP SDVVF	1670	TGCACAACATGGTTTAGGACTCCGAGTGATGTGG TCTTC
847	CTTWFRTA SDVVF	1671	TGCACAACATGGTTTAGGACTGCGAGTGATGTGG TCTTC
848	CTTWDYD LSVVF	1672	TGCACAACGTGGGATTACGGTCTGAGTGTCGTCT TC
849	CTARDTSL SPGGVF	1673	TGCACAGCAAGGGATAACCAGCCTGAGTCCTGGCG GGGTCTTC
850	CSTWNTRP SDVVF	1674	TGCTCAACATGGAATACGAGGCCGAGTGATGTGG TGTTTC
851	CSTWESSL TTVVF	1675	TGTTCAACATGGGAGAGCAGTTTGACTACTGTGG TCTTC
852	CSTWDTSL TNVLF	1676	TGCTCAACATGGGATAACCAGCCTCACTAATGTGC TATTC
853	CSTWDTSL SGVVF	1677	TGCTCAACATGGGATAACCAGCCTGAGTGGAGTAG TCTTC
854	CSTWDHSL KAALF	1678	TGCTCAACATGGGATCACAGCCTGAAAGCTGCAC TGTTTC
855	CSTWDAR LSVRVF	1679	TGCTCAACCTGGGATGCGAGGCTGAGTGTCGGG TGTTTC
856	CSSYTSSST WVF	1680	TGCTCCTCATATACAAGCAGCAGCACTTGGGTGT TC
857	CSSYATRG LRVLF	1681	TGCAGCTCATAACGCAACCCGCGGCCTTCGTGTGTT GTTC
858	CSSWDATL SVRIF	1682	TGTTTCATCATGGGACGCCACCCTGAGTGTTTCGCAT ATTC
859	CQVWEGS SDHWVF	1683	TGTCAGGTGTGGGAGGGTAGTAGTGATCATTGGG TGTTTC
860	CQTWDNR LSAVVF	1684	TGCCAAACCTGGGATAACAGACTGAGTGCTGTGG TGTTTC

861	CQTDWHS LHVGVF	1685	TGTCAAACGTGGGATCACAGCCTGCATGTTGGGG TGTTTC
862	CQSYDDIL NVWVL	1686	TGCCAGTCCTATGACGACATCTTGAATGTTTGGGT CCTT
863	CNTWDKS LTSELF	1687	TGCAATACATGGGATAAGAGTTTGACTTCTGAAC TCTTC
864	CLTWDRSL NVRVF	1688	TGCTTAACATGGGATCGCAGCCTGAATGTGAGGG TGTTTC
865	CLTWDHSL TAYVF	1689	TGCCTAACATGGGACCACAGCCTGACTGCTTATG TCTTC
866	CLTRDTSL SAPVF	1690	TGCTTAACAAGGGATAACCAGTCTGAGTGCCCCTG TGTTTC
867	CKTWESGL NFGHVF	1691	TGCAAAACATGGGAAAGTGGCCTTAATTTTGGCC ACGTCTTC
868	CKTWDTSL SAVVF	1692	TGCAAAACATGGGATAACCAGCCTGAGTGCTGTGG TCTTC
869	CGVWDVS LGAGVF	1693	TGCGGAGTCTGGGATGTCAGTCTGGGTGCTGGGG TGTTTC
870	CGVWDTT PSAVLF	1694	TGCGGAGTCTGGGATAACCACCCCGAGTGCCGTTT TTTTTC
871	CGVWDTT LSAVLF	1695	TGCGGAGTCTGGGATAACCACCTGAGTGCCGTTT TTTTTC
872	CGVWDTS LGVF	1696	TGCGGAGTATGGGATAACCAGCCTGGGGGTCTTC
873	CGVWDTN LGKWF	1697	TGCGGGGTATGGGATAACCAACCTGGGTAAATGGG TTTTTC
874	CGVWDTG LDAGWVF	1698	TGTGGAGTTTGGGATACTGGCCTGGATGCTGGTT GGGTGTTT
875	CGVWDNV LEAYVF	1699	TGCGGAGTGTGGGATAACGTCCTGGAGGCCTATG TCTTC
876	CGVWDISL SANWVF	1700	TGCGGAGTCTGGGATATCAGCCTGAGTGCTAATT GGGTGTTT
877	CGVWDHS LGIWAF	1701	TGCGGAGTATGGGATCACAGCCTGGGGATTTGGG CCTTC

878	CGVWDDIL TAEVF	1702	TGCGGAGTTTGGGATGATATTCTGACTGCTGAAG TG TTC
879	CGVRDTSL GVF	1703	TGCGGAGTTCGGGATACCAGCCTGGGGGTCTTC
880	CGTYDTSL PAWVF	1704	TGCGGAACATACGATACGAGCCTGCCTGCTTGGG TG TTT
881	CGTYDNL VFGYVF	1705	TGCGGAACTTACGATAATCTTGTATTTGGTTATGT CTTC
882	CGTYDDRL REVF	1706	TGCGGAACATACGATGATAGACTCAGAGAGGTGT TC
883	CGTWVTSL SAGVF	1707	TGCGGAACGTGGGTACCAGCCTGAGTGCTGGGG TG TTC
884	CGTWVSSL TTVVF	1708	TGCGGAACATGGGTAGCAGCCTGACTACTGTAG TATTC
885	CGTWVSSL NVWVF	1709	TGCGGAACATGGGTAGCAGCCTGAACGTCTGGG TG TTC
886	CGTWVGR FWVF	1710	TGCGGAACATGGGTGGCAGGTTTTGGGTATTC
887	CGTWSGG PSGHWLF	1711	TGCGGAACATGGTCTGGCGGCCCGAGTGGCCATT GGTTG TTC
888	CGTWSGG LSGHWLF	1712	TGCGGAACATGGTCTGGCGGCCTGAGTGGCCATT GGTTG TTC
889	CGTWQTG REAVLF	1713	TGCGGAACGTGGCAGACCGGCCGGGAGGCTGTCC TATTT
890	CGTWQSR LRWVF	1714	TGCGGAACGTGGCAGAGCAGGCTGAGGTGGGTGT TC
891	CGTWQSR LGWVF	1715	TGCGGAACGTGGCAGAGCAGGCTGGGGTGGGTGT TC
892	CGTWPRSL SAVWVF	1716	TGCGGAACATGGCCTAGGAGCCTGAGTGCTGTTT GGGTG TTC
893	CGTWNNY LSAGDVVF	1717	TGCGGAACATGGAATAACTACCTGAGTGCTGGCG ATGTGGTTTTTC
894	CGTWLGS QSPYWVF	1718	TGCGGAACATGGCTTGGCAGCCAGAGTCCTTATT GGGTCTTC

895	CGTWHTG LSAYVF	1719	TGCGGAACATGGCATAACCGGCCTGAGTGCTTATG TCTTC
896	CGTWHSTL SAGHWVF	1720	TGCGGAACATGGCATAGTACCCTGAGTGCTGGCC ATTGGGTGTTC
897	CGTWHSSL STWVF	1721	TGCGGAACATGGCATAGTAGCCTGAGTACTTGGG TGTTTC
898	CGTWHSSL SAYVF	1722	TGCGGAACATGGCATAGCAGCCTGAGTGCCTATG TCTTC
899	CGTWHSSL SAVVF	1723	TGCGGAACATGGCATAGCAGCCTGAGTGCTGTGG TATTC
900	CGTWHSG LSGWVF	1724	TGCGGAACGTGGCATTCCGGCCTGAGTGGGTGGG TTTTTC
901	CGTWHNT LRNVIF	1725	TGCGGAACATGGCATAACACCCTGCGTAATGTGA TATTC
902	CGTWHAS LTAVF	1726	TGCGGAACATGGCATGCCAGCCTGACTGCTGTGT TC
903	CGTWGWY GSQRGVVF	1727	TGCGGGACATGGGGATGGTATGGCAGCCAGAGA GGCGTCGTCTTC
904	CGTWGWY GGQRGVV F	1728	TGCGGGACATGGGGATGGTATGGCGGCCAGAGA GGCGTCGTCTTC
905	CGTWGTSL SAWVF	1729	TGCGGAACCTGGGGAACCAGCCTGAGTGCTTGGG TGTTTC
906	CGTWGSSL TTGLF	1730	TGCGGAACCTGGGGTAGCAGCCTGACTACTGGCC TGTTTC
907	CGTWGSSL TAYVF	1731	TGCGGAACATGGGGTAGCAGCCTGACTGCCTATG TCTTC
908	CGTWGSSL SVVF	1732	TGCGGAACATGGGGTAGCAGCCTGAGTGTTGTGT TC
909	CGTWGSSL SGGVF	1733	TGCGGAACATGGGGTAGCAGCCTGAGTGGTGGGG TGTTTC
910	CGTWGSSL SAYWVF	1734	TGCGGAACATGGGGTAGCAGCCTGAGTGCTTATT GGGTGTTC

911	CGTWGSSL SAYVVF	1735	TGCGGAACATGGGGTAGCAGCCTGAGTGCTTATG TGGTG TTC
912	CGTWGSSL SAYVF	1736	TGCGGAACATGGGGTAGCAGCCTGAGTGCTTATG TCTTC
913	CGTWGSSL SAVVF	1737	TGCGGAACGTGGGGTAGTAGCCTGAGTGCTGTGG TG TTC
914	CGTWGSSL SAPYVF	1738	TGCGGAACATGGGGTAGCAGCCTGAGTGCTCCTT ATGTCTTC
915	CGTWGSSL SAPVF	1739	TGCGGAACATGGGGTAGCAGCCTGAGTGCCCCGG TG TTC
916	CGTWGSSL SAGVF	1740	TGCGGAACATGGGGTAGCAGCCTGAGTGCTGGGG TG TTC
917	CGTWGSSL SAGLF	1741	TGCGGAACTTGGGGTAGCAGCCTGAGTGCTGGAC TG TTC
918	CGTWGSSL SAGALF	1742	TGCGGAACATGGGGTAGCAGCCTGAGTGCTGGGG CACTCTTC
919	CGTWGSSL RAWVF	1743	TGCGGAACATGGGGCAGTAGCCTGCGTGCTTGGG TG TTC
920	CGTWFTSL ASGVF	1744	TGCGGAACTTGGTTTACTAGTCTGGCTAGTGGGG TTTTTC
921	CGTWETSL SVVVI	1745	TGCGGAACTTGGGAGACCAGTCTGAGTGTCGTGG TCATC
922	CGTWETSL SGVF	1746	TGCGGAACATGGGAGACCAGCCTGAGTGGTGTCT TC
923	CGTWETSL SDWVF	1747	TGCGGAACATGGGAAACCAGCCTGAGTGATTGGG TATTC
924	CGTWETSL SAGVF	1748	TGCGGAACATGGGAGACCAGCCTGAGTGCTGGGG TATTC
925	CGTWETSL NYVAF	1749	TGCGGAACATGGGAAACCAGCCTTAATTATGTGG CCTTC
926	CGTWETSL NTWLL	1750	TGCGGAACATGGGAGACCAGCCTGAATACTTGGT TGCTC
927	CGTWETSE SGNYIF	1751	TGCGGAACATGGGAGACCAGCGAGAGTGGTAATT ACATCTTC

928	CGTWETRL GTWVI	1752	TGCGGAACATGGGAAACCAGACTGGGTACTTGGG TGATC
929	CGTWETQ LYWVF	1753	TGCGGAACATGGGAGACCCAGTTATATTGGGTGT TC
930	CGTWETG LSAGEVF	1754	TGCGGAACATGGGAGACTGGCCTAAGTGCTGGAG AGGTGTTC
931	CGTWESTL SVFLF	1755	TGCGGAACTTGGGAAAGCACCTGAGTGTTTTCC TATTC
932	CGTWESSL TVVVF	1756	TGCGGGACATGGGAAAGTAGCCTGACTGTTGTGG TCTTC
933	CGTWESSL TGVVF	1757	TGCGGAACATGGGAAAGTAGCCTGACTGGAGTGG TATTC
934	CGTWESSL TGFVF	1758	TGCGGAACATGGGAAAGCAGCCTGACTGGTTTTG TCTTC
935	CGTWESSL SVGVF	1759	TGTGGAACATGGGAGAGCAGCCTGAGTGTTGGGG TGTTTC
936	CGTWESSL SEWVF	1760	TGCGGAACCTGGGAAAGTAGCCTCAGTGAATGGG TGTTTC
937	CGTWESSL SAVF	1761	TGCGGAACATGGGAGAGCAGCCTGAGTGCTGTAT TC
938	CGTWESSL SAGYIF	1762	TGCGGAACATGGGAGAGCAGCCTGAGTGCTGGTT ATATCTTC
939	CGTWESSL SAGVF	1763	TGCGGAACATGGGAGAGCAGCCTGAGTGCTGGA GTGTTC
940	CGTWESSL SAGPVF	1764	TGCGGAACATGGGAAAGCAGCCTGAGCGCTGGCC CGGTGTTC
941	CGTWESSL SAGGQVF	1765	TGCGGAACATGGGAAAGCAGCCTGAGTGCTGGA GGCCAGGTGTTC
942	CGTWESSL SAFGGYVF	1766	TGCGGAACATGGGAGAGCAGCCTGAGTGCCTTCG GCGGTTATGTCTTC
943	CGTWESSL RVWVF	1767	TGCGGAACATGGGAAAGCAGCCTGAGGGTTTGGG TGTTTC
944	CGTWESSL FTGPWVF	1768	TGCGGAACATGGGAAAGCAGCCTCTTACTGGGC CTTGGGTGTTC

945	CGTWESLS ATYVF	1769	TGCGGAACATGGGAGAGCCTGAGTGCCACCTATG TCTTC
946	CGTWESGL SAGVF	1770	TGCGGAACATGGGAGAGCGGCCTGAGTGCTGGTG TCTTC
947	CGTWESDF WVF	1771	TGCGGAACATGGGAAAGCGACTTTTGGGTGTTT
948	CGTWENR LSAVVF	1772	TGCGGTACATGGGAAAACAGACTGAGTGCTGTGG TCTTC
949	CGTWENR LSAGVF	1773	TGCGGAACATGGGAAAACAGACTGAGTGCCGGG GTATTC
950	CGTWEISL TTSVVF	1774	TGCGGAACATGGGAAATCAGCCTGACTACTTCTG TGGTATTC
951	CGTWEISL STSVVF	1775	TGCGGAACATGGGAAATCAGCCTGAGTACTTCTG TGGTATTC
952	CGTWEGSL SVVF	1776	TGCGGAACATGGGAAGGCAGCCTCAGTGTTGTTT TC
953	CGTWEGSL RVF	1777	TGCGGAACATGGGAAGGCAGCCTGAGGGTGTTT
954	CGTWEGSL RHVF	1778	TGCGGAACATGGGAGGGCAGCCTGAGGCACGTGT TC
955	CGTWDYS PVRAGVF	1779	TGCGGAACATGGGATTACAGCCCTGTACGTGCTG GGGTGTTC
956	CGTWDYS LSVYLF	1780	TGCGGAACGTGGGATTACAGCCTGAGTGTTTATC TCTTC
957	CGTWDYS LSSGVVF	1781	TGCGGAACATGGGATTACAGCCTGAGTTCTGGCG TGGTATTC
958	CGTWDYS LSAWVF	1782	TGCGGAACATGGGATTACAGCCTGAGTGCCTGGG TGTTTC
959	CGTWDYS LSAEVF	1783	TGCGGAACATGGGATTACAGTCTGAGTGCTGAGG TGTTTC
960	CGTWDYS LRRAlF	1784	TGCGGAACATGGGATTACAGCCTGCGTCGTGCGA TATTC
961	CGTWDWS LILQLF	1785	TGCGGAACATGGGATTGGAGCCTCATTCTTCAAT TGTTTC

962	CGTWDVT LHTGVF	1786	TGCGGAACATGGGATGTCACCTTGCATACTGGGG TGTTTC
963	CGTWDVT LHIGVF	1787	TGCGGAACATGGGATGTCACCTTGCATATTGGGG TGTTTC
964	CGTWDVT LHAGVF	1788	TGCGGAACATGGGATGTCACCTTGCATGCTGGGG TGTTTC
965	CGTWDVS LYSGGVF	1789	TGCGGAACATGGGATGTCAGTTTGTATAGTGGCG GGGTCTTC
966	CGTWDVS LTSFVF	1790	TGTGGAACATGGGATGTCAGCCTGACTTCTTTCGT CTTC
967	CGTWDVS LSVGL	1791	TGCGGAACATGGGATGTCAGCCTGAGTGTTGGGG TGCTC
968	CGTWDVS LSAGDVVF	1792	TGCGGAACGTGGGATGTCAGCCTGAGTGCTGGCG ATGTAGTTTTC
969	CGTWDVS LNVVVF	1793	TGCGGAACATGGGATGTCAGCCTGAATGTCGTGG TTTTTC
970	CGTWDVS LNTQVF	1794	TGCGGAACATGGGATGTCAGCCTGAATACTCAGG TGTTTC
971	CGTWDVS LGALF	1795	TGCGGCACATGGGATGTGAGCCTGGGTGCGCTGT TC
972	CGTWDVN LKTUVF	1796	TGCGGAACGTGGGACGTTAATCTGAAAACGTGTCG TTTTTC
973	CGTWDVIL SAEVF	1797	TGCGGAACATGGGATGTCATCCTGAGTGCTGAGG TATTC
974	CGTWDTT VSAVVF	1798	TGCGGAACATGGGATACCACCGTGAGTGCTGTGG TTTTTC
975	CGTWDTT LTAWVF	1799	TGCGGAACATGGGATACCACCCTGACTGCCTGGG TGTTTC
976	CGTWDTT LSVFLF	1800	TGCGGAACATGGGACACCACCTTGAGTGTTTTCC TATTC
977	CGTWDTS VSAGVF	1801	TGCGGGACTTGGGATACCAGTGTGAGTGCTGGGG TGTTTC
978	CGTWDTS VISWVF	1802	TGCGGAACATGGGATACCAGTGTGATTTCTTGGG TTTTTC

979	CGTWDTS RSSLYVVF	1803	TGCGGAACATGGGATACCAGTCGGAGTTCTCTCT ATGTGGTCTTC
980	CGTWDTS RSAWVF	1804	TGCGGAACATGGGATACCAGCCGGAGTGCTTGGG TATTC
981	CGTWDTS RNPGGIF	1805	TGCGGAACATGGGATACCAGCCGGAATCCTGGAG GAATTTTC
982	CGTWDTS RGHVF	1806	TGCGGAACATGGGACACCAGTCGGGGTCATGTTT TC
983	CGTWDTSP STGQVLF	1807	TGCGGAACATGGGATACCAGCCCGAGTACTGGCC AGGTGCTTTTC
984	CGTWDTSP SAWVF	1808	TGCGGAACATGGGATACCAGCCCGAGTGCCTGGG TGTTTC
985	CGTWDTSL TWVF	1809	TGCGGAACATGGGATACTAGCCTGACCTGGGTGT TC
986	CGTWDTSL TWFAVF	1810	TGCGGAACATGGGATACCAGCCTGACGTGGTTTCG CAGTGTTTC
987	CGTWDTSL TVVVF	1811	TGCGGAACATGGGATACCAGCCTGACTGTTGTGG TATTC
988	CGTWDTSL TTSWVF	1812	TGCGGAACATGGGATACCAGCCTGACTACTTCTT GGGTGTTTC
989	CGTWDTSL TTGPFWVF	1813	TGCGGAACATGGGATACCAGCCTGACCACTGGTC CTTTTTGGGTGTTTC
990	CGTWDTSL TPFYVF	1814	TGCGGAACATGGGATACCAGCCTGACTCCTTTTT ATGTCTTC
991	CGTWDTSL TAYVF	1815	TGCGGAACATGGGATACCAGCCTGACTGCTTATG TCTTC
992	CGTWDTSL TAWVF	1816	TGCGGAACATGGGATACCAGCCTGACTGCTTGGG TGTTTC
993	CGTWDTSL TAWGVF	1817	TGCGGAACATGGGATACCAGCCTGACTGCGTGGG GGGTGTTTC
994	CGTWDTSL TAVVL	1818	TGCGGCACATGGGATACCAGCCTGACTGCGGTGG TTCTC
995	CGTWDTSL TARVF	1819	TGCGGAACCTGGGATACCAGCCTGACTGCTCGGG TTTTTC

996	CGTWDTSL TAIVF	1820	TGCGGAACATGGGATACCAGCCTGACTGCGATTG TCTTC
997	CGTWDTSL TAGVF	1821	TGCGGAACATGGGATACCAGCCTGACTGCTGGTG TCTTC
998	CGTWDTSL SVYVF	1822	TGCGGAACATGGGATACCAGCCTGAGTGTTTATG TCTTC
999	CGTWDTSL SVVF	1823	TGCGGAACATGGGATACCAGCCTGAGTGTGGTGT TC
1000	CGTWDTSL SVGEF	1824	TGCGGGACATGGGATACCAGCCTGAGTGTGGGG AATTC
1001	CGTWDTSL STWVF	1825	TGCGGAACATGGGATACCAGCCTGAGTACTTGGG TGTTTC
1002	CGTWDTSL STVVF	1826	TGCGGAACATGGGATACCAGCCTGAGTACTGTGG TATTC
1003	CGTWDTSL STGQVLF	1827	TGCGGAACATGGGATACCAGCCTGAGTACTGGCC AGGTGCTTTTC
1004	CGTWDTSL STGPLWVF	1828	TGCGGCACATGGGATACCAGCCTGAGCACTGGTC CTCTTTGGGTGTTC
1005	CGTWDTSL SSYVF	1829	TGCGGAACTTGGGATACCAGCCTGAGTTCTTATG TCTTC
1006	CGTWDTSL SSVVF	1830	TGCGGAACATGGGATACCAGCCTGAGTTCTGTGG TCTTC
1007	CGTWDTSL SSRYIF	1831	TGCGGAACATGGGATACCAGCCTGAGTTCTAGAT ACATATTC
1008	CGTWDTSL SSRFIF	1832	TGCGGAACATGGGATACCAGCCTGAGTTCTAGAT TCATATTC
1009	CGTWDTSL SSGWVF	1833	TGCGGAACATGGGATACCAGCCTGAGTTCTGGGT GGGTGTTC
1010	CGTWDTSL SRYVF	1834	TGCGGAACATGGGATACCAGCCTGAGTCGGTATG TGTTTC
1011	CGTWDTSL SQWLF	1835	TGCGGAACTTGGGATACCAGTCTGAGTCAATGGC TGTTTC
1012	CGTWDTSL SPGLWVF	1836	TGCGGAACATGGGATACCAGCCTGAGTCCTGGCC TTTGGGTGTTC

1013	CGTWDTSL SNYVF	1837	TGCGGAACATGGGATAACCAGCCTGAGTAATTATG TCTTC
1014	CGTWDTSL SIWVF	1838	TGCGGAACATGGGATAACCAGCCTAAGTATTTGGG TGTTTC
1015	CGTWDTSL SIGPFWVF	1839	TGCGGCACATGGGATAACCAGCCTGAGCATTGGTC CTTTTTGGGTGTTC
1016	CGTWDTSL SGWVF	1840	TGCGGAACATGGGATAACCAGCCTGAGTGGTTGGG TGTTTC
1017	CGTWDTSL SGTVF	1841	TGCGGAACATGGGATAACCAGCCTGAGTGGTACAG TGTTTC
1018	CGTWDTSL SGGQVF	1842	TGCGGAACATGGGATACTAGTCTGAGTGGTGGCC AGGTGTTC
1019	CGTWDTSL SGGIF	1843	TGCGGAACATGGGATAACCAGCCTGAGTGGTGGGA TATTC
1020	CGTWDTSL SGEDVVI	1844	TGCGGAACATGGGATAACCAGCCTGAGTGGTGAGG ATGTGGTAATC
1021	CGTWDTSL SFLYAF	1845	TGCGGAACATGGGATAACCAGCCTGAGTTTCCTTT ATGCTTTC
1022	CGTWDTSL SEVVF	1846	TGCGGAACATGGGATAACCAGCCTGAGTGAGGTCG TATTC
1023	CGTWDTSL SEVF	1847	TGCGGAACATGGGATAACCAGCCTGAGTGAAGTGT TC
1024	CGTWDTSL SENWVF	1848	TGCGGAACATGGGATACTAGCCTGAGTGAAAATT GGGTGTTC
1025	CGTWDTSL SAYIF	1849	TGCGGAACATGGGATAACCAGCCTGAGTGCCTACA TATTC
1026	CGTWDTSL SAVVL	1850	TGCGGAACATGGGATAACCAGCCTGAGTGCTGTGG TACTC
1027	CGTWDTSL SAVF	1851	TGCGGAACATGGGATAACCAGCCTGAGTGCTGTTT TC
1028	CGTWDTSL SARVF	1852	TGCGGAACATGGGATAACCAGCCTGAGTGCCCGGG TGTTTC
1029	CGTWDTSL SARQVF	1853	TGCGGCACATGGGATAACCAGCCTGAGTGCCCGCC AGGTATTC

1030	CGTWDTSL SALVF	1854	TGCGGAACATGGGATACCAGCCTGAGTGCTTTGG TTTTTC
1031	CGTWDTSL SAKVF	1855	TGCGGAACATGGGATACCAGCCTGAGTGCTAAGG TGTTTC
1032	CGTWDTSL SAKIF	1856	TGCGGAACATGGGATACCAGCCTGAGTGCGAAAA TCTTC
1033	CGTWDTSL SAKAVF	1857	TGCGGAACATGGGATACCAGCCTGAGTGCCAAGG CGGTATTC
1034	CGTWDTSL SAHAVF	1858	TGCGGAACATGGGATACCAGCCTGAGTGCCCATG CTGTGTTC
1035	CGTWDTSL SAGYVF	1859	TGCGGAACATGGGATACCAGCCTGAGTGCTGGCT ATGTCTTC
1036	CGTWDTSL SAGRWF	1860	TGCGGAACATGGGACACCAGTCTGAGTGCTGGCC GCTGGGTGTTC
1037	CGTWDTSL SAGIF	1861	TGCGGAACATGGGATACCAGCCTGAGTGCTGGGA TATTC
1038	CGTWDTSL SAGFRVF	1862	TGCGGAACATGGGATACCAGCCTGAGTGCTGGTG GGTTCCGGGTCTTC
1039	CGTWDTSL SAGAF	1863	TGCGGAACATGGGATACCAGCCTGAGTGCTGGGG CATTC
1040	CGTWDTSL SADWFF	1864	TGCGGAACATGGGATACCAGTCTGAGTGCTGATT GGTTTTTC
1041	CGTWDTSL SADEYVF	1865	TGCGGAACATGGGATACCAGCCTGAGTGCTGATG AATATGTCTTC
1042	CGTWDTSL SAAWVF	1866	TGCGGCACATGGGATACCAGCCTGAGTGCGGCTT GGGTGTTC
1043	CGTWDTSL SAALF	1867	TGCGGAACATGGGATACCAGCCTGAGTGCTGCGC TATTC
1044	CGTWDTSL SAAGVF	1868	TGCGGAACATGGGATACCAGCCTGAGTGCTGCGG GGGTTTTTC
1045	CGTWDTSL RVVVF	1869	TGCGGAACATGGGATACCAGCCTGAGAGTTGTGG TTTTTC
1046	CGTWDTSL RTWVF	1870	TGCGGAACATGGGATACCAGCCTGAGAACCTGGG TATTC

1047	CGTWDTSL RGAVF	1871	TGCGGAACGTGGGATACCAGCCTGAGGGGTGCAG TGTTTC
1048	CGTWDTSL RAVVF	1872	TGCGGAACATGGGATACCAGCCTGCGTGCTGTGG TATTC
1049	CGTWDTSL NVVYVF	1873	TGCGGAACATGGGATACAAGCCTGAATGTAGTTT ATGTCTTC
1050	CGTWDTSL NTYLF	1874	TGCGGAACATGGGATACCAGCCTCAACACCTACC TGTTTC
1051	CGTWDTSL NFAWLF	1875	TGCGGAACATGGGATACTAGCCTGAACTTCGCTT GGCTGTTC
1052	CGTWDTSL LVWLF	1876	TGCGGCACATGGGATACCAGCCTTCTTGTGTGGC TTTTTC
1053	CGTWDTSL KTWVF	1877	TGCGGAACATGGGATACCAGTCTGAAGACGTGGG TGTTTC
1054	CGTWDTSL IVWVF	1878	TGCGGAACATGGGATACCAGTCTGATTGTCTGGG TGTTTC
1055	CGTWDTSL ITGVF	1879	TGCGGAACATGGGATACCAGCCTAATTAAGTGGGG TGTTTC
1056	CGTWDTSL ISVVF	1880	TGCGGAACATGGGATACCAGCCTGATTAGCGTGG TATTC
1057	CGTWDTSL IAYVF	1881	TGCGGAACATGGGATACCAGCCTGATTGCTTATG TCTTC
1058	CGTWDTSL HTELF	1882	TGCGGAACATGGGATACCAGCCTGCACACTGAGT TGTTTC
1059	CGTWDTSL GSYVF	1883	TGCGGAACTTGGGATACCAGCCTGGGTTCTTATG TCTTC
1060	CGTWDTSL GSLWVF	1884	TGCGGAACATGGGATACCAGCCTGGGTTCTCTTT GGGTGTTC
1061	CGTWDTSL GSGVF	1885	TGCGGTACATGGGATACCAGCCTGGGTTCTGGGG TATTC
1062	CGTWDTSL GGRGVF	1886	TGCGGAACTTGGGATACCAGTCTGGGTGGTAGAG GGGTCTTC
1063	CGTWDTSL GAWVF	1887	TGCGGAACATGGGATACCAGCCTGGGTGCTTGGG TGTTTC

1064	CGTWDTSL GAVVF	1888	TGCGGAACATGGGATAACCAGCCTGGGTGCCGTGG TATTC
1065	CGTWDTSL GAGVF	1889	TGCGGAACATGGGATAACCAGCCTGGGTGCTGGGG TATTC
1066	CGTWDTSL GAGLF	1890	TGCGGAACATGGGATAACCAGCCTGGGTGCTGGCC TATTC
1067	CGTWDTSL DAVVF	1891	TGCGGAACATGGGATAACCAGTCTGGATGCTGTGG TTTTC
1068	CGTWDTSL DAVLF	1892	TGCGGGACTTGGGATAACCAGCCTGGATGCTGTGC TG TTC
1069	CGTWDTSL AWVF	1893	TGCGGAACATGGGATAACCAGCCTGGCTTGGGTGT TC
1070	CGTWDTSL ATGLF	1894	TGCGGAACATGGGATAACCAGCCTGGCGACTGGAC TG TTC
1071	CGTWDTSL APVVF	1895	TGCGGGACATGGGATAACCAGCCTGGCCCCTGTAG TCTTC
1072	CGTWDTR LTIVIF	1896	TGCGGAACATGGGACACCCGCCTGACTATTGTGA TCTTC
1073	CGTWDTR LSVWLF	1897	TGTGGAACATGGGACACCAGGCTGAGTGTTTGGC TG TTC
1074	CGTWDTR LSVGVF	1898	TGCGGAACGTGGGACACCAGACTGAGTGTTGGGG TTTTC
1075	CGTWDTR LSTVIF	1899	TGCGGCACATGGGATAACCAGACTGAGTACTGTAA TTTTC
1076	CGTWDTR LSSVVF	1900	TGCGGAACATGGGATAACCCGCCTGAGTTCTGTGG TCTTC
1077	CGTWDTR LSIVVF	1901	TGCGGAACATGGGATAACCCGCCTGAGTATTGTGG TTTTC
1078	CGTWDTR LSAYVVF	1902	TGCGGAACATGGGATAACCAGACTGAGTGCCTATG TGGTATTC
1079	CGTWDTR LSAWVF	1903	TGCGGAACCTGGGACACCCGCCTGAGTGCGTGGG TG TTC
1080	CGTWDTR LSAVVF	1904	TGCGGAACATGGGATAACCAGACTGAGTGCTGTGG TG TTC

1081	CGTWDTR LSAGLF	1905	TGCGGAACATGGGATACCCGCCTGAGTGCTGGGT TG TTC
1082	CGTWDTR LSAGGVF	1906	TGCGGAACATGGGATACCAGACTGAGTGCTGGTG GGGTG TTC
1083	CGTWDTR LNVWLF	1907	TGCGGAACATGGGATACCAGATTGAATGTGTGGC TATTC
1084	CGTWDTN REVVLL	1908	TGCGGAACATGGGATACCAACCGGGAAGTTGTGC TCCTC
1085	CGTWDTN LRAHVF	1909	TGCGGAACATGGGATACCAACCTGCGTGCCCATG TCTTC
1086	CGTWDTN LPAVVF	1910	TGCGGAACATGGGATACTAATCTGCCCGCTGTAG TG TTC
1087	CGTWDTN LGGVF	1911	TGCGGAACATGGGACACCAATTTGGGTGGGGTGT TC
1088	CGTWDTIV SIGVF	1912	TGCGGAACATGGGATACCATCGTGAGTATTGGGG TG TTC
1089	CGTWDTIL SAVVF	1913	TGCGGAACATGGGATACCATCCTGAGTGCGGTGG TG TTC
1090	CGTWDTIL SAEVF	1914	TGCGGCACATGGGATACCATCCTGAGTGCTGAGG TG TTC
1091	CGTWDTH LGVVF	1915	TGCGGAACATGGGATACCCACCTGGGTGTGGTTT TC
1092	CGTWDTG PSPHWLF	1916	TGCGGAACATGGGATACCGGCCCGAGCCCTCATT GGCTG TTC
1093	CGTWDTG LTFGGVF	1917	TGCGGAACATGGGATACCGGCCTGACTTTTGGAG GCGTG TTC
1094	CGTWDTG LTAFFVF	1918	TGCGGAACATGGGATACCGGCCTGACTGCTTTTG TCTTC
1095	CGTWDTG LSVWVF	1919	TGCGGAACATGGGATACCGGCCTGAGTGTTTGGG TG TTC
1096	CGTWDTG LSTGIF	1920	TGCGGAACATGGGATACCGGCCTGAGTACTGGGA TTTTC
1097	CGTWDTG LSSLLF	1921	TGCGGAACATGGGATACCGGCCTGAGTTCCTGC TCTTC

1098	CGTWDTG LSIVVF	1922	TGCGGAACGTGGGACACCGGCCTGAGTATTGTGG TG TTC
1099	CGTWDTG LSFVVF	1923	TGCGGAACGTGGGACACCGGCCTGAGTTTTGTGG TG TTC
1100	CGTWDTG LSAWVF	1924	TGCGGAACATGGGATACCGGCCTGAGTGCTTGGG TG TTC
1101	CGTWDTG LSAGVVF	1925	TGCGGAACATGGGATACCGGCCTGAGTGCTGGTG TGGTATTC
1102	CGTWDTG LRGWIF	1926	TGCGGAACATGGGATACCGGTCTGAGGGGGTTGGA TTTTTC
1103	CGTWDTE LSAGVF	1927	TGCGGAACATGGGATACCGAGCTAAGTGCGGGG GTCTTC
1104	CGTWDTA LTAGVF	1928	TGCGGAACGTGGGATACCGCCCTGACTGCTGGGG TG TTC
1105	CGTWDTA LSLVVF	1929	TGCGGAACATGGGATACTGCCCTGAGTCTTGTGG TCTTC
1106	CGTWDTA LSAWLF	1930	TGCGGAACATGGGATACCGCCCTGAGTGCTGGC TG TTC
1107	CGTWDTA LSAGVF	1931	TGCGGCACATGGGATACCGCCCTGAGTGCTGGGG TG TTC
1108	CGTWDTA LRGVLF	1932	TGCGGAACATGGGATACCGCCCTGCGTGGCGTGC TG TTC
1109	CGTWDTA LKEWLF	1933	TGCGGAACATGGGATACCGCCCTGAAAGAATGGC TG TTC
1110	CGTWDRT LTAGDVLF	1934	TGCGGAACATGGGATAGGACCCTGACTGCTGGCG ATGTGCTCTTC
1111	CGTWDRS VTYVF	1935	TGCGGAACATGGGATAGAAGCGTGACTTATGTCT TC
1112	CGTWDRS RNEWVF	1936	TGCGGAACATGGGATCGCAGCCGAAATGAATGG GTGTTC
1113	CGTWDRS LTVWVF	1937	TGCGGAACATGGGATCGCAGTCTGACTGTTTGGG TCTTC
1114	CGTWDRS LTPGWLF	1938	TGCGGAACATGGGATCGCAGCCTGACTCCTGGGT GGTTGTTC

1115	CGTWDRS LTAWVF	1939	TGCGGAACATGGGATAGAAGCCTGACTGCTTGGG TG TTC
1116	CGTWDRS LSVVVF	1940	TGCGGAACATGGGACCGCAGCCTGAGTGTTGTGG TATTC
1117	CGTWDRS LSVVF	1941	TGCGGCACATGGGATCGCAGCCTGAGTG TAGTCT TC
1118	CGTWDRS LSVQLF	1942	TGCGGAACATGGGATAGGAGCCTGAGTGTTCAAT TG TTC
1119	CGTWDRS LSVLWVF	1943	TGCGGAACATGGGATCGCAGCCTCAGTGTTCTTT GGGTG TTC
1120	CGTWDRS LSVGLF	1944	TGCGGAACATGGGATCGCAGCCTGAGTGTTGGAT TATTC
1121	CGTWDRS LSTWVF	1945	TGCGGAACATGGGATCGCAGCCTGAGTACTTGGG TG TTC
1122	CGTWDRS LSTHWVL	1946	TGCGGAACATGGGATAGAAGCCTGAGTACTCATT GGGTG CTC
1123	CGTWDRS LSTHWVF	1947	TGCGGAACATGGGATAGAAGCCTGAGTACTCATT GGGTG TTC
1124	CGTWDRS LSSAVF	1948	TGCGGAACCTGGGATCGAAGCCTGAGTTCTGCGG TG TTC
1125	CGTWDRS LSPSYVF	1949	TGCGGAACATGGGACAGAAGCCTGAGTCCCTCTT ATGTCTTC
1126	CGTWDRS LSGEVF	1950	TGCGGAACATGGGATAGGAGCCTGAGTG GTGAG GTG TTC
1127	CGTWDRS LSGAVF	1951	TGCGGAACATGGGATAGGAGCCTGAGTG GTGCGG TG TTC
1128	CGTWDRS LSAVAF	1952	TGCGGAACATGGGATCGCAGCCTGAGTGCTGTGG CATT C
1129	CGTWDRS LSAGGEF	1953	TGCGGAACATGGGATAGGAGCCTGAGTGCCGGG GGGGAATTC
1130	CGTWDRS LSAFWVF	1954	TGCGGAACATGGGATCGCAGCCTGAGTGCTTTTT GGGTG TTC
1131	CGTWDRS LSAAVF	1955	TGCGGAACATGGGATAGGAGCCTGAGTGCTGCGG TG TTC

1132	CGTWDRS LSAALF	1956	TGCGGAACATGGGATAGGAGCCTGAGTGCTGCAC TCTTC
1133	CGTWDRS LRVF	1957	TGCGGAACATGGGATCGCAGCCTGAGAGTGTTTC
1134	CGTWDRS LNWVF	1958	TGCGGTACATGGGACAGAAGCCTTAATTGGGTGT TC
1135	CGTWDRS LNVYVF	1959	TGCGGAACATGGGATCGCAGCCTGAATGTTTATG TCTTC
1136	CGTWDRS LNVGVF	1960	TGCGGAACATGGGATAGGAGCCTGAATGTTGGGG TGTTTC
1137	CGTWDRS LHVVF	1961	TGCGGAACATGGGATCGGAGCCTGCATGTGGTCT TC
1138	CGTWDRS LGGWVF	1962	TGTGGAACATGGGATCGCAGCCTGGGTGGTTGGG TGTTTC
1139	CGTWDRS LGAFWVF	1963	TGCGGAACATGGGATCGCAGCCTGGGTGCTTTTT GGGTGTTTC
1140	CGTWDRS LFWVF	1964	TGCGGAACATGGGATAGAAGCCTGTTTTGGGTGT TC
1141	CGTWDRS LAAGVF	1965	TGCGGAACGTGGGATCGCAGCCTGGCTGCTGGGG TGTTTC
1142	CGTWDRR LSGVVF	1966	TGCGGAACATGGGATAGGAGGTTGAGTGGTGTCG TATTC
1143	CGTWDRR LSDVVF	1967	TGCGGAACGTGGGATCGCCGCCTAAGTGATGTGG TATTC
1144	CGTWDRR LSAVVF	1968	TGCGGAACATGGGATAGGAGGCTGAGTGCTGTGG TATTC
1145	CGTWDRR LNVAFF	1969	TGCGGAACATGGGATAGACGCCTGAATGTTGCGT TCTTC
1146	CGTWDRR LLAVF	1970	TGTGGAACATGGGATAGGAGGCTGCTTGCTGTTT TC
1147	CGTWDRN LRAVVF	1971	TGCGGAACTGGGATAGGAACCTGCGCGCCGTGG TCTTC
1148	CGTWDRL SAGVF	1972	TGCGGAACATGGGATAGGCTGAGTGCTGGGGTGT TC

1149	CGTWDRG PNTGVF	1973	TGCGGAACATGGGATAGAGGCCCGAATACTGGG GTATTC
1150	CGTWDRG LNTVYVF	1974	TGCGGAACATGGGATAGAGGCCTGAATACTGTTT ACGTCTTC
1151	CGTWDNY VSAPWVF	1975	TGCGGAACATGGGATAACTATGTGAGTGCCCCTT GGGTGTTC
1152	CGTWDNY LSAGDVVF	1976	TGCGGAACATGGGATAACTACCTGAGTGCTGGCG ATGTGGTTTTTC
1153	CGTWDNY LRAGVF	1977	TGCGGAACATGGGATAACTACCTGAGAGCTGGGG TCTTC
1154	CGTWDNY LGAVVF	1978	TGCGGAACATGGGACAATTATCTGGGTGCCGTGG TTTTTC
1155	CGTWDNY LGAGVF	1979	TGCGGAACATGGGATAACTACCTGGGTGCGGGGG TGTTTC
1156	CGTWDNT VSAPWVF	1980	TGCGGAACATGGGATAACACCGTGAGTGCCCCTT GGGTTTTTTC
1157	CGTWDNT LSLWVF	1981	TGCGGAACATGGGATAACACCCTGAGTCTTTGGG TGTTTC
1158	CGTWDNT LSAGVF	1982	TGCGGAACATGGGATAACACCCTGAGTGCTGGGG TCTTC
1159	CGTWDNT LLTVLF	1983	TGCGGAACATGGGACAACACTCTGCTTACTGTGT TATTC
1160	CGTWDNR LSSVIF	1984	TGCGGAACATGGGATAACAGACTGAGTAGTGTGA TTTTTC
1161	CGTWDNR LSAVVF	1985	TGCGGAACATGGGATAACAGGTTGAGTGCTGTGG TCTTC
1162	CGTWDNR LSAGGIF	1986	TGCGGAACATGGGATAACAGGCTGAGTGCTGGTG GGATATTC
1163	CGTWDNR LSAEVF	1987	TGCGGAACATGGGATAACAGACTGAGTGCTGAGG TGTTTC
1164	CGTWDNR LRVGVL	1988	TGTGGAACATGGGATAACAGACTGCGTGTTGGGG TTCTC
1165	CGTWDNR LLENVF	1989	TGCGGAACATGGGATAATCGCCTGCTTGAGAATG TCTTC

1166	CGTWDNN LRAVF	1990	TGCGGAACATGGGATAACAACCTGCGTGCTGTCT TC
1167	CGTWDNN LRAGVF	1991	TGCGGAACTTGGGATAATAACCTGCGTGCTGGAG TGTTTC
1168	CGTWDNN LGGGRVF	1992	TGCGGAACATGGGACAACAATTTGGGCGGTGGCC GGGTGTTTC
1169	CGTWDNN LGAGVL	1993	TGCGGAACATGGGATAACAACCTGGGTGCTGGCG TCCTC
1170	CGTWDNN LGAGVF	1994	TGCGGAACATGGGATAACAACCTGGGTGCTGGCG TCTTC
1171	CGTWDNIL SAAVF	1995	TGCGGAACTTGGGATAACATCCTGAGCGCTGCGG TGTTTC
1172	CGTWDNIL DAGVF	1996	TGCGGAACTTGGGATAACATCTTGATGCAGGGG TTTTTC
1173	CGTWDND LSGWLF	1997	TGCGGAACATGGGATAACGACCTGAGTGGTTGGC TGTTTC
1174	CGTWDND LSAWVF	1998	TGCGGAACATGGGATAACGACCTGAGTGCCTGGG TGTTTC
1175	CGTWDLT LGGVVF	1999	TGCGGAACATGGGATCTCACCTGGGTGGTGTGG TGTTTC
1176	CGTWDLSL SAGVF	2000	TGCGGAACATGGGATCTCAGCCTGAGTGCTGGGG TATTC
1177	CGTWDLSL KEWVF	2001	TGCGGAACATGGGATCTCAGCCTGAAAGAATGGG TGTTTC
1178	CGTWDLSL DAVVF	2002	TGCGGAACGTGGGATCTCAGCCTGGATGCTGTTG TTTTTC
1179	CGTWDLK VF	2003	TGCGGAACTTGGGACCTGAAGGTTTTTC
1180	CGTWDKT LSVWVF	2004	TGCGGAACATGGGATAAGACTCTGAGTGTTTGGG TGTTTC
1181	CGTWDKS LSVWVF	2005	TGCGGAACATGGGATAAGAGCCTGAGTGTTTGGG TGTTTC
1182	CGTWDKS LSGVVF	2006	TGCGGAACATGGGATAAGAGCCTGAGTGGTGTGG TATTT

1183	CGTWDKS LSDWVF	2007	TGCGGAACATGGGATAAGAGCCTGAGTGATTGGG TG TTC
1184	CGTWDKS LSALVF	2008	TGCGGAACATGGGATAAGAGCCTGAGTGCTTTGG TTTT C
1185	CGTWDKS LSAGVF	2009	TGCGGAACATGGGATAAGAGCCTGAGTGCTGGCG TCTTC
1186	CGTWDKS LSADV F	2010	TGCGGAACATGGGATAAGAGCCTGAGTGCCGACG TCTTC
1187	CGTWDKR LTIVVF	2011	TGCGGAACATGGGATAAACGCCTGACTATTGTGG TCTTC
1188	CGTWDKR LSAWVL	2012	TGCGGAACATGGGATAAACGCCTGAGTGCCCTGGG TGCTC
1189	CGTWDKN LRAVVF	2013	TGCGGAACATGGGATAAGAACCTGCGTGCTGTGG TCTTC
1190	CGTWDITL SGFVF	2014	TGCGGAACATGGGATATCACCTGAGTGGGTTTG TCTTC
1191	CGTWDITL HTGVF	2015	TGCGGAACATGGGATATCACCTGCATACTGGAG TATTC
1192	CGTWDISV TVVF	2016	TGCGGAACATGGGATATCAGTGTGACTGTGGTGT TC
1193	CGTWDISV RGYAF	2017	TGCGGAACATGGGATATCAGTGTGAGGGGTTATG CCTTC
1194	CGTWDISR WVF	2018	TGCGGAACATGGGATATCAGCCGTTGGGTTTT C
1195	CGTWDISP SAWVF	2019	TGCGGAACATGGGATATCAGCCCGAGTGCTTGGG TG TTC
1196	CGTWDISL SVWVF	2020	TGCGGAACATGGGATATTAGCCTGAGTGTCTGGG TG TTC
1197	CGTWDISL SVVF	2021	TGCGGAACATGGGATATCAGCCTGAGTGTTGTAT TC
1198	CGTWDISL SSVVF	2022	TGCGGAACTTGGGATATCAGCCTGAGTTCTGTGG TG TTC
1199	CGTWDISL SHWLF	2023	TGCGGAACATGGGATATCAGCCTGAGTCACTGGT TG TTC

1200	CGTWDISL SGWVF	2024	TGCGGAACATGGGATATCAGTCTGAGTGGTTGGG TG TTC
1201	CGTWDISL SGRVF	2025	TGCGGAACATGGGATATCAGCCTGAGTGGTCGAG TG TTC
1202	CGTWDISL SAWAF	2026	TGCGGAACATGGGACATCAGCCTGAGTGCTTGGG CG TTC
1203	CGTWDISL SAVVF	2027	TGCGGAACATGGGATATCAGCCTGAGTGCTGTGG TTTT C
1204	CGTWDISL SAVIF	2028	TGCGGGACATGGGACATCAGCCTGAGTGCTGTGA TATTC
1205	CGTWDISL SAV F	2029	TGCGGAACATGGGATATCAGCCTGAGTGCTGTGT TC
1206	CGTWDISL SARVF	2030	TGCGGAACATGGGATATCAGCCTGAGTGCCCGGG TG TTC
1207	CGTWDISL SALVF	2031	TGCGGAACATGGGATATCAGCCTGAGTGCCCTGG TG TTC
1208	CGTWDISL SAHVF	2032	TGCGGAACATGGGATATTAGCCTGAGTGCCCATG TCTTC
1209	CGTWDISL SAGVVF	2033	TGCGGAACATGGGATATCAGCCTGAGTGCTGGGG TGGTATTC
1210	CGTWDISL SAGPYVF	2034	TGCGGAACATGGGATATCAGCCTGAGTGCCGGCC CTTATGTCTTC
1211	CGTWDISL SAGGVF	2035	TGCGGCACATGGGATATCAGCCTGAGTGCTGGAG GGGTG TTC
1212	CGTWDISL SAEVF	2036	TGCGGAACATGGGATATCAGCCTGAGTGCTGAGG TTTT C
1213	CGTWDISL SAAVF	2037	TGCGGAACATGGGATATCAGCCTGAGTGCTGCTG TG TTC
1214	CGTWDISL RAVF	2038	TGCGGAACATGGGATATCAGCCTGCGTGCTGTGT TC
1215	CGTWDISL NTGVF	2039	TGCGGAACATGGGATATTAGCCTGAATACTGGGG TG TTC
1216	CGTWDISL NNYVF	2040	TGCGGAACATGGGATATCAGCCTAAATAATTATG TCTTC

1217	CGTWDISL IAGVF	2041	TGCGGAACATGGGATATCAGCCTAATTGCTGGGG TATTC
1218	CGTWDISL HTWLF	2042	TGCGGAACATGGGATATCAGCCTGCATACTTGGC TGTTTC
1219	CGTWDIRL TDELLF	2043	TGCGGAACATGGGATATCCGCCTGACCGATGAGC TGTTATTC
1220	CGTWDIRL SGFVF	2044	TGCGGAACATGGGATATCAGACTGAGCGGTTTTG TTTTTC
1221	CGTWDINL GAGGLYV F	2045	TGCGGAACATGGGATATCAACCTGGGTGCTGGGG GCCTTTATGTCTTC
1222	CGTWDIIL SAEVF	2046	TGCGGAACATGGGATATCATCCTGAGTGCTGAGG TATTC
1223	CGTWDHT LSAVF	2047	TGCGGAACATGGGATCACACCCTGAGTGCTGTCT TC
1224	CGTWDHT LLTVLF	2048	TGCGGAACATGGGACCACACTCTGCTTACTGTGT TATTC
1225	CGTWDHS LTAVVF	2049	TGCGGAACATGGGATCACAGCCTGACTGCTGTGG TATTC
1226	CGTWDHS LTAGIF	2050	TGCGGAACCTGGGATCACAGCCTGACTGCTGGGA TATTC
1227	CGTWDHS LSVVLF	2051	TGCGGAACATGGGATCACAGCCTGAGTGTTGTAT TATTC
1228	CGTWDHS LSLVF	2052	TGCGGAACATGGGATCACAGCCTGAGTTTGGTAT TC
1229	CGTWDHS LSIGVF	2053	TGCGGAACATGGGATCACAGCCTGTCTATTGGGG TTTTTC
1230	CGTWDHS LSAGVF	2054	TGCGGAACATGGGATCACAGCCTGAGTGCTGGGG TGTTTC
1231	CGTWDHS LSAFVF	2055	TGTGGAACCTGGGATCACAGCCTGAGTGCTTTTCG TGTTTC
1232	CGTWDHS LSAAVF	2056	TGCGGAACATGGGATCACAGTCTGAGTGCTGCTG TTTTTC

1233	CGTWDHN LRAVF	2057	TGCGGAACATGGGACCACAATCTGCGTGCTGTCT TC
1234	CGTWDFTL SVGRF	2058	TGCGGGACATGGGATTTACCCCTGAGTGTTGGGC GCTTC
1235	CGTWDFTL SAPVF	2059	TGCGGAACATGGGATTTACCCCTGAGTGCTCCTG TCTTC
1236	CGTWDFS VSAGWVF	2060	TGCGGAACGTGGGATTTACAGCGTGAGTGCTGGGT GGGTGTTTC
1237	CGTWDFSL TTWLF	2061	TGCGGAACGTGGGATTTACAGTCTTACTACCTGGTT ATTC
1238	CGTWDFSL SVWVF	2062	TGCGGAACATGGGATTTACAGCCTGAGTGTTTGGG TGTTTC
1239	CGTWDFSL STGVF	2063	TGCGGAACATGGGATTTACAGCCTGAGTACTGGGG TTTTTC
1240	CGTWDFSL SGVVF	2064	TGCGGCACATGGGATTTACAGCCTGAGTGGTGTGG TATTC
1241	CGTWDFSL SGFVF	2065	TGCGGAACATGGGATTTACAGCCTGAGTGGTTTTCG TGTTTC
1242	CGTWDFSL SAGVF	2066	TGCGGAACATGGGATTTACAGCCTGAGTGCTGGGG TGTTTC
1243	CGTWDET VRGWVF	2067	TGCGGAACATGGGATGAAACCGTGAGAGGTTGG GTGTTTC
1244	CGTWDES LSWVF	2068	TGCGGAACATGGGATGAAAGTCTGAGAAGCTGG GTGTTTC
1245	CGTWDER QTDESYVF	2069	TGCGGAACTTGGGATGAGAGGCAGACTGATGAGT CCTATGTCTTC
1246	CGTWDER LVAGQVF	2070	TGCGGAACATGGGATGAGAGACTCGTTGCTGGCC AGGTCTTC
1247	CGTWDER LSPGAFF	2071	TGCGGAACATGGGATGAGAGACTGAGTCCTGGAG CTTTTTTC
1248	CGTWDEK VF	2072	TGCGGAACATGGGATGAGAAGGTGTTTC
1249	CGTWDEG QTTDFVF	2073	TGCGGAACCTGGGATGAAGGCCAGACTACTGATT TCTTTGTCTTC

1250	CGTWDDT LAGVVF	2074	TGCGGAACATGGGATGACACCCTGGCTGGTGTGG TCTTC
1251	CGTWDDR LTSAVF	2075	TGCGGAACATGGGATGACAGGCTGACTTCTGCGG TCTTC
1252	CGTWDDR LFVVVF	2076	TGCGGAACATGGGATGACAGACTGTTTGTGTGG TATTC
1253	CGTWDDN LRGWVF	2077	TGCGGAACATGGGATGATAACCTGAGAGGTTGGG TGTTTC
1254	CGTWDDN LRGVVF	2078	TGCGGAACATGGGATGACAACTGCGTGGTGTTCG TGTTTC
1255	CGTWDDN LNIGRVF	2079	TGCGGAACCTGGGATGACAATTTGAATATTGGAA GGGTGTTC
1256	CGTWDDIL SAVIF	2080	TGCGGAACATGGGATGACATCCTGAGTGCTGTGA TATTC
1257	CGTWDDIL RGWVF	2081	TGCGGAACATGGGATGATATCCTGAGAGGTTGGG TGTTTC
1258	CGTWDAT LSPGWLF	2082	TGCGGAACATGGGATGCCACCCTGAGTCCTGGGT GGTTATTC
1259	CGTWDAS VTSWVF	2083	TGCGGAACATGGGATGCCAGCGTGACTTCTTGGG TGTTTC
1260	CGTWDAS LTSVVF	2084	TGCGGAACATGGGATGCCAGCCTGACTTCTGTGG TCTTC
1261	CGTWDAS LSWVVF	2085	TGCGGAACATGGGATGCCAGCCTGAGTGTTTGGG TGTTTC
1262	CGTWDAS LSVPWVF	2086	TGCGGAACATGGGATGCCAGCCTGAGTGTTTCCTT GGGTGTTC
1263	CGTWDAS LSVAVF	2087	TGCGGAACATGGGATGCCAGCCTGAGTGTGGCGG TATTC
1264	CGTWDAS LSTWVF	2088	TGCGGAACATGGGATGCCAGCCTGAGTACCTGGG TATTC
1265	CGTWDAS LSGVVF	2089	TGCGGAACATGGGATGCCAGCCTGAGTGGTGTGG TATTC
1266	CGTWDAS LSGGGEF	2090	TGCGGAACATGGGATGCCAGCCTGAGTGGTGGGG GAGAATTC

1267	CGTWDAS LSAGVF	2091	TGCGGAACATGGGATGCCAGCCTGAGTGCTGGGG TG TTC
1268	CGTWDAS LSAGLF	2092	TGCGGAACATGGGATGCCAGCCTGAGTGCTGGGC TTTTC
1269	CGTWDAS LSAEVF	2093	TGTGGCACATGGGATGCCAGCCTGAGTGCTGAAG TCTTC
1270	CGTWDAS LSADFWVF	2094	TGCGGAACATGGGATGCCAGCCTGAGTGCTGACT TTTGGGTG TTC
1271	CGTWDAS LRVFF	2095	TGCGGAACATGGGATGCCAGCCTGAGAGTCTTCT TC
1272	CGTWDAS LRAVVL	2096	TGCGGAACATGGGATGCCAGTCTGAGGGCTGTGG TACTC
1273	CGTWDAS LNIWVF	2097	TGCGGAACATGGGATGCCAGCCTGAATATTTGGG TTTTC
1274	CGTWDAS LKNLVF	2098	TGCGGGACATGGGATGCCAGCCTGAAGAATCTGG TCTTC
1275	CGTWDAS LGAWVF	2099	TGCGGAACATGGGATGCCAGCCTGGGTGCCTGGG TATTC
1276	CGTWDAS LGAVVF	2100	TGCGGAACATGGGATGCCAGCCTGGGTGCTGTGG TCTTC
1277	CGTWDAS LGAGVF	2101	TGCGGAACATGGGATGCCAGCCTGGGTGCGGGGG TCTTC
1278	CGTWDAR LSGLYVF	2102	TGCGGAACATGGGATGCTAGGCTGAGTGGCCTTT ATGTCTTC
1279	CGTWDAR LGGAVF	2103	TGTGGAACCTGGGATGCGAGACTGGGTGGTGCAG TCTTC
1280	CGTWDAN LRAGVF	2104	TGCGGAACATGGGATGCCAATCTGCGTGCTGGGG TCTTC
1281	CGTWDAIL SGWVF	2105	TGCGGAACATGGGATGCTATCATAAGTGGTTGGG TG TTC
1282	CGTWDAG QSVWVF	2106	TGCGGAACATGGGATGCCGGCCAGAGTGTTTGGG TG TTC
1283	CGTWDAG LTGLYVF	2107	TGCGGCACATGGGATGCCGGGCTGACTGGCCTTT ATGTCTTC

1284	CGTWDAG LSVYVF	2108	TGCGGAACTGGGATGCCGGTCTGAGTGTTTATG TCTTC
1285	CGTWDAG LSTGVF	2109	TGCGGGACATGGGATGCCGGCCTGAGTACTGGGG TCTTC
1286	CGTWDAG LSGDVF	2110	TGCGGAAACATGGGATGCCGGCCTGAGTGGGGACG TTTTTC
1287	CGTWDAG LSAGYVF	2111	TGCGGAAACATGGGATGCCGGCCTGAGTGCTGGTT ATGTCTTC
1288	CGTWDAG LRVWVF	2112	TGCGGAAACATGGGATGCCGGCCTGCGTGTTTGGG TGTTTC
1289	CGTWDAG LREIF	2113	TGCGGAAACATGGGATGCCGGCCTGAGGGAAATTT TC
1290	CGTWASSL SSWVF	2114	TGCGGAAACATGGGGCCAGCAGCCTGAGTTCTTGGG TGTTTC
1291	CGTWAGS LSGHVF	2115	TGCGGAAACATGGGGCTGGCAGCCTGAGTGGTCATG TCTTC
1292	CGTWAGS LSAAWVF	2116	TGCGGAAACATGGGGCTGGCAGCCTGAGTGCCGCTT GGGTGTTTC
1293	CGTWAGS LNVYWVF	2117	TGCGGAAACATGGGGCTGGCAGCCTGAATGTTTATT GGGTGTTTC
1294	CGTWAGN LRPNWVF	2118	TGCGGAAACATGGGGCTGGCAACCTGAGACCTAATT GGGTGTTTC
1295	CGTRGSLG GAVF	2119	TGCGGAAACAAGGGGTAGCCTGGGTGGTGCGGTGT TC
1296	CGTRDTTL SVPVF	2120	TGCGGAAACAAGGGATAACCACCCTGAGTGTCCCGG TGTTTC
1297	CGTRDTSL NIEIF	2121	TGCGGAAACACGGGATAACCAGCCTCAATATTGAAA TCTTC
1298	CGTRDTSL NDVF	2122	TGTGGAACACGGGATAACCAGCCTGAATGATGTCT TC
1299	CGTRDTRL SIVVF	2123	TGCGGAAACACGGGATAACCCGCCTGAGTATTGTGG TTTTTC
1300	CGTRDTIL SAEVF	2124	TGCGGCACACGGGATAACCATCCTGAGTGCTGAGG TGTTTC

1301	CGTRDRSL SGWVF	2125	TGCGGAACACGGGATAGAAGCCTGAGTGGTTGGG TG TTC
1302	CGSWYYN VFLF	2126	TGCGGATCATGGTATTACAATGTCTTCCTTTTC
1303	CGSWHSSL NLVVF	2127	TGCGGATCTTGGCATAGCAGCCTCAACCTTGTCGT CTTC
1304	CGSWGSG LSAPYVF	2128	TGCGGATCATGGGGTAGTGGCCTGAGTGCCCCTT ATGTCTTC
1305	CGSWESGL GAWLF	2129	TGCGGTTTCGTGGGAAAGCGGCCTGGGTGCTTGGC TG TTC
1306	CGSWDYG LLL F	2130	TGCGGATCCTGGGATTACGGCCTCCTACTCTTC
1307	CGSWDVS LTAVF	2131	TGCGGTTTCATGGGATGTCAGCCTGACTGCTGTTTT C
1308	CGSWDVS LNVGIF	2132	TGCGGATCCTGGGATGTCAGTCTCAATGTTGGCA TTTTTC
1309	CGSWDTTL RAWVF	2133	TGCGGATCATGGGATAACCACCCTGCGTGCTTGGG TG TTC
1310	CGSWDTSP VRAWVF	2134	TGCGGCTCGTGGGATAACCAGCCCTGTCCGTGCTT GGGTG TTC
1311	CGSWDTSL SVWVF	2135	TGCGGATCATGGGATAACCAGCCTGAGTGTTTGGG TG TTC
1312	CGSWDTSL SAEVF	2136	TGCGGATCATGGGATAACCAGCCTGAGTGCTGAGG TG TTC
1313	CGSWDTSL RAWVF	2137	TGCGGCTCGTGGGATAACCAGCCTGCGTGCTTGGG TG TTC
1314	CGSWDTSL RAWAF	2138	TGCGGCTCGTGGGATAACCAGCCTGCGTGCTTGGG CG TTC
1315	CGSWDTSL DARLF	2139	TGCGGATCATGGGATAACCAGCCTGGATGCTAGGC TG TTC
1316	CGSWDTIL LVYVF	2140	TGCGGATCATGGGATAACCATCCTGCTTGTCTATGT CTTC
1317	CGSWDRW QAAVF	2141	TGCGGATCATGGGATCGCTGGCAGGCTGCTGTCT TC

1318	CGSWDRSL SGYVF	2142	TGCGGATCATGGGATAGGAGCCTGAGTGGGTATG TCTTC
1319	CGSWDRSL SAYVF	2143	TGCGGATCATGGGATAGAAGCCTGAGTGCTTATG TCTTC
1320	CGSWDRSL SAVVF	2144	TGCGGATCATGGGATAGGAGCCTGAGTGCCGTGG TTTTTC
1321	CGSWDNT LGVVLF	2145	TGCGGATCATGGGATAACACCTGGGGTGTGTTCT CTTC
1322	CGSWDNR LSTVIF	2146	TGCGGATCGTGGGATAACAGACTAAGTACTGTCA TCTTC
1323	CGSWDNR LNTVIF	2147	TGCGGAAGCTGGGATAATCGATTGAACACTGTGA TTTTTC
1324	CGSWDLSP VRVLVF	2148	TGCGGTTTCATGGGATCTCAGCCCTGTACGTGTCCT TGTGTTC
1325	CGSWDLSL SAVVF	2149	TGCGGATCATGGGATCTCAGCCTGAGTGCTGTGC TTTTTC
1326	CGSWDKN LRAVLF	2150	TGCGGATCATGGGATAAAAACCTGCGTGCTGTGC TGTTTC
1327	CGSWDISL SAGVF	2151	TGCGGCTCATGGGATATCAGCCTGAGTGCTGGGG TGTTTC
1328	CGSWDIRL SAEVF	2152	TGCGGATCATGGGATATCAGACTGAGTGCAGAGG TCTTC
1329	CGSWDIKL NIGVF	2153	TGCGGATCATGGGACATCAAACCTGAATATTGGGG TATTC
1330	CGSWDFSL NYFVF	2154	TGCGGATCATGGGATTTTCAGTCTCAATTATTTTGT CTTC
1331	CGSWDAS LSTEVF	2155	TGCGGATCATGGGATGCCAGCCTGAGTACTGAGG TGTTTC
1332	CGSWDAG LRGWVF	2156	TGCGGATCCTGGGATGCCGGCCTGCGTGGCTGGG TTTTTC
1333	CGRWESSL GAVVF	2157	TGCGGAAGATGGGAGAGCAGCCTGGGGTGTGTTGG TTTTTC
1334	CGRWDFSL SAYVF	2158	TGCGGAAGATGGGATTTTAGTCTGAGTGCTTATG TCTTC

1335	CGQWDND LSVWVF	2159	TGCGGACAATGGGATAACGACCTGAGTGTTTGGG TG TTC
1336	CGPWHSSV TSGHVL	2160	TGCGGACCCTGGCATAGCAGCGTGACTAGTGGCC ACGTGCTC
1337	CGLWDAS LSAPTWVF	2161	TGCGGATTATGGGATGCCAGCCTGAGTGCTCCTA CTTGGGTG TTC
1338	CGIWHTSL SAWVF	2162	TGTGGAATATGGCACACTAGCCTGAGTGCTTGGG TG TTC
1339	CGIWDYSL DTWVF	2163	TGCGGAATATGGGATTACAGCCTGGATACTTGGG TG TTC
1340	CGIWDTSL SAWVF	2164	TGCGGCATATGGGATACCAGCCTGAGTGCTTGGG TG TTC
1341	CGIWDTRL SVYVF	2165	TGCGGAATTTGGGATACCAGGCTGAGTGTTTATG TCTTC
1342	CGIWDTRL SVYIF	2166	TGCGGAATTTGGGATACCAGGCTGAGTGTTTATA TCTTC
1343	CGIWDTNL GYLF	2167	TGTGGAATATGGGATACGAATCTGGGTTATCTCTT C
1344	CGIWDTGL SAVVF	2168	TGCGGTATATGGGATACCGGCCTGAGTGCTGTGG TATTC
1345	CGIWDRSL SAWVF	2169	TGCGGAATATGGGATCGCAGCCTGAGTGCTTGGG TG TTT
1346	CGIRDTRL SVYVF	2170	TGCGGAATTCGGGATACCAGGCTGAGTGTTTATG TCTTC
1347	CGGWSSRL GVGPVF	2171	TGCGGAGGATGGAGTAGCAGACTGGGTGTTGGCC CAGTGTTT
1348	CGGWGSG LSAWVF	2172	TGCGGAGGATGGGGTAGCGGCCTGAGTGCTTGGG TG TTC
1349	CGGWDTS LSAWVF	2173	TGCGGAGGATGGGATACCAGCCTGAGTGCTTGGG TG TTC
1350	CGGWDRG LDAWVF	2174	TGCGGAGGATGGGATAGGGGCCTGGATGCTTGGG TTTTC
1351	CGAWRNN VWVF	2175	TGCGGAGCATGGCGTAATAACGTGTGGGTG TTC

1352	CGAWNRR LNPHSHW VF	2176	TGCGGAGCATGGAACAGGCGCCTGAATCCTCATT CTCATTGGGTGTTC
1353	CGAWHMK LSAVF	2177	TGCGGAGCCTGGCACAACTGAGCGCGGTCT TC
1354	CGAWGSS LRASVF	2178	TGCGGAGCATGGGGTAGCAGCCTGAGAGCTAGTG TCTTC
1355	CGAWGSG LSAWVF	2179	TGCGGAGCATGGGGTAGCGGCCTGAGTGCTTGGG TGTTTC
1356	CGAWESSL SAPYVF	2180	TGCGGAGCATGGGAAAGTAGCCTGAGTGCCCCTT ATGTCTTC
1357	CGAWESSL NVGLI	2181	TGCGGAGCATGGGAGAGCAGCCTCAATGTTGGAC TGATC
1358	CGAWESG RSAGVVF	2182	TGCGGAGCATGGGAGAGCGGCCGGAGTGCTGGG GTGGTGTTC
1359	CGAWDYS VSGWVF	2183	TGCGGAGCTTGGGATTACAGTGTGAGTGGTTGGG TGTTTC
1360	CGAWDYS LTAGVF	2184	TGCGGAGCATGGGATTACAGCCTGACTGCCGGAG TATTC
1361	CGAWDYR LSAVLF	2185	TGCGGAGCCTGGGATTACAGACTGAGTGCCGTGC TATTC
1362	CGAWDVR LDVGVF	2186	TGCGGAGCGTGGGATGTTCGTCTGGATGTTGGGG TGTTTC
1363	CGAWDTY SYVF	2187	TGCGGAGCATGGGATACCTACAGTTATGTCTTC
1364	CGAWDTT LSGVVF	2188	TGCGGAGCATGGGATACGACCCTGAGTGGTGTGG TATTC
1365	CGAWDTT LSAVIF	2189	TGCGGAGCGTGGGATACTACCCTGAGTGCTGTGA TATTC
1366	CGAWDTS QGASYVF	2190	TGCGGCGCATGGGATACCAGCCAGGGTGCGTCTT ATGTCTTT
1367	CGAWDTS PVRAGVF	2191	TGCGGAGCATGGGATACCAGCCCTGTACGTGCTG GGGTGTTC

1368	CGAWDTS LWLF	2192	TGCGGAGCATGGGATAACCAGCCTGTGGCTTTTC
1369	CGAWDTS LTVYVF	2193	TGCGGAGCATGGGATAACCAGCCTGACTGTTTATG TCTTC
1370	CGAWDTS LTAGVF	2194	TGCGGAGCATGGGACACCAGTCTGACTGCTGGGG TGTTTC
1371	CGAWDTS LSTVVF	2195	TGCGGAGCTTGGGATAACCAGCCTGAGTACTGTGG TTTTTC
1372	CGAWDTS LSSRYIF	2196	TGCGGAGCATGGGATAACCAGCCTGAGTTCTAGAT ACATATTC
1373	CGAWDTS LSGYVF	2197	TGCGGAGCATGGGATAACCAGCCTGAGTGGTTATG TCTTC
1374	CGAWDTS LSGWVF	2198	TGCGGAGCCTGGGATAACCAGCCTGAGTGGCTGGG TGTTTC
1375	CGAWDTS LSGVLF	2199	TGCGGAGCATGGGATAACCAGTCTGAGTGGTGTGC TATTC
1376	CGAWDTS LSGLVF	2200	TGCGGAGCTTGGGATAACCAGCTTGAGTGGTCTTG TTTTTC
1377	CGAWDTS LSGFVF	2201	TGCGGAGCTTGGGATAACCAGCTTGAGTGGTTTTG TTTTTC
1378	CGAWDTS LSGEVF	2202	TGCGGAGCATGGGATAACCAGCCTGAGTGGTGAGG TCTTT
1379	CGAWDTS LSDFVF	2203	TGCGGAGCTTGGGATAACCAGCTTGAGTGATTTTG TTTTTC
1380	CGAWDTS LRTAIF	2204	TGCGGAGCATGGGATAACCAGCCTGCGAACTGCGA TATTC
1381	CGAWDTS LRLF	2205	TGCGGAGCATGGGATAACCAGCCTGCGGCTTTTC
1382	CGAWDTS LNVHVF	2206	TGCGGAGCATGGGATAACCAGCCTGAATGTTTCATG TCTTC
1383	CGAWDTS LNKWVF	2207	TGCGGAGCATGGGATAACCAGCCTCAATAAATGGG TGTTTC
1384	CGAWDTR LSARLF	2208	TGCGGAGCATGGGATAACCCGCCTCAGTGCGCGGC TGTTTC

1385	CGAWDTR LRGFIF	2209	TGCGGAGCATGGGATACCAGACTGAGGGGTTTTA TTTTC
1386	CGAWDTN LGNVLL	2210	TGCGGAGCATGGGATACTAATTTGGGGAATGTTC TCCTC
1387	CGAWDTN LGKWF	2211	TGCGGGGCATGGGATACCAACCTGGGTAAATGGG TTTTC
1388	CGAWDTG LEWYVF	2212	TGCGGAGCATGGGATACCGGCCTTGAGTGGTATG TTTTT
1389	CGAWDRT SGLWLF	2213	TGCGGAGCATGGGATAGGACTTCTGGATTGTGGC TTTTC
1390	CGAWDRS LVAGLF	2214	TGCGGAGCGTGGGATCGTAGCCTGGTTGCTGGAC TCTTC
1391	CGAWDRS LTVYVF	2215	TGCGGAGCGTGGGATAGAAGCCTGACTGTTTATG TCTTC
1392	CGAWDRS LSGYVF	2216	TGCGGAGCATGGGATAGAAGCCTGAGTGGTTATG TCTTC
1393	CGAWDRS LSAYVF	2217	TGCGGAGCATGGGATAGAAGCCTGAGTGCTTATG TCTTC
1394	CGAWDRS LSAVVF	2218	TGCGGAGCATGGGATAGAAGCCTGAGTGCGGTGG TATTC
1395	CGAWDRS LSAGVF	2219	TGCGGAGCATGGGATCGCAGCCTGAGTGCTGGGG TTTTC
1396	CGAWDRS LRIVVF	2220	TGCGGAGCGTGGGATCGCAGCCTGCGTATTGTGG TATTC
1397	CGAWDRS LRAYVF	2221	TGCGGAGCATGGGATAGAAGTCTGAGGGCTTACG TCTTC
1398	CGAWDRS LNVWLF	2222	TGCGGAGCATGGGATAGAAGTCTGAATGTTTGGC TGTTTC
1399	CGAWDRG LNVGWLF	2223	TGCGGCGCCTGGGATAGGGGCCTGAATGTCGGTT GGCTTTTC
1400	CGAWDNR LSILAF	2224	TGCGGCGCATGGGATAATAGACTGAGTATTTTGG CCTTC
1401	CGAWDND LTAYVF	2225	TGCGGAGCTTGGGATAATGACCTGACAGCTTATG TCTTC

1402	CGAWDFS LTPLF	2226	TGCGGGGCATGGGATTTTCAGCCTGACTCCTCTCTT C
1403	CGAWDDY RGSIVVF	2227	TGCGGAGCCTGGGATGACTATCGGGGTGTGAGTA TTTATGTCTTC
1404	CGAWDDR PSSAVVF	2228	TGTGGAGCATGGGATGACCGGCCTTCGAGTGCCG TGGTTTTTC
1405	CGAWDDR LTVVVF	2229	TGCGGAGCATGGGATGACAGACTGACTGTCGTTG TTTTTC
1406	CGAWDDR LGAVF	2230	TGCGGAGCGTGGGATGACAGGCTGGGTGCTGTGT TC
1407	CGAWDAS LNPGRF	2231	TGCGGAGCGTGGGATGCCAGCCTGAATCCTGGCC GGGCATTC
1408	CGAWDAG LREIF	2232	TGCGGAGCATGGGATGCCGCCTGAGGGAAATTT TC
1409	CGAWAGS PSPWVF	2233	TGCGGAGCTTGGGCTGGCAGTCCGAGTCCTTGGG TTTTTC
1410	CGAFDTTL SAGVF	2234	TGCGGAGCATTCGACACCACCCTGAGTGCTGGCG TTTTTC
1411	CETWESSL SVGVF	2235	TGCGAAACATGGGAGAGCAGCCTGAGTGTTGGGG TCTTC
1412	CETWESSL RVWVF	2236	TGCGAAACATGGGAAAGCAGCCTGAGGGTTTGGG TGTTTC
1413	CETWDTSL SGGVF	2237	TGCGAAACGTGGGATACCAGCCTGAGTGGTGGGG TGTTTC
1414	CETWDTSL SDFYVF	2238	TGCGAAACATGGGATACCAGCCTGAGTGACTTTT ATGTCTTC
1415	CETWDTSL SALF	2239	TGCGAAACATGGGATACCAGCCTGAGTGCCCTCT TC
1416	CETWDTSL RAEVF	2240	TGCGAAACATGGGATACCAGCCTGCGTGCTGAAG TCTTC
1417	CETWDTSL NVVVF	2241	TGCGAAACATGGGATACCAGCCTGAATGTTGTGG TATTC
1418	CETWDTSL GAVVF	2242	TGCGAAACATGGGATACCAGCCTGGGTGCCGTGG TGTTTC

1419	CETWDRSL SGVVF	2243	TGCGAAACATGGGATAGAAGCCTGAGTGGTGTGG TATTC
1420	CETWDRSL SAVVF	2244	TGCGAAACATGGGATAGGAGCCTGAGTGCTTGGG TGTTT
1421	CETWDRSL SAVVF	2245	TGCGAAACATGGGATCGCAGCCTGAGTGCTGTGG TCTTC
1422	CETWDRG LSVVVF	2246	TGCGAGACGTGGGATAGAGGCCTGAGTGTTGTGG TTTTC
1423	CETWDRG LSAVVF	2247	TGCGAAACATGGGATAGGGGCCTGAGTGCAGTGG TATTC
1424	CETWDHT LSVVIF	2248	TGCGAAACATGGGATCACACCCTGAGTGTTGTGA TATTC
1425	CETWDASL TVVLF	2249	TGCGAAACATGGGATGCCAGCCTGACTGTTGTGT TATTC
1426	CETWDASL SAGVF	2250	TGCGAAACATGGGATGCCAGCCTGAGTGCTGGGG TGTTT
1427	CETWDAG LSEVVF	2251	TGCGAAACGTGGGATGCCGGCCTGAGTGAGGTGG TGTTT
1428	CETFDTSL SVVVF	2252	TGCGAAACATTTGATAACCAGCCTGAGTGTTGTAG TCTTC
1429	CETFDTSL NIVVF	2253	TGCGAAACATTTGATAACCAGCCTAAATATTGTAG TCTTT
1430	CESWDRSR IGVVF	2254	TGCGAATCATGGGATAGAAGCCGATTGGTGTGG TCTTC
1431	CESWDRSL SARVY	2255	TGCGAAAGTTGGGACAGGAGTCTGAGTGCCCGGG TGTAC
1432	CESWDRSL RAVVF	2256	TGCGAATCCTGGGATAGGAGCCTGCGTGCCGTGG TCTTC
1433	CESWDRSL IVVF	2257	TGCGAATCTTGGGATCGTAGTTTGATTGTGGTGT C
1434	CESWDNN LNEVVF	2258	TGCGAAAGTTGGGATAACAATTTAAATGAGGTGG TTTTT
1435	CEIWESSPS ADDLVF	2259	TGCGAAATATGGGAGAGCAGCCCGAGTGCTGACG ATTTGGTGTTC

1436	CEAWDTSL SGAVF	2260	TGCGAAGCATGGGATAACCAGCCTGAGTGGTGC GG TG TTC
1437	CEAWDTSL SAGVF	2261	TGCGAAGCATGGGATAACCAGCCTGAGTGCCGGGG TG TTC
1438	CEAWDTSL GGGVF	2262	TGCGAAGCATGGGATAACCAGCCTGGGTGGTGGGG TG TTC
1439	CEAWDRS LTGSLF	2263	TGCGAAGCATGGGATCGCAGCCTGACTGGTAGCC TG TTC
1440	CEAWDRG LSAVVF	2264	TGCGAAGCGTGGGATAGGGGCCTGAGTGCAGTGG TATTC
1441	CEAWDNIL STVVF	2265	TGCGAAGCCTGGGATAACATCCTGAGTACTGTGG TG TTC
1442	CEAWDISL SAGVF	2266	TGCGAAGCATGGGACATCAGCCTGAGTGCTGGGG TG TTC
1443	CEAWDAD LSGAVF	2267	TGCGAAGCATGGGATGCCGACCTGAGTGGTGC GG TG TTC
1444	CATWTGSF RTGHYVF	2268	TGCGCAACATGGACTGGTAGTTTCAGAACTGGCC ATTATGTCTTC
1445	CATWSSSP RGWVF	2269	TGCGCAACATGGAGTAGCAGTCCCAGGGGGTGGG TG TTC
1446	CATWHYS LSAGR VF	2270	TGCGCAACATGGCATTACAGCCTGAGTGCTGGCC GAGTG TTC
1447	CATWHTSL SIVQF	2271	TGCGCAACATGGCATAACCAGCCTGAGTATTGTGC AG TTC
1448	CATWHSTL SADVLF	2272	TGCGCAACATGGCATAGCACCTGAGTGCTGATG TGCTTTTC
1449	CATWHSSL SAGRLF	2273	TGCGCAACATGGCATAGCAGCCTGAGTGCTGGCC GACTCTTC
1450	CATWHIAR SAWVF	2274	TGCGCAACATGGCATATCGCTCGGAGTGCCTGGG TG TTC
1451	CATWGSS QSAVVF	2275	TGCGCAACATGGGGTAGTAGTCAGAGTGCCGTGG TATTC
1452	CATWGSSL SAGGVF	2276	TGCGCAACATGGGGTAGCAGCCTGAGTGCTGGGG GTGTTTTC

1453	CATWEYSL SVVLF	2277	TGTGCAACATGGGAATACAGCCTGAGTGTTGTGC TGTTTC
1454	CATWETTR RASVVF	2278	TGCGCAACATGGGAGACCACCCGACGTGCCTCTT TTGTCTTC
1455	CATWETSL NVYVVF	2279	TGCGCAACATGGGAGACCAGCCTGAATGTTTATG TCTTC
1456	CATWETSL NVVVF	2280	TGCGCAACATGGGAAACTAGCCTGAATGTTGTGG TCTTC
1457	CATWETSL NLYVVF	2281	TGCGCAACATGGGAGACCAGCCTGAATCTTTATG TCTTC
1458	CATWETG LSAGEVVF	2282	TGCGCAACATGGGAGACTGGCCTAAGTGCTGGAG AGGTGTTC
1459	CATWESTL SVVVF	2283	TGCGCGACGTGGGAGAGTACCCTAAGTGTTGTGG TTTTTC
1460	CATWESSL SIFVVF	2284	TGCGCAACGTGGGAGAGCAGCCTGAGTATTTTTG TCTTC
1461	CATWESSL NTFYVVF	2285	TGCGCAACATGGGAAAGCAGCCTCAACACTTTTT ATGTCTTC
1462	CATWESR VDTRGLLF	2286	TGCGCAACATGGGAGAGTAGGGTGGATACTCGAG GGTTGTTATTC
1463	CATWESGL SGAGVVF	2287	TGCGCAACATGGGAGAGCGGCCTGAGTGGTGCGG GGGTGTTC
1464	CATWEGSL NTFYVVF	2288	TGCGCAACATGGGAAGGCAGCCTCAACACTTTTT ATGTCTTC
1465	CATWDYS LSAVVF	2289	TGCGCAACTTGGGATTATAGCCTGAGTGCTGTGG TGTTTC
1466	CATWDYR LSIVVF	2290	TGCGCAACATGGGATTACAGACTGAGTATTGTGG TATTC
1467	CATWDYN LGAAVF	2291	TGCGCAACATGGGATTATAACCTGGGAGCTGCGG TGTTTC
1468	CATWDVT LGVLHF	2292	TGCGCCACATGGGATGTCACCCTGGGTGTCTTGC ATTTC
1469	CATWDTT LSVWVF	2293	TGCGCAACATGGGATACAACACTGAGTGTCTGGG TCTTC

1470	CATWDTT LSVVL	2294	TGCGCAACATGGGATACCACCCTGAGTGTAGTAC TTTTC
1471	CATWDTT LSVEVF	2295	TGCGCAACATGGGATACCACCCTGAGTGTTGAGG TCTTC
1472	CATWDTSP SLSGFVV	2296	TGCGCAACATGGGATACCAGCCCCAGCCTGAGTG GTTTTTGGGTGTTC
1473	CATWDTSL TGVVF	2297	TGCGCAACATGGGATACCAGCCTGACTGGTGTGG TATTC
1474	CATWDTSL TGAVF	2298	TGCGCAACATGGGATACCAGCCTGACTGGTGC GG TGTTTC
1475	CATWDTSL TAWVF	2299	TGCGCAACATGGGATACCAGCCTGACTGCCTGGG TATTC
1476	CATWDTSL TAVVF	2300	TGCGCAACATGGGATACCAGCCTGACTGCTGTGG TTTTC
1477	CATWDTSL TAKVF	2301	TGCGCAACATGGGATACTAGCCTGACTGCTAAGG TGTTTC
1478	CATWDTSL SVVVF	2302	TGCGCAACATGGGACACCAGCCTGAGTGTTGTGG TTTTC
1479	CATWDTSL SVGVF	2303	TGCGCTACTTGGGATACCAGCCTGAGTGTTGGGG TATTT
1480	CATWDTSL SSWVF	2304	TGCGCAACATGGGATACCAGCCTGAGTTCTTGGG TGTTTC
1481	CATWDTSL SGGVL	2305	TGCGCAACATGGGATACCAGCCTGAGTGGTGGGG TACTC
1482	CATWDTSL SGGVF	2306	TGCGCAACATGGGATACCAGCCTGAGTGGTGGGG TGTTTC
1483	CATWDTSL SGGRVF	2307	TGCGCAACATGGGATACCAGCCTGAGTGGTGGCC GAGTGTTTC
1484	CATWDTSL SGDRVF	2308	TGCGCAACATGGGATACCAGCCTGAGTGGTGACC GAGTGTTTC
1485	CATWDTSL SEGVF	2309	TGCGCAACGTGGGATACTAGCCTGAGTGAAGGGG TGTTTC
1486	CATWDTSL SAVVL	2310	TGCGCAACCTGGGATACCAGCCTGAGTGCCGTGG TGCTC

1487	CATWDTSL SAVF	2311	TGCGCAACATGGGATACCAGCCTGAGTGCTGTCT TC
1488	CATWDTSL SARVF	2312	TGCGCGACATGGGATACCAGCCTGAGTGCTCGGG TGTTTC
1489	CATWDTSL SALF	2313	TGCGCAACATGGGATACCAGCCTGAGTGCCTTAT TC
1490	CATWDTSL SAHVF	2314	TGCGCAACATGGGATACCAGCCTGAGTGCTCATG TCTTC
1491	CATWDTSL SAGRVF	2315	TGCGCAACATGGGATACCAGCCTGAGTGCTGGCC GGGTGTTTC
1492	CATWDTSL SAEVF	2316	TGCGCAACATGGGATACCAGCCTGAGTGCGGAGG TCTTC
1493	CATWDTSL SADAGGG VF	2317	TGCGCAACATGGGATACCAGCCTGAGTGCTGATG CTGGTGGGGGGGTCTTC
1494	CATWDTSL RVVVF	2318	TGCGCAACATGGGATACCAGCCTGCGTGTCGTGG TATTC
1495	CATWDTSL RGVF	2319	TGCGCAACATGGGATACCAGCCTGAGAGGGGTGT TC
1496	CATWDTSL PAWVF	2320	TGCGCAACATGGGATACCAGCCTGCCTGCGTGGG TGTTTC
1497	CATWDTSL NVGVF	2321	TGTGCAACATGGGATACCAGCCTGAATGTTGGGG TATTC
1498	CATWDTSL GIVLF	2322	TGCGCAACATGGGATACCAGCCTGGGTATTGTGT TATTT
1499	CATWDTSL GARVVF	2323	TGCGCAACATGGGACACCAGCCTGGGTGCGCGTG TGGTCTTC
1500	CATWDTSL GALF	2324	TGTGCAACGTGGGATACCAGTCTAGGTGCCTTGT TC
1501	CATWDTSL ATGLF	2325	TGCGCAACATGGGATACCAGCCTGGCGACTGGAC TGTTTC
1502	CATWDTSL AAWVF	2326	TGCGCAACATGGGATACCAGCCTGGCTGCCTGGG TATTC

1503	CATWDTR LSAVVF	2327	TGCGCAACCTGGGATACCAGGCTGAGTGCTGTGG TCTTC
1504	CATWDTR LSAGVF	2328	TGCGCAACATGGGATACCAGGCTGAGTGCTGGGG TGTTTC
1505	CATWDTR LLITVF	2329	TGTGCAACGTGGGACACACGTCTACTTATTACGG TTTTTC
1506	CATWDTL LSVELF	2330	TGCGCAACATGGGACACCCTCCTGAGTGTTGAAC TCTTC
1507	CATWDTG RNPHVVF	2331	TGCGCAACATGGGATACTGGCCGCAATCCTCATG TGGTCTTC
1508	CATWDTG LSSVLF	2332	TGCGCAACATGGGATACCGGCCTGTCTTCGGTGT TGTTTC
1509	CATWDTG LSAVF	2333	TGCGCAACGTGGGATACCGGCCTGAGTGCGGTTT TC
1510	CATWDRT LSIGVF	2334	TGCGCTACGTGGGATAGGACCCTGAGTATTGGAG TCTTC
1511	CATWDRS VTAVLF	2335	TGCGCAACGTGGGATCGCAGTGTGACTGCTGTGC TCTTC
1512	CATWDRS LSGVVF	2336	TGCGCAACCTGGGATAGGAGCCTGAGTGGTGTGG TGTTTC
1513	CATWDRS LSAVVF	2337	TGCGCAACATGGGATAGAAGCCTGAGTGCTGTGG TCTTC
1514	CATWDRS LSAVPWVF	2338	TGCGCAACATGGGATAGAAGCCTGAGTGCTGTTC CTTGGGTGTTC
1515	CATWDRS LSAGVF	2339	TGCGCAACATGGGATCGCAGCCTGAGTGCTGGGG TGTTTC
1516	CATWDRS LRAGVF	2340	TGCGCAACGTGGGATAGGAGCCTGCGTGCTGGGG TGTTTC
1517	CATWDRS LNVYVL	2341	TGCGCAACATGGGATCGCAGTCTGAATGTTTATG TCCTC
1518	CATWDRIL SAEVF	2342	TGCGCAACGTGGGATCGCATCCTGAGCGCTGAGG TGTTTC
1519	CATWDRG LSTGVF	2343	TGCGCAACGTGGGATAGAGGCCTGAGTACTGGGG TGTTTC

1520	CATWDNY LGAAVF	2344	TGCGCAACATGGGATAACTACCTGGGGTGCTGCCG TGTTTC
1521	CATWDNT PSNIVVF	2345	TGCGCAACATGGGATAACACGCCTTCGAATATTG TGGTATTC
1522	CATWDNT LSVWVF	2346	TGCGCAACATGGGATAATACACTGAGTGTGTGGG TCTTC
1523	CATWDNT LSVNWVF	2347	TGCGCAACATGGGATAACACCCTGAGTGTCAATT GGGTGTTTC
1524	CATWDNT LNVFYVF	2348	TGCGCAACCTGGGATAACACACTGAATGTCTTTT ATGTTTTTC
1525	CATWDNR LSSVVF	2349	TGTGCGACATGGGATAATCGGCTCAGTTCTGTGG TCTTC
1526	CATWDNR LSAGVL	2350	TGCGCAACATGGGATAACCGCCTGAGTGCTGGGG TGCTC
1527	CATWDNR LSAGVF	2351	TGCGCAACGTGGGATAACAGGCTGAGTGCTGGGG TGTTTC
1528	CATWDNR DWVF	2352	TGCGCAACATGGGATAACAGGGATTGGGTCTTC
1529	CATWDNN LGAGVF	2353	TGCGCAACATGGGATAACAACCTGGGGTGCTGGGG TGTTTC
1530	CATWDNK LTSGVF	2354	TGCGCAACATGGGATAACAAGCTGACTTCTGGGG TCTTC
1531	CATWDNIL SAWVF	2355	TGCGCAACATGGGATAACATCCTGAGTGCCTGGG TGTTT
1532	CATWDNDI HSGLF	2356	TGCGCAACCTGGGACAACGATATACATTCTGGGC TGTTTC
1533	CATWDLSL SALF	2357	TGCGCAACTTGGGATCTCAGCCTGAGTGCCCTGTT C
1534	CATWDITL SAEVF	2358	TGCGCAACATGGGATATCACCTGAGTGCTGAGG TGTTTC
1535	CATWDISP SAGGVF	2359	TGCGCAACGTGGGATATCAGCCCAGTGCTGGCG GGGTGTTTC
1536	CATWDISL STGRAVF	2360	TGCGCAACATGGGATATCAGTCTAAGTACTGGCC GGGCTGTGTTTC

1537	CATWDISL SQVF	2361	TGCGCAACATGGGATATCAGTCTGAGTCAGGTAT TC
1538	CATWDIRL SSGVF	2362	TGCGCAACATGGGATATCAGGCTGAGTAGTGGAG TGTTTC
1539	CATWDIGP SAGGVF	2363	TGCGCAACGTGGGATATCGGCCCGAGTGCTGGCG GGGTGTTTC
1540	CATWDHS RAGVLF	2364	TGCGCAACATGGGATCACAGCCGGGCTGGTGTGC TATTC
1541	CATWDHS PSVGEVF	2365	TGCGCAACATGGGATCACAGTCCGAGTGTTGGAG AAGTCTTC
1542	CATWDHS LRVGVF	2366	TGCGCAACATGGGATCACAGCCTGCGTGTTGGGG TGTTTC
1543	CATWDHS LNIGVF	2367	TGCGCAACATGGGATCACAGCCTGAACATTGGGG TGTTTC
1544	CATWDHS LGLWAF	2368	TGCGCAACATGGGATCACAGCCTGGGTCTTTGGG CATTC
1545	CATWDHN LRLVF	2369	TGCGCCACATGGGATCACAATCTGCGTCTTGTTTT C
1546	CATWDHIL ASGVF	2370	TGCGCGACTTGGGATCACATCCTGGCTTCTGGGG TGTTTC
1547	CATWDFSL SVWVF	2371	TGCGCAACATGGGATTTTCAGCCTGAGTGTTTGGG TGTTTC
1548	CATWDFSL SAWVF	2372	TGCGCAACATGGGATTTTCAGCCTGAGTGCTTGGG TGTTTC
1549	CATWDDT LTAGVF	2373	TGCGCAACATGGGATGACACCCTCACTGCTGGTG TGTTTC
1550	CATWDDR LSAVLF	2374	TGCGCAACATGGGACGACAGGCTGAGTGCTGTGC TTTTC
1551	CATWDDR LDAAVF	2375	TGCGCAACATGGGATGACAGGCTGGATGCTGCGG TGTTTC
1552	CATWDAT LNTGVF	2376	TGCGCAACATGGGATGCGACCCTGAATACTGGGG TGTTTC
1553	CATWDAS LSVWLL	2377	TGCGCAACATGGGATGCCAGCCTGAGTGTTTGGC TGCTC

1554	CATWDAS LSGGVF	2378	TGCGCGACATGGGATGCCAGCCTGAGTGGTGGGG TGTTTC
1555	CATRD TTL SAVLF	2379	TGCGCAACACGGGATACCACCCTCAGCGCCGTTTC TGTTTC
1556	CATLGSSL SLWVF	2380	TGCGCTACATTGGGTAGTAGCCTGAGTCTCTGGG TGTTTC
1557	CATIETSLP AWVF	2381	TGCGCAACAATCGAAACTAGCCTGCCTGCCTGGG TATTC
1558	CATGDRSL TVEVF	2382	TGCGCAACAGGGGACAGAAGCCTGACTGTTGAGG TATTC
1559	CATGDLGL TIVF	2383	TGCGCTACAGGGGATCTCGGCCTGACCATAGTCT TC
1560	CASWDYR GRSGWVF	2384	TGCGCATCATGGGATTACAGGGGGAGATCTGGTT GGGTGTTTC
1561	CASWD TTL NVGVF	2385	TGCGCATCATGGGATACCACCCTGAATGTTGGGG TGTTTC
1562	CASWD TTL GFVLF	2386	TGCGCTTCATGGGATACCACCCTGGGTTTTGTGTT ATTC
1563	CASWD TSL SGGYVF	2387	TGCGCATCATGGGATACCAGCCTGAGTGGTGGTT ATGTCTTC
1564	CASWD TSL RAGVF	2388	TGCGCATCATGGGATACCAGCCTCCGTGCTGGGG TGTTTC
1565	CASWD TSL GAGVF	2389	TGCGCATCATGGGATACCAGCCTGGGTGCTGGGG TGTTTC
1566	CASWDRG LSAVVF	2390	TGCGCATCATGGGACAGAGGCCTGAGTGCAGTGG TGTTTC
1567	CASWDNV LRGVVF	2391	TGTGCTAGTTGGGATAACGTCCTGCGTGGTGTGG TATTC
1568	CASWDNR LTAVVF	2392	TGCGCGTCATGGGATAACAGGCTGACTGCCGTGG TTTTTC
1569	CASWDAS LSVAF	2393	TGCGCATCATGGGATGCAAGCCTGTCCGTCGCTTT C
1570	CASWDAG LSSYVF	2394	TGCGCTTCGTGGGATGCCGGCCTGAGTTCTTATGT CTTC

1571	CASGDTSL SGVIF	2395	TGCGCATCCGGGGATAACCAGCCTGAGTGGTGTGA TATTC
1572	CARWHTS LSIWVF	2396	TGCGCAAGATGGCATAACGAGCCTAAGTATTTGGG TCTTC
1573	CAIWDTGL SPGQVAF	2397	TGCGCAATATGGGATAACCGGCCTGAGTCCTGGCC AAGTTGCCTTC
1574	CAAWHSG LGLPVF	2398	TGCGCAGCATGGCATAGCGGCCTGGGTCTCCCGG TCTTC
1575	CAAWDYS LSAGVF	2399	TGCGCAGCATGGGATTACAGCCTGAGTGCTGGGG TGTTTC
1576	CAAWDTR LRVRLF	2400	TGCGCAGCCTGGGATACTACCCTGCGTGTTAGGC TGTTTC
1577	CAAWDTS LTAWVF	2401	TGCGCAGCATGGGATAACCAGCCTGACTGCCTGGG TTTTTC
1578	CAAWDTS LSGGVF	2402	TGCGCAGCATGGGATAACCAGCTTGAGTGGTGGGG TGTTTC
1579	CAAWDTS LSGEAVF	2403	TGCGCAGCATGGGATAACCAGCCTGAGTGGCGAGG CTGTGTTC
1580	CAAWDTS LSGAVF	2404	TGCGCAGCATGGGATAACCAGCTTGAGTGGTGC GG TGTTTC
1581	CAAWDTS LSAWVF	2405	TGCGCAGCATGGGATAACCAGCCTGAGTGCCTGGG TGTTTC
1582	CAAWDTS LSAGVF	2406	TGCGCAGCATGGGATAACCAGCCTGAGTGCTGGGG TATTC
1583	CAAWDTS LDTYVF	2407	TGCGCAGCATGGGATAACCAGCCTGGATACTTATG TCTTC
1584	CAAWDTR LSGVLF	2408	TGCGCTGCATGGGATAACCCGTCTGAGTGGTGTGT TATTC
1585	CAAWDTR LSAGVF	2409	TGCGCAGCATGGGATAACCAGGCTGAGTGCTGGGG TGTTTC
1586	CAAWDRS LSTGVF	2410	TGCGCAGCATGGGATCGCAGTCTGAGTACTGGAG TTTTTC
1587	CAAWDIRR SVLF	2411	TGCGCAGCGTGGGATATCCGCCGGTCTGTCTTTT C

1588	CAAWDHT QRLSF	2412	TGCGCTGCGTGGGATCACACTCAGCGTCTTTCCTT C
1589	CAAWDHS LSAGQVF	2413	TGCGCAGCATGGGATCACAGCCTGAGTGCTGGCC AGGTGTTC
1590	CAAVDTG LKEWVF	2414	TGCGCAGCAGTCGATACTGGTCTGAAAGAATGGG TGTTTC

[00210] CDR-участки подвергали предварительному скринингу на отсутствие аминокислотных отклонений, скрытых участков сплайсинга или сайтов рестрикции нуклеотидов. Вариант CDR наблюдали по меньшей мере у двух индивидуумов, и он включал одиночные, двойные и тройные мутации фактически на уровне зародышевой линии. Порядок сборки показан на **фиг. 21С**.

[00211] Конструировали домены VH, к которым относятся IGHV1-69 и IGHV3-30. Каждый из двух доменов VH тяжелой цепи собирали с их соответствующими инвариантными 4 каркасными элементами (FW1, FW2, FW3, FW4) и 3 вариабельными CDR (H1, H2, H3) элементами. В случае IGHV1-69 417 вариантов конструировали для H1 и 258 вариантов конструировали для H2. В случае IGHV3-30 535 вариантов конструировали для H1 и 165 вариантов конструировали для H2. В случае CDR H3 использовали ту же кассету, что и для IGHV1-69 и IGHV3-30, поскольку в обеих сконструированных конструкциях был использован идентичный FW4, а также поскольку край FW3 также был идентичен как у IGHV1-69, так и у IGHV3-30. CDR H3 содержит N-концевой и C-концевой элемент, которые комбинаторно соединяли с центральным средним элементом для создания разнообразия 1×10^{10} . N-концевой и средний элемент перекрывались глициновым кодоном «GGG». Средний и C-концевой элемент перекрывались глициновым кодоном «GGT». CDR H3 содержал 5 подпулов, которые собирали отдельно. Различные N-концевые и C-концевые элементы содержали последовательности, которые показаны в **таблице 14**.

Таблица 14. Последовательности для N-концевых и C-концевых элементов

Элемент	SEQ ID NO	Последовательность
Стебель А	2415	CARDLRELECEEWT XXX SRGPCVDPRGVAGSFDVW
Стебель В	2416	CARDMYYDF XXX EVVPADDAFDIW
Стебель С	2417	CARDGRGSLPRPKGGP XXX YDSSSEDSGGAFDIW
Стебель D	2418	CARANQHF XXX GYHYYGMDVW

Стебель E	2419	CAKHMSMQ XXX RADLVGDAFDVW
-----------	------	---------------------------

[00212] Пример 12. Усиление связывающих GPCR GLP1R белков

[00213] Антитела, имеющие участки CDR-H3 с вариантами фрагментов связывающего GPCR белка, создавали с помощью описанными в настоящем документе способов и подвергали пэннингу с помощью клеточных способов для выявления вариантов, которые были усилены в отношении связывания с конкретными GPCR, как описано в примере 10.

[00214] Выявляли варианты С-концевого пептида GLP (представлены в таблице 15), которые при включении в участок CDR-H3 антитела были многократно и селективно усилены в отношении связывания с GLP1R GPCR.

Таблица 15. Последовательности GLP1, встроенные в CDR-H3

SEQ ID NO	Последовательность
2420	CAKHMSMQEGAVTGEGQAAKEFIAWLKGRVRADLVGDAFDVW
2421	CARDGRGSLPRPKGGPQTVGEGQAAKEFIAWLKGGGLTYDSSSEDSGGAFDIW
2422	CAKHMSMQDYLVIGEGQAAKEFIAWLKGGPARADLVGDAFDVW
2423	CAKHMSMQEGAVTGEGQDAKEFIAWLKGRVRADLVGDAFDVW
2424	WAKHMSMQEGAVTGEGQAAKEFIAWLKGRVRADLVGDAFDVW
2425	CARDGRGSLPRPKGGPQTVGEGQAAKEFIAWLKGRVRADLVGDAFDVW
2426	CARANQH FYEQEFTSDVSSYLEGQAAKEFIAWLKGGIRGYHYHYYGMDVW
2427	CARANQHFTLHGEGQAAKEFIAWLKGRGQIDIGYHYHYYGMDVW
2428	CARANQHFLGAGVSSYLEGQAAKEFIAWLKGDITGYHYHYYGMDVW
2429	CARANQHFLDKGTFTSDVSSYLEGQAAKEFIAWLKGIYPGYHYHYYGMDVW
2430	CARANQHFGTSLSAGEGQAAKEFIAWLKGGSQYDSSSEDSGGAFDIW
2431	CARANQHFGFLHAQGEGQAAKEFIAWLKGGSTYGYHYHYYGMDVW
2432	CARANQHFGGKGEQAAKEFIAWLKGGGSGAGYHYHYYGMDVW
2433	CAKQMSMQEGAVTGEGQAAKEFIAWLKGRVRADLVGDAFDVW
2434	CAKHMSMQEGAVTGEGQAAKEFIAWLKGGPARADLVGDAFDVW
2435	CAKHMSMQEGAVTGEGQAAKEFIAWLKGGGLTYDSSSEDSGGAFDIW
2436	CAKHMSMQDYLVIGEGQAAKEFIAWLKGRVRADLVGDAFDVW

[00215] Несмотря на то, что в настоящем документе были показаны и описаны предпочтительные варианты осуществления настоящего раскрытия, для специалистов в настоящей области техники будет очевидно, что такие варианты осуществления приведены лишь в качестве примера. Специалистам в настоящей области техники теперь станут очевидны многочисленные варианты, изменения и замены без отступления от настоящего раскрытия. Следует понимать, что при реализации на практике настоящего раскрытия можно использовать различные альтернативы описанным в настоящем документе вариантам осуществления настоящего раскрытия. Предполагается, что приведенная далее формула настоящего изобретения определяет объем настоящего раскрытия и что посредством ее охватываются способы и структуры в пределах объема настоящей формулы изобретения, а также их эквиваленты.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Антитело, причем антитело содержит CDR-H3, содержащий последовательность под любым из SEQ ID NO: 2420-2436.

2. Антитело, причем антитело содержит CDR-H3, содержащий последовательность под любым из SEQ ID NO: 2420-2436; и причем антитело представляет собой моноклональное антитело, поликлональное антитело, биспецифическое антитело, мультиспецифическое антитело, привитое антитело, человеческое антитело, гуманизированное антитело, синтетическое антитело, химерное антитело, камелизированное антитело, одноцепочечные Fv (scFv), одноцепочечное антитело, Fab-фрагмент, F(ab')₂-фрагмент, Fd-фрагмент, Fv-фрагмент, однодоменное антитело, выделенный определяющий комплементарность участок (CDR), диатело, фрагмент, состоящий только из одного мономерного переменного домена, связанные дисульфидными связями Fv (sdFv), интратело, антиидиотипическое (анти-Id) антитело или их антигенсвязывающие фрагменты.

3. Способ ингибирования активности GLP1R, предусматривающий введение антитела по п 1 или п. 2.

4. Способ лечения метаболического нарушения, предусматривающий введение нуждающемуся в том субъекту антитела по п. 1 или п. 2.

5. Способ по п. 4, причем метаболическое нарушение представляет собой сахарный диабет II типа или ожирение.

6. Библиотека нуклеиновых кислот, содержащая множество нуклеиновых кислот, причем каждая нуклеиновая кислота кодирует последовательность, которая при трансляции кодирует остов иммуноглобулина, причем остов иммуноглобулина содержит петлю CDR-H3, которая содержит связывающий GPCR домен, и причем каждая нуклеиновая кислота содержит последовательность, кодирующую вариант последовательности связывающего GPCR домена.

7. Библиотека нуклеиновых кислот по п. 6, причем длина петли CDR-H3 составляет от приблизительно 20 до приблизительно 80 аминокислот.

8. Библиотека нуклеиновых кислот по п. 6, причем длина петли CDR-H3 составляет от приблизительно 80 до приблизительно 230 пар оснований.

9. Библиотека нуклеиновых кислот по п. 6, причем остов иммуноглобулина дополнительно содержит один или несколько доменов, выбранных из переменного домена легкой цепи (VL), переменного домена тяжелой цепи (VH), константного домена легкой цепи (CL) и константного домена тяжелой цепи (CH).

10. Библиотека нуклеиновых кислот по п. 9, причем домен VH представляет собой IGHV1-18, IGHV1-69, IGHV1-8 IGHV3-21, IGHV3-23, IGHV3-30/33m, IGHV3-28, IGHV3-74, IGHV4-39 или IGHV4-59/61.

11. Библиотека нуклеиновых кислот по п. 9, причем домен VL представляет собой IGKV1-39, IGKV1-9, IGKV2-28, IGKV3-11, IGKV3-15, IGKV3-20, IGKV4-1, IGLV1-51 или IGLV2-14.

12. Библиотека нуклеиновых кислот по п. 9, причем длина домена VH составляет от приблизительно 90 до приблизительно 100 аминокислот.

13. Библиотека нуклеиновых кислот по п. 9, причем длина домена VL составляет от приблизительно 90 до приблизительно 120 аминокислот.

14. Библиотека нуклеиновых кислот по п. 9, причем длина домена VH составляет от приблизительно 280 до приблизительно 300 пар оснований.

15. Библиотека нуклеиновых кислот по п. 9, причем длина домена VL составляет от приблизительно 300 до приблизительно 350 пар оснований.

16. Библиотека нуклеиновых кислот по п. 6, причем библиотека содержит по меньшей мере 10^5 неидентичных нуклеиновых кислот.

17. Библиотека нуклеиновых кислот по п. 6, причем остов иммуноглобулина содержит один домен иммуноглобулина.

18. Библиотека нуклеиновых кислот по п. 6, причем остов иммуноглобулина содержит пептид не более чем из 100 аминокислот.

19. Библиотека нуклеиновых кислот по п. 6, причем связывающие GPCR домены содержат пептидомиметик или низкомолекулярный миметик.

20. Библиотека белков, содержащая множество белков, причем каждый из этих белков из множества белков содержит остов иммуноглобулина, причем остов иммуноглобулина содержит петлю CDR-H3, которая содержит вариант последовательности связывающего GPCR домена.

21. Библиотека белков по п. 20, причем длина петли CDR-H3 составляет от приблизительно 20 до приблизительно 80 аминокислот.

22. Библиотека белков по п. 20, причем остов иммуноглобулина дополнительно содержит один или несколько доменов, выбранных из переменного домена легкой цепи (VL), переменного домена тяжелой цепи (VH), константного домена легкой цепи (CL) и константного домена тяжелой цепи (CH).

23. Библиотека белков по п. 20, причем домен VH представляет собой IGHV1-18, IGHV1-69, IGHV1-8 IGHV3-21, IGHV3-23, IGHV3-30/33m, IGHV3-28, IGHV3-74, IGHV4-39 или IGHV4-59/61.

24. Библиотека белков по п. 20, причем домен VL представляет собой IGKV1-39, IGKV1-9, IGKV2-28, IGKV3-11, IGKV3-15, IGKV3-20, IGKV4-1, IGLV1-51 или IGLV2-14.

25. Библиотека белков по п. 20, причем длина домена VH составляет от приблизительно 90 до приблизительно 100 аминокислот.

26. Библиотека белков по п. 20, причем длина домена VL составляет от приблизительно 90 до приблизительно 120 аминокислот.

27. Библиотека белков по п. 20, причем множество белков применяют для получения библиотеки пептидомиметиков.

28. Библиотека белков по п. 20, причем библиотека белков содержит антитела.

29. Библиотека белков, содержащая множество белков, причем множество белков содержит последовательность, содержащую различные связывающие GPCR домены, и причем длина каждого связывающего GPCR домена составляет от приблизительно 20 до приблизительно 80 аминокислот.

30. Библиотека белков по п. 29, причем библиотека белков содержит пептиды.

31. Библиотека белков по п. 29, причем библиотека белков содержит иммуноглобулины.

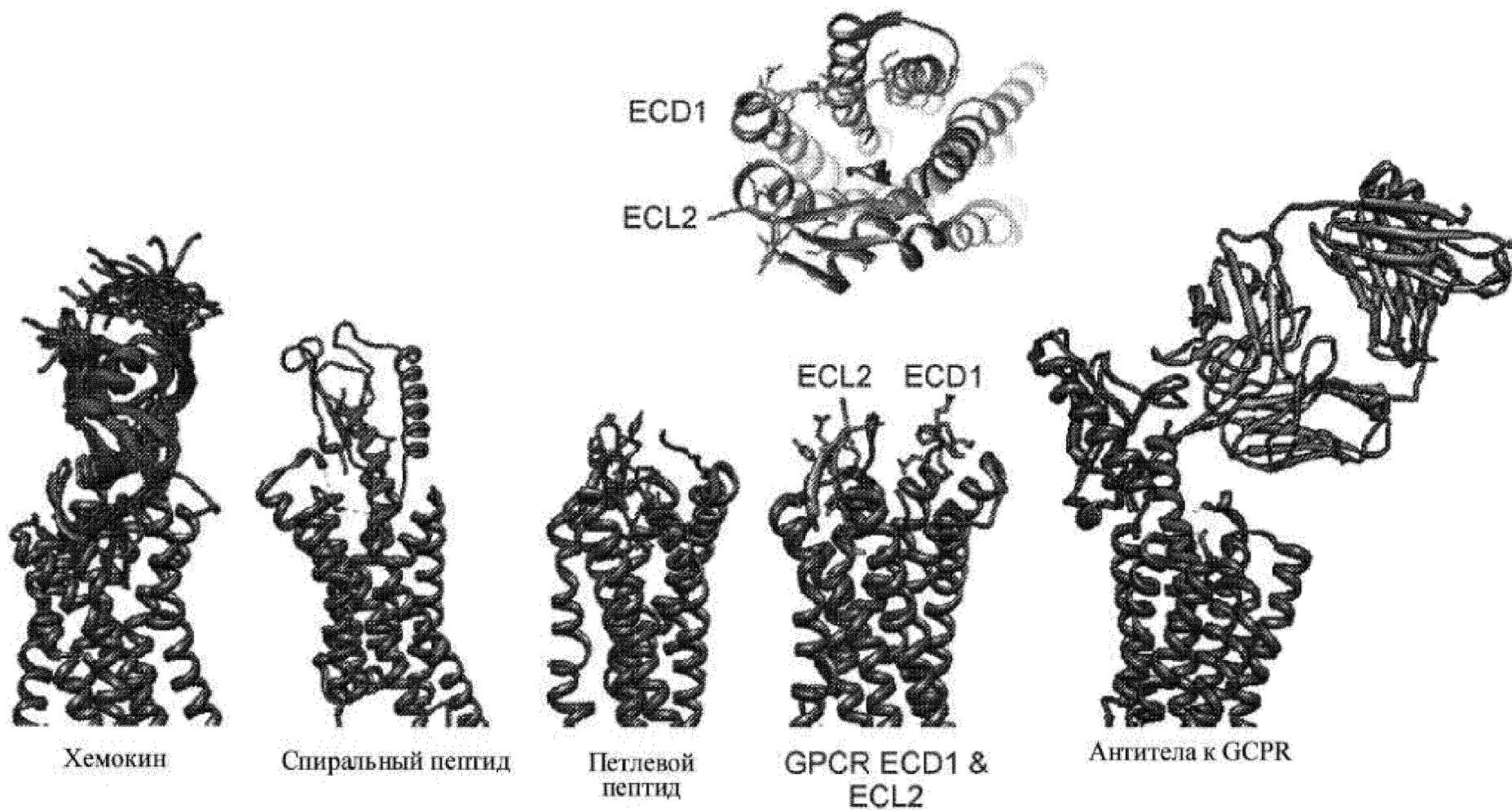
32. Библиотека белков по п. 29, причем библиотека белков содержит антитела.

33. Библиотека белков по п. 29, причем множество белков применяют для получения библиотеки пептидомиметиков.

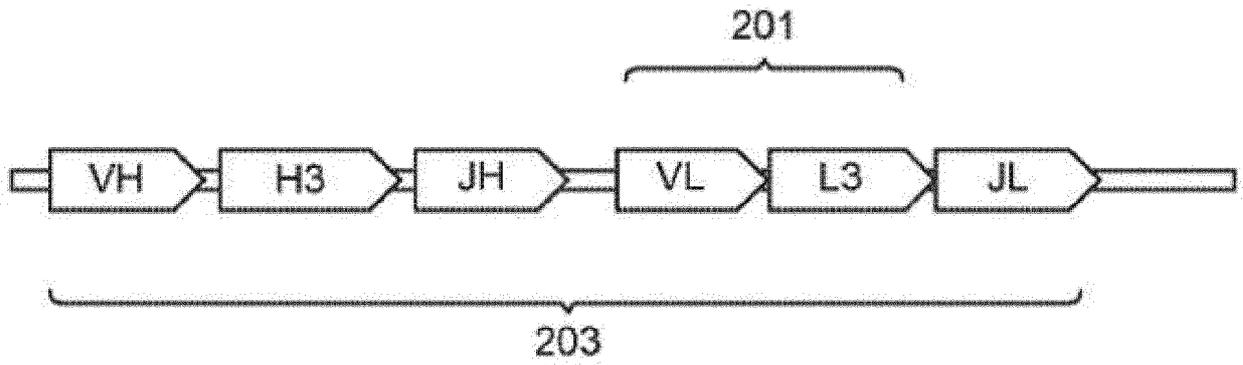
34. Библиотека векторов, содержащая библиотеку нуклеиновых кислот по любому из пп. 6-19.

35. Клеточная библиотека, содержащая библиотеку нуклеиновых кислот по любому из пп. 6-19.

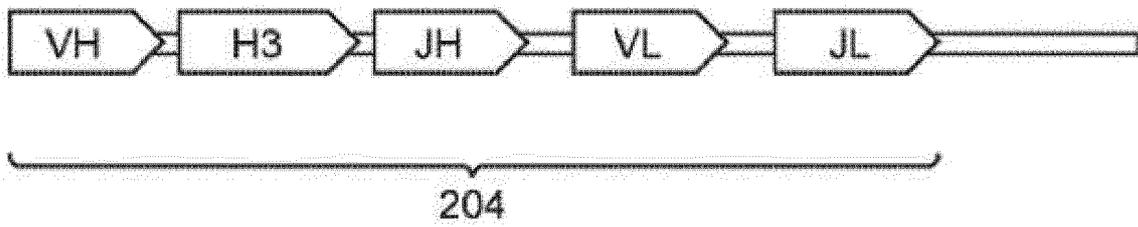
36. Клеточная библиотека, содержащая библиотеку белков по любому из пп. 20-33.



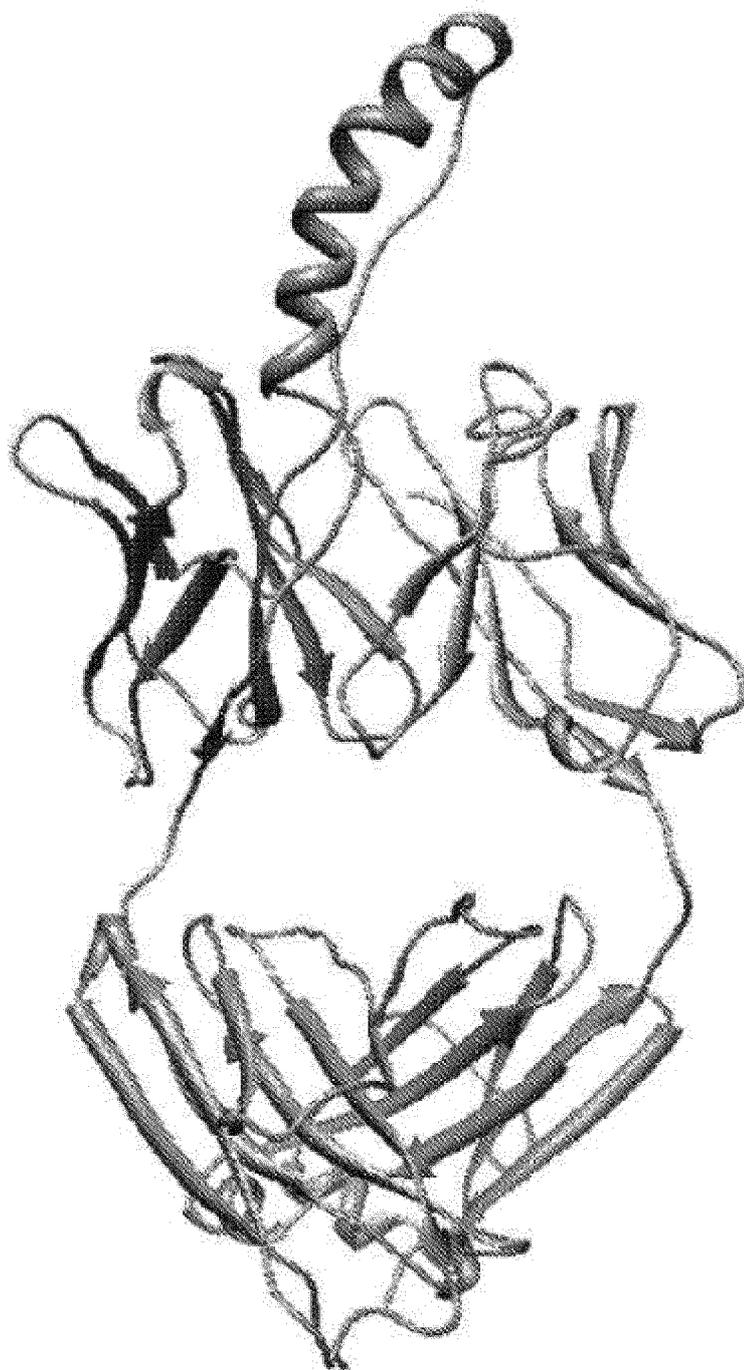
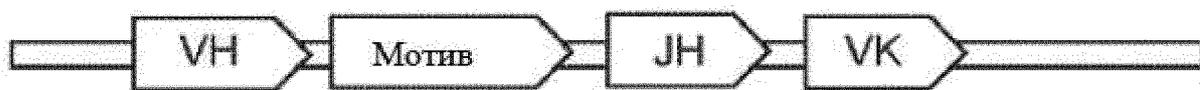
Фиг. 1



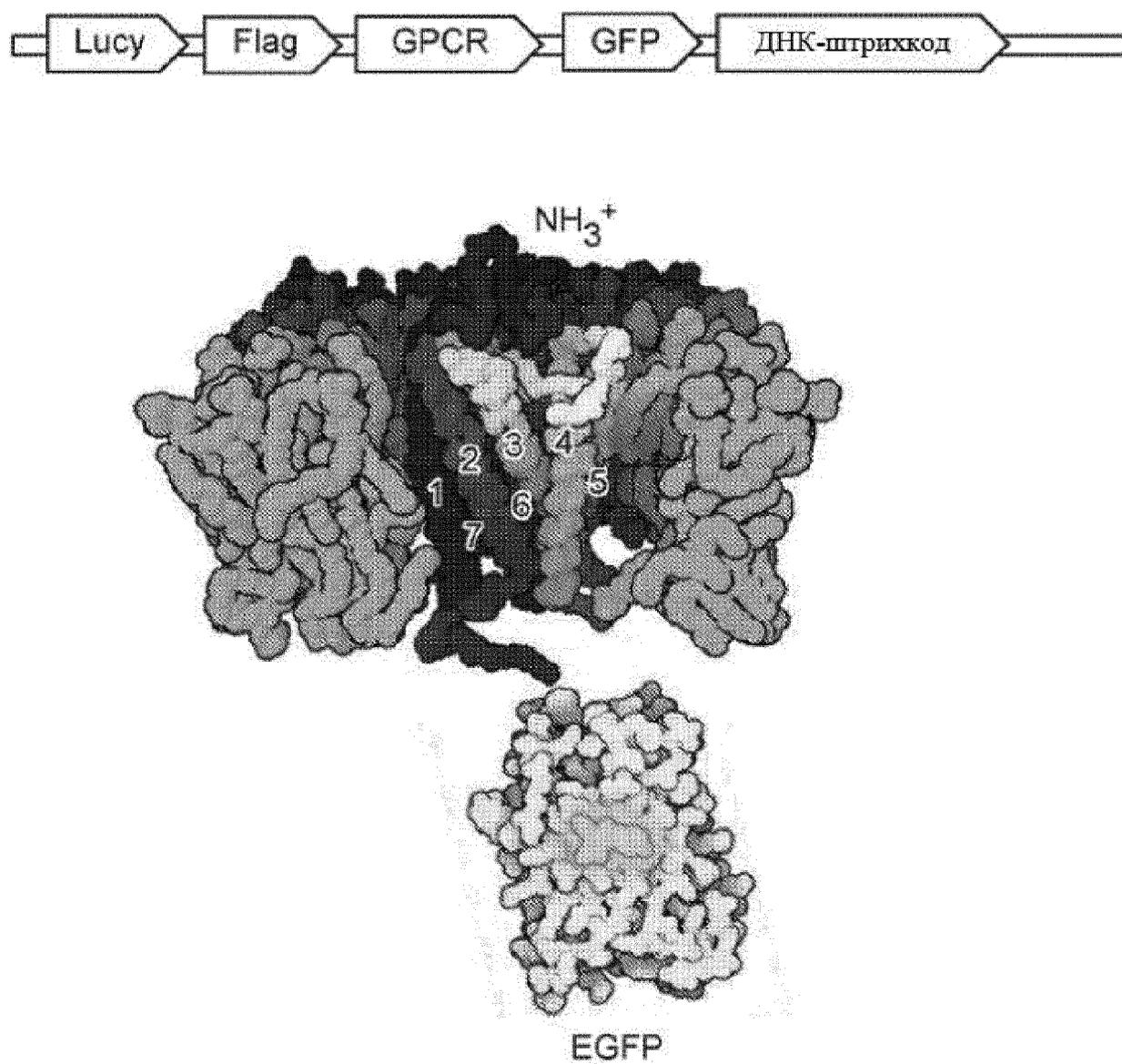
Фиг. 2А



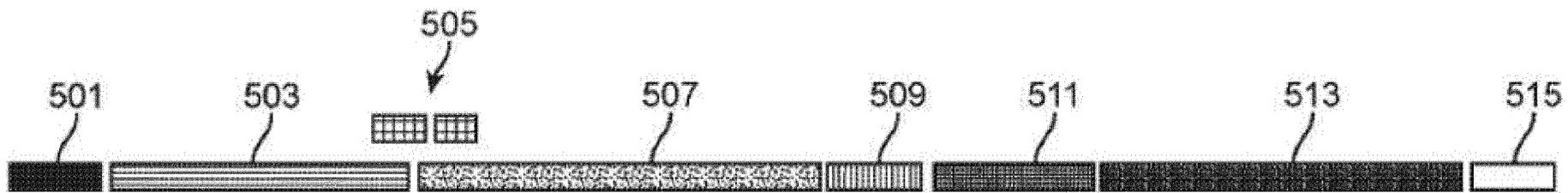
Фиг. 2В



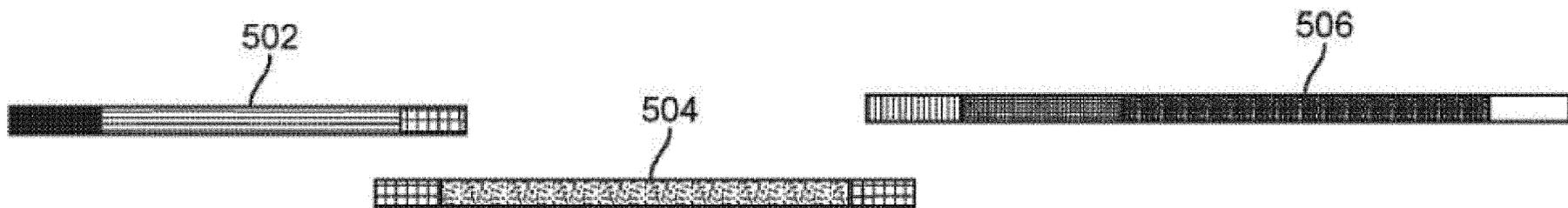
Фиг. 3



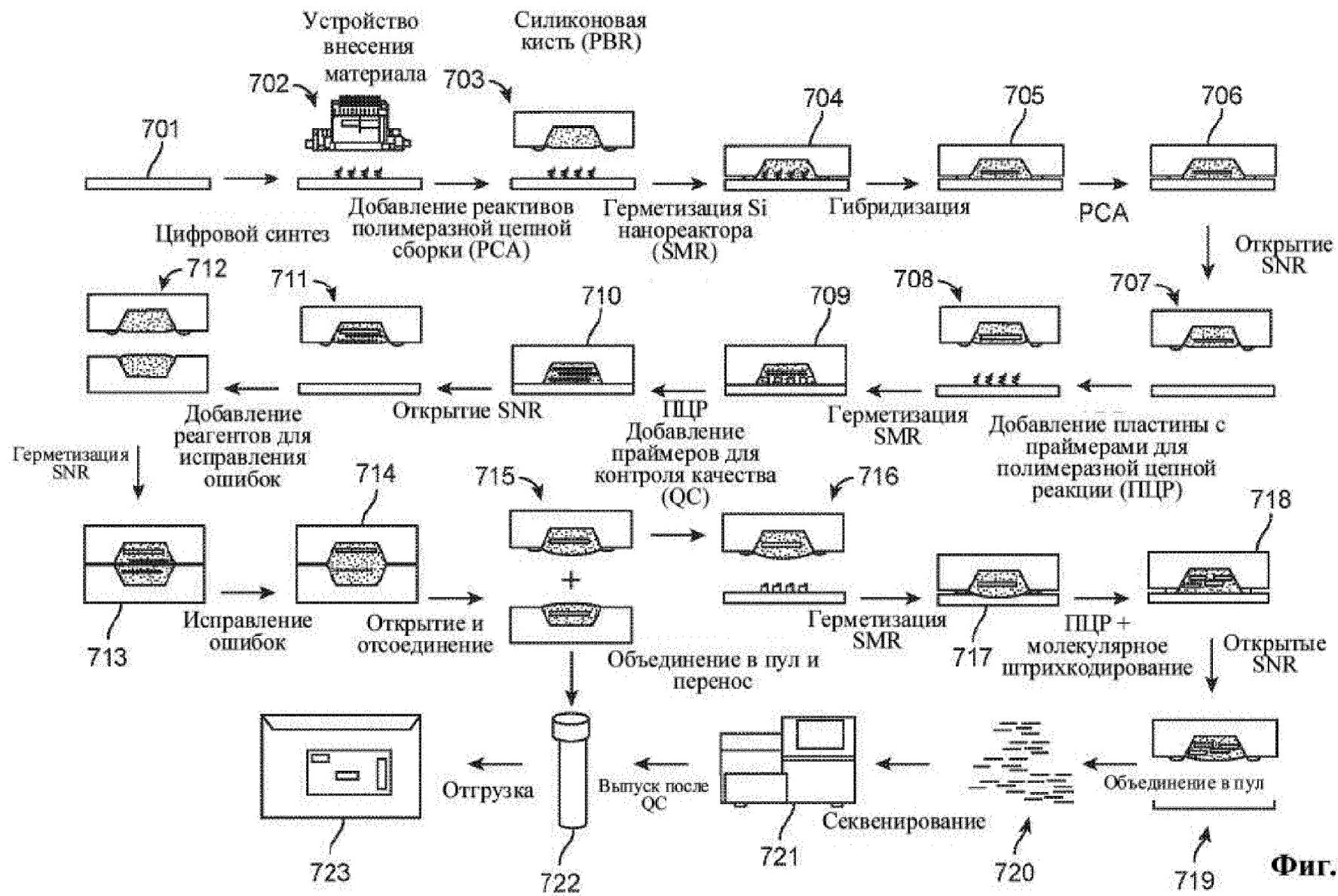
Фиг. 4



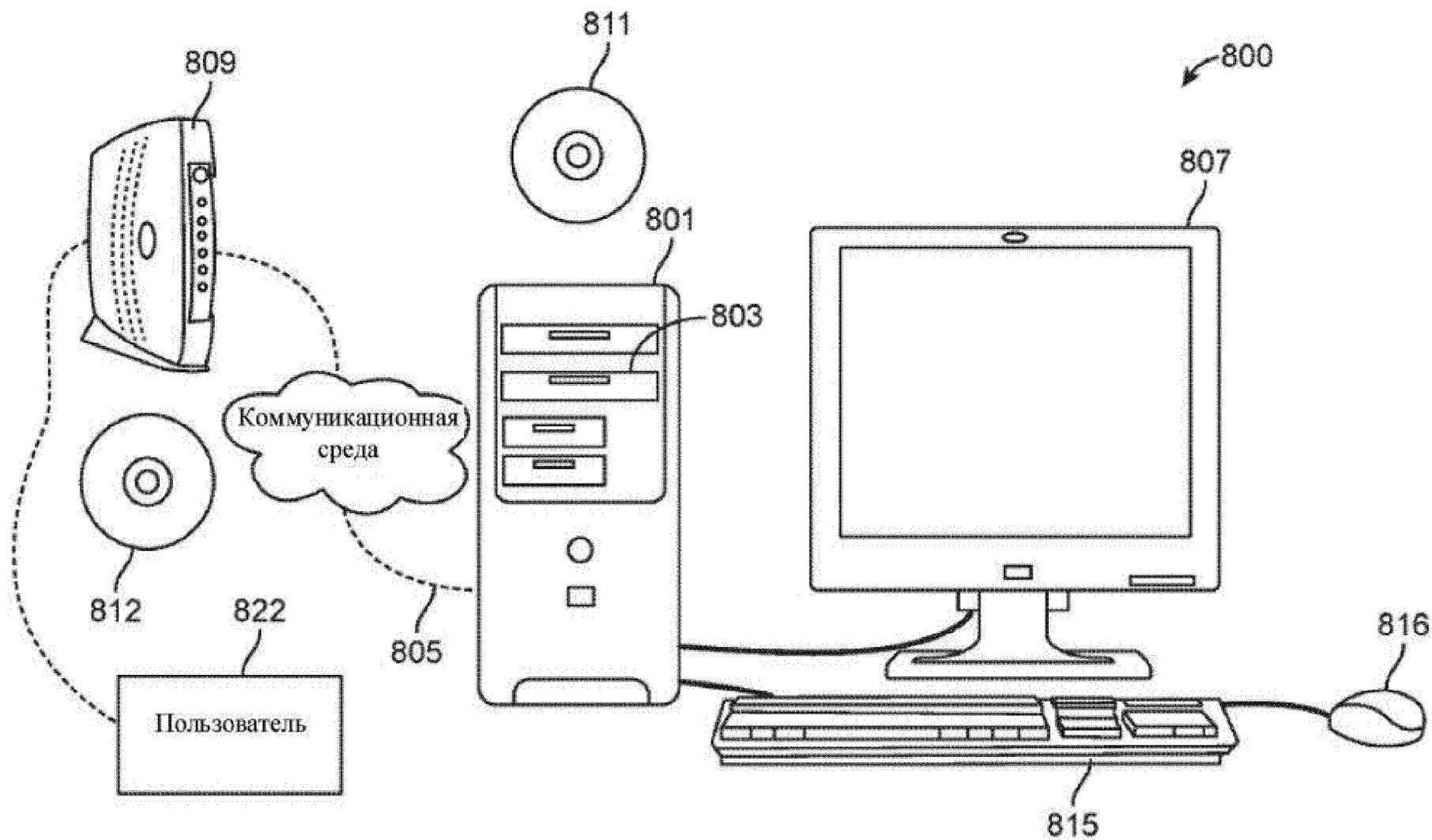
Фиг. 5



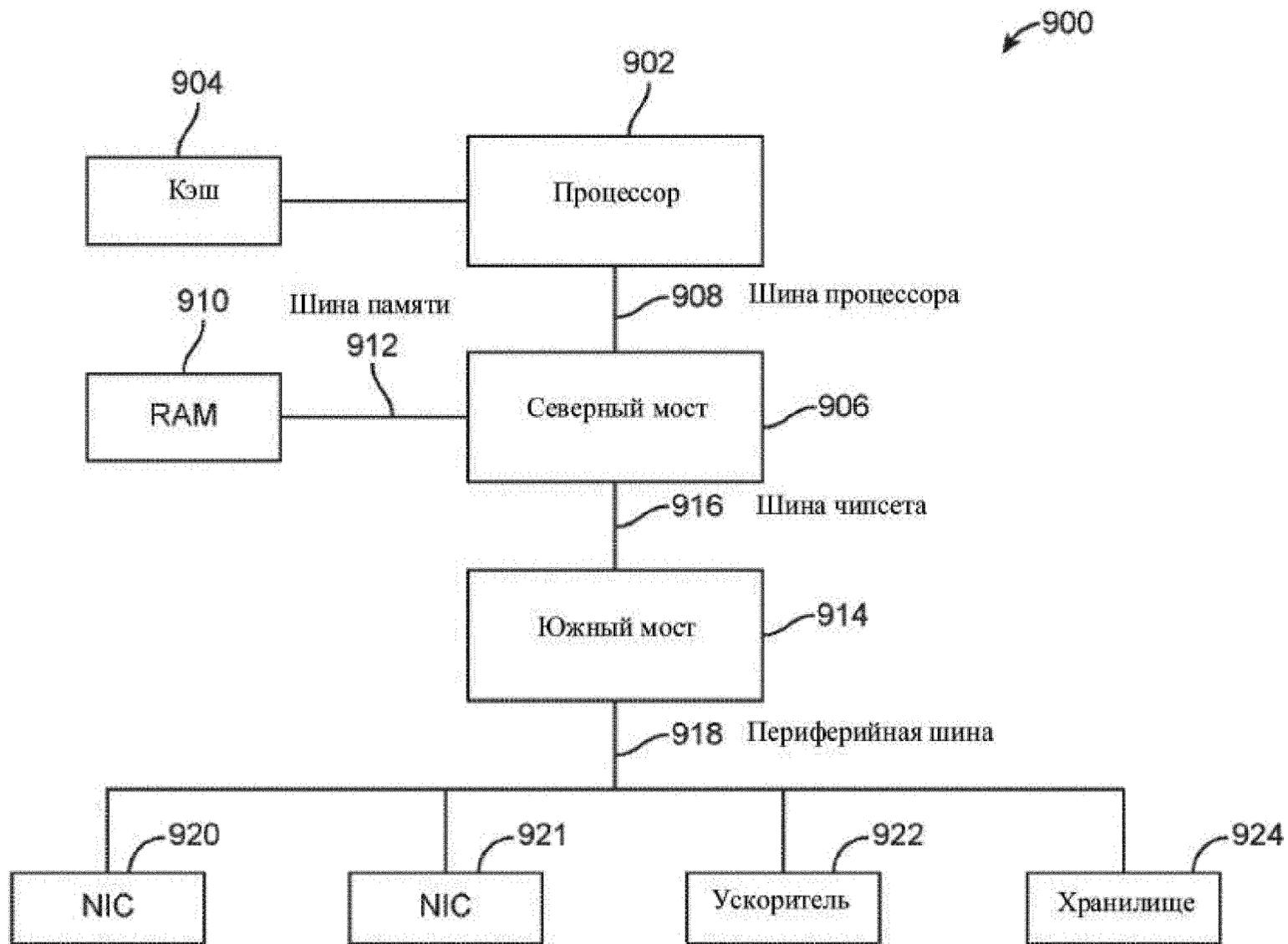
Фиг. 6



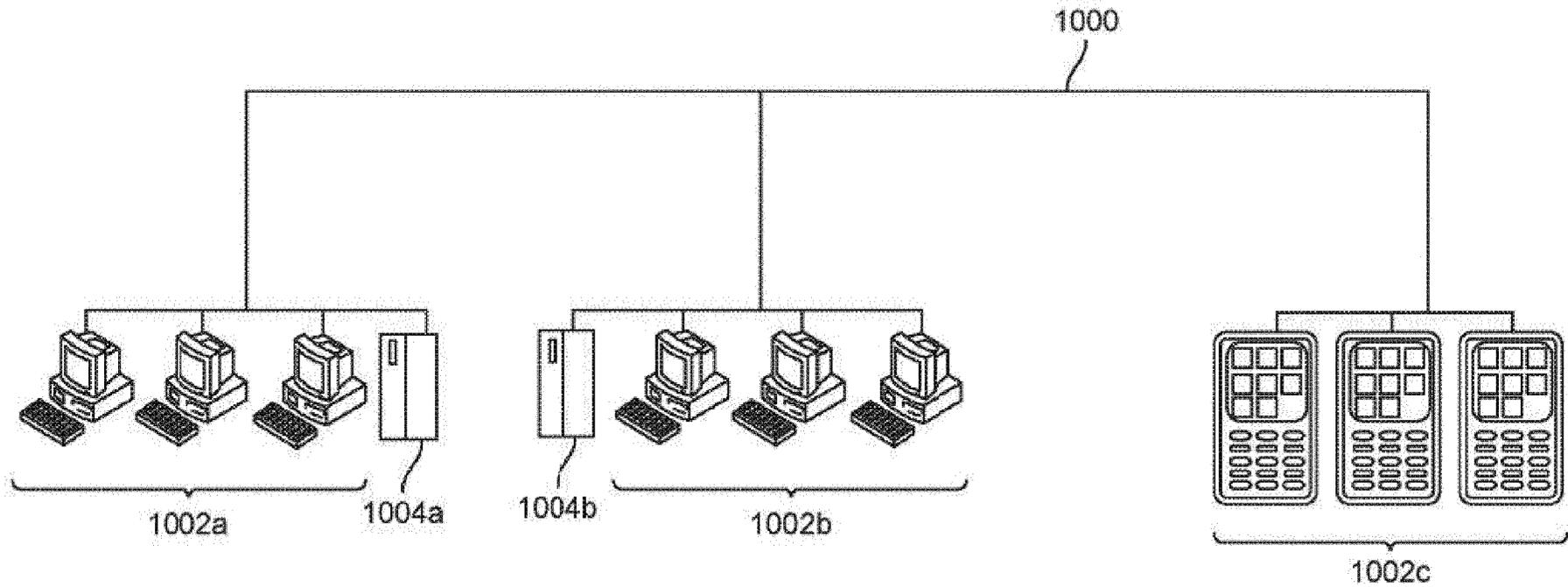
Фиг. 7



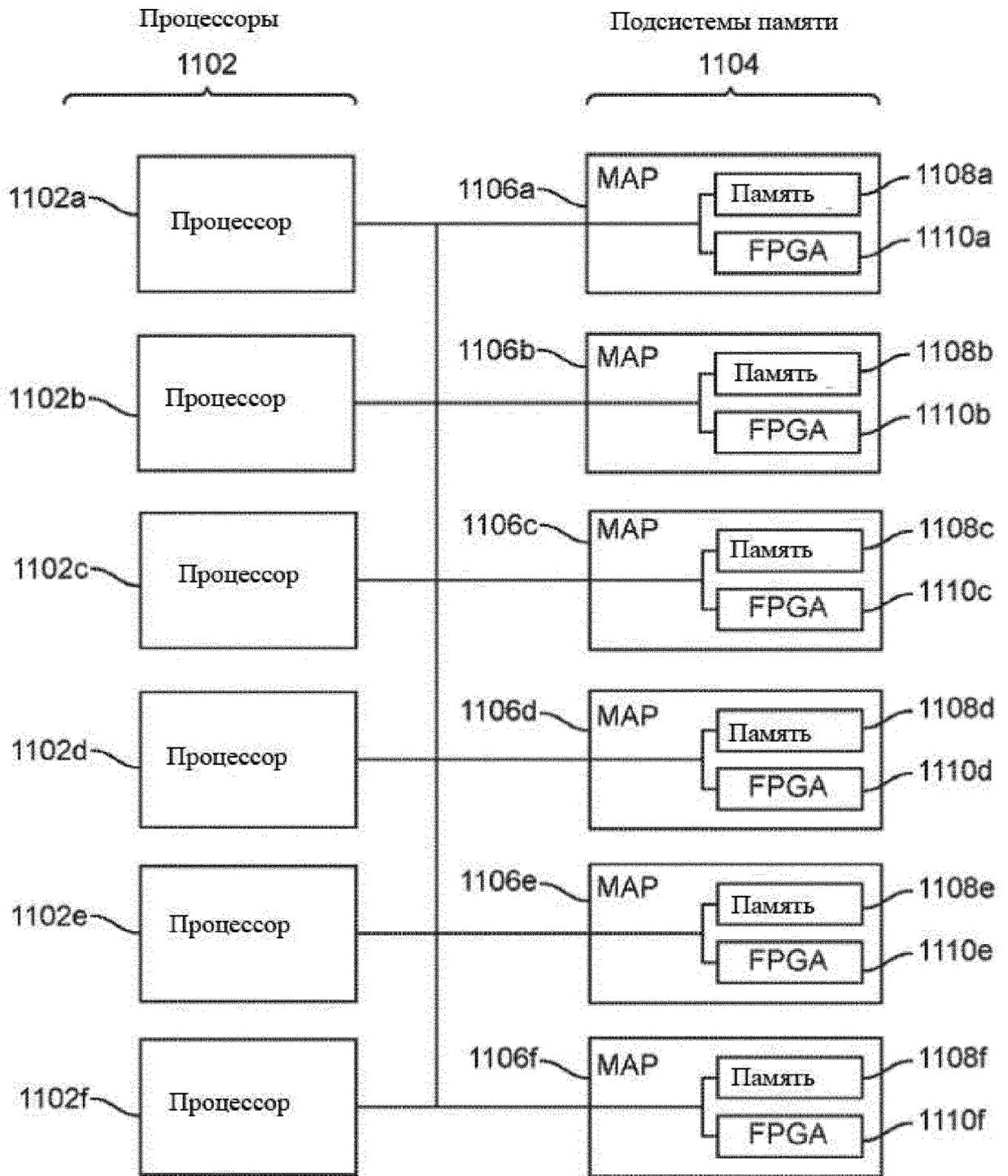
Фиг. 8



Фиг. 9



Фиг. 10

**Фиг. 11**

9 IGHV1-69;IGKV3-20 *

IGHV1-69;CAMQFTSFYIWGSYFAAENYNNEMDWW;	COQYSSFPITF ;ATFSSYAFN;GRIIPALGVNTYA;RASPNYGRSLA;GASGHPT;IGKV3-20
IGHV1-69;CARAVINYGSCTSLESDYYYYYYMEVW;	COQYGRSOTF ;DTFSDYAIT;GGIVPIFGTSKKA;RASQSVTSQHLLA;GASSRAT;IGKV3-20
IGHV1-69;CARGAALQSMCVQITRSNYFSFGMDVW;	COQYGDTPFITF;GGFSNYAIS;GGTIPIFDTHYA;RASRSVTSNYLA;GASSRAT;IGKV3-20
IGHV1-69;CARGPERWARDIAEGRPGYHHYYLDVW;	COQYGSSPRTF ;DTLSYIIT;GDIIPDLGTTDYE;RASQTVSGKFVA;GASNRAT;IGKV3-20
IGHV1-69;CARVNLQSSPSQVLRARPTYYYYGMDVW;	COQYGSSPLETF ;GTFSSYAIS;GGIIPFGTANYA;RASQSVSSSYLA;GASSRAT;IGKV3-20
IGHV1-69;CASGLLYQFDTSLSLTYNKFFYGMVDW;	COHYGDSTGTF ;GSELYAIS;GRITPVNGKANHA;RASQSVSTSLA;AASRVT;IGKV3-20
IGHV1-69;CASGPLYGFDTSSLTYSOLFYGMDVW;	COHYGDSTGTF ;GTLVTHAIT;GRITPINGKANHA;RASQSVRSTSLA;AASRAT;IGKV3-20
IGHV1-69;CATQSTNDEYDGLLCCDSYFMDVW;	COQYGTHTTF ;GIFSNFVVS;GGIFPLGGIPSYA;RASQSFSDVYLA;ATSSRAP;IGKV3-20
IGHV1-69;CAVAFNNDYLRGNRSPCRSWLESW;	COQYGTSPTE ;GTFSSYAIS;GGIIPFGSADYA;RASQSISSLSLA;GASSRAT;IGKV3-20

7 IGHV1-69;IGKV2-28

IGHV1-69;CAGGKHHRIWVENAISYDYYYGMDVW;	CMQALOTPPYTE;GLFYSYGIN;GGIIPFLGPANYA;RSSQSLQHSNGNNYLD;LGSNRAS;IGKV2-28
IGHV1-69;CARDRCQYSYVWGSYRYLEYGMDVW;	CMQALOTPRLTE;GTFSTYGIN;GGIIPLFNITIHIA;RSSQSLHNSNGNNYLD;LGSNRAS;IGKV2-28
IGHV1-69;CARDRSVRGVVSEVSPFHHHGMVDW;	CMQALQTPWTF;ASFQSYTIN;GRITPELLGSIHLA;RSSQSLRHSNGHHYVD;GASNRAS;IGKV2-28
IGHV1-69;CARNSRWKQAGATSVGAKPCNQRFYYGMDVW;	CMQALQPAHTF;GTFSTYSLG;GMVIPALRITNLA;KSSQSLLDNSNGYNYLD;LGSNRAS;IGKV2-28
IGHV1-69;CASAVVIDIPCKPKGRGNFGYYGLDWW;	CMQALQTPMYTE;GTFSSYAIT;GGIIPLFVDVKSS;RSSQSLHNSNGYNYLD;LGSNRAS;IGKV2-28
IGHV1-69;CASAVVIBIPCKPKTRCNFCYYGLDWW;	CMQALQNPMTF;GTFSTYGIS;GALIPLFDTRISA;RSSQSLHNSNGYNYLD;LGSNRAS;IGKV2-28
IGHV1-69;CASLPDSSDWWYSSPGGEYYGLDWW;	CMQALQTPPTF;GTFSSYAIS;GGIIPVFGTVHYA;RSSQSLHNSNGYNYLD;LGSNRAS;IGKV2-28

5 IGHV1-69;IGKV4-1

IGHV1-69;CAGGIWAGELYPSLPGYNYGMDVW;	CHQYSSPYSE;DTFSRYSIL;GGIIPMSATANYA;KSSQPTILYSLNNKNYLA;WASTRGS;IGKV4-1
IGHV1-69;CARDSDHLTLVGGVIDYHYYGMDVW;	COQYFSTPPTF;GTFSTYVIT;GRVITPLEKADYA;KSSQSLHSSSSKSYLT;WASTRES;IGKV4-1
IGHV1-69;CARERTDTSGDIGDSNLYYYYGMDVW;	COQYNSPRIF;DTFSRHAIIS;GGIIPYDAANYA;RASQSVLHSSNHRNYLG;WASTRES;IGKV4-1
IGHV1-69;CARGEDLKKDSSGHYYWDRIPIYSYFDSW;	COQYITPQTF;GTFNTYAIS;GGIIPIFGKPDYA;KSSQSLFYSSNNKTYLA;WASTRES;IGKV4-1
IGHV1-69;CARGGSPRITMVRILSYYYYYMDVW;	COQYSLPTF;GTFSGHAIS;GGIIPLFDTSNYA;KSSQSVLYSSNNKNYLA;WASTRES;IGKV4-1

8 IGHV3-30/33rn;IGKV2-28

IGHV3-30/33rn;CAKAMGQSGTTQSVSWGYYFYALDWW;	CMQALQTPYTF ;FTLSGYGMH;AVISFNNGSNKYA;RSSQSLLYANRYNYLD;LGSNRAS;IGKV2-28
IGHV3-30/33rn;CAKDRAREDLRFLDWPSTPPHYDMDVW;	CMQALQTEPTE ;FSFKNFAMH;SLISYDGKKSIV;RSSQSLLESNGYNYLN;LGSNRAS;IGKV2-28
IGHV3-30/33rn;CAKGEENLVLSPAKIYPRKSASIHVALDWW;	CMQALQTEPTE ;VTFSGFGLH;ALISNDGNIKHFA;RSSQSLHNSNGNNYLD;LGSTRAS;IGKV2-28
IGHV3-30/33rn;CARGELVVLSPAKIYPRKSASIHVALDWW;	CMQALQTEPTE ;VFSFAYGLH;ALISSDGSFTQYG;RSSQSLHNSNGYNYVD;LGSNRAS;IGKV2-28
IGHV3-30/33rn;CAKGGDQSSASQNAAGYYYYYGMDVW;	CMQTELEPTE ;FSVSSYAMN;SVISYDGRSEHYV;RSSQSLSRNKYIYLB;FGSNRAS;IGKV2-28
IGHV3-30/33rn;CARALGQGGTSQSVSWGYYFYALEVW;	CMQALQSPYTF ;FTLSGYGMH;AVISPNNGNKNYLA;RSSQSLYRNRYNYLD;LGSTRAS;IGKV2-28
IGHV3-30/33rn;CARGSRMAHCLSTSQSHYYGMDVW;	CMQTELEPTE ;FMFNMYGTH;ALISHDGNKNKYA;RSSQSLHNSNGYNYLD;LGSNRAS;IGKV2-28
IGHV3-30/33rn;CARGIEGQCNSPTQYPPCCYYGLEVW;	CMQSLQTPPTF ;FTFSNYRMH;AVISYDGNIKKYA;RSSQSLHNSNGYNYVD;LGSTRAS;IGKV2-28

Фиг. 12А

7 IGHV3-30/33rn; IGLV2-14

IGHV3-30/33rn; CAKDLGVQDYGFWSGYNFYGMVDW;	CSSYTGSSSTLVVF ; FAFSTYAMH; TLTSYDGSEKYYG; TGTSSSDVGGSNFVS; EVINRPS
IGHV3-30/33rn; CARALPSPKKIKVGVPSLFTTTTIVVYYYIMDWW;	CNSYIGSRNLVIF ; FTFSSYTFY; ADISDDGTNKYYA; TGISSDVGDYSCVVS; DVTNRPS
IGHV3-30/33rn; CARDFKTIEYYDMLSGSSIHYHYLMDVW;	CSSYKDTNSLVF ; FTFSDYALH; AVISYDGAYEYYA; SGSSSDIGGNYVVS; DWRDRPS
IGHV3-30/33rn; CARDHPSITLIVVVVFHTATGYAMDAA;	CASYTTSSTLLE ; FTFSEHALH; AVISNNGINKYYA; TGTSSSDIGGKYVS; EVSNRPA
IGHV3-30/33rn; CARDHPSLTLIVVVVFHTATGYTLDDVW;	CSSYTSSSTLLE ; FTFSEYAMH; AVISNTGINKYYA; TGTSGDVGAYNYVS; EVSNRPS
IGHV3-30/33rn; CARGDTAYQISTNQVTGDDFWIGYYPGALDYW;	CTSCTSCSTTTSTVWF ; FTFKNGYVH; GVIWYDGRNTHYA; TGTSSSDVGGYDYVS; EVSNRPS
IGHV3-30/33rn; CARYLQVGGTQVGNWRDYFFAMDVW;	CHSYTGPNTLEF ; FTLLTKHSMH; ATISYHGSHEYV; TGAIRDVGGNNYVS; EVDNRPS

6 IGHV3-30/33rn; IGKV3-20

IGHV3-30/33rn; CAKDNQYYDMWSSFYTYFYHYMDVW;	CQYYGGAPYIF ; FFSYGIH; AVISYDGKKNKYA; RASQTLATTSLA; GASNRAT; IGKV3-20
IGHV3-30/33rn; CARDAPYQSGGSQVELRDYFYGMVDW;	CQOYGRSSITF ; FTFNNSGMH; AVIWDGSYTDYV; RASQSVGRATSLA; GTSNRAT; IGKV3-20
IGHV3-30/33rn; CARDLSWESDNFWSALVRKQEDYSGLDWW;	CQHYGSSPLFIF ; FTFSRYGMN; AVVWFDDGSKNYA; RTSQSVSSSYVA; AASRRAT; IGKV3-20
IGHV3-30/33rn; CARDVGRIGSFALFGMVVATNIFDFW;	CQOYDKSPWTF ; FTFSCHAMH; SMISDNGINKLYS; RASQSVTNMYLA; GVSSRAA; IGKV3-20
IGHV3-30/33rn; CAREETSRIAAIGGVTHYFYGMVDW;	CQOYGDSPRTF ; FTFSYAMN; ALIEYDGSNKYYT; RASQSVASTYLA; GASSRAT; IGKV3-20
IGHV3-30/33rn; CARGAPGGHTAMVGYFYFYGMVDW;	CQOYGNSPVYTF; FTFSSYAMH; AVISYDGSNKYYA; RASQSVSSSYLA; GASSRAT; IGKV3-20

5 IGHV3-30/33rn; IGLV3-1

IGHV3-30/33rn; CAKAVYFEYSSGWSLTOYYYYGMVDW;	COAWDSSTVVF; FTFSSYGMH; AVISYDGSNKYYA; SGDKLGDKYAC; QDNKRPS; IGLV3-1
IGHV3-30/33rn; CAKDAAGYGDHDDDFEKDFGMVDW;	COAWDTRIVVF; FTLSNYGTH; ASISFDGNTYYA; SADNLRHRYVF; QNDRRPS; IGLV3-1
IGHV3-30/33rn; CAKDQLGYSITSTQVREYYYYGMVDW;	COAWDSSVVF; LTFSSYGMH; AVISYDGSNKYYA; SGDKLGDKYVC; QVTKRPS; IGLV3-1
IGHV3-30/33rn; CAKDVRDSIQWLMLFNYYSGMDVW;	CDSRDSSGDPHVLF; FTFINYDMH; AVISYDGSKYYV; QGDSLRIYYAS; GKNNRPS; IGLV3-19
IGHV3-30/33rn; CARDESLDGLMITFGGIIVTALDWW;	CSSPDSSGYHYVF; FTIGGYGFH; ALISDDGSSKSYA; QGDSLRRYYAS; GKNNRPS; IGLV3-19
IGHV3-30/33rn; CARDPPGGQSSQYKASYYYWGLDWW;	COAWDTTIGVF; FPFSSYVMH; ASISLDGNTYYA; SGDKLGDKYTC; QDKKRPS; IGLV3-1
IGHV3-30/33rn; CAREAGYDFSSGYAHKRTLDFDYW;	CCSRYTGGKQGVF; LTFGVYSWH; AVISPDGSTTSYA; QGDSLRYFAS; GNNNRPS; IGLV3-19
IGHV3-30/33rn; CARGPNHYDSSGYEGSYHYGMVDW;	CQVWDTRTYVVF; FTFRNYALH; ALISYDGNSENYE; SGDNLGNKYVS; QDTKRPS; IGLV3-1

5 IGHV3-30/33rn; IGLV1-51

IGHV3-30/33rn; CAKVGIAARPGGAIVPQYYYYFIDVW;	CGTWDNSLRGVVF ; FTFSTYGIH; AAISHDGRNKYYA; SGSSSNIVNNYVC; DNDKRPS; IGLV1-51
IGHV3-30/33rn; CAKVPSYSSGWDPRSLTYQNGMDVW;	CGAWDSSLSAVVF ; FTFRKYAMY; AVISYDGNNEYYE; SGSSSNIGNNYVS; DNNKRPS; IGLV1-51
IGHV3-30/33rn; CARDGGYYSGSVYHFLDYFYGMVDW;	CGTWDNSLRSAGVVF; FTFSSYAMH; AVILHDGSKNYA; SGSSSNIGNNYVS; DNNKRPS; IGLV1-51
IGHV3-30/33rn; CARDLLGVYFVVIIPDNPNISHYDDMEVW;	; FTFSSYVMH; AVISYDGTNKNYA; ; ; IGLV1-51
IGHV3-30/33rn; CARHLNYYDSSGFYFPQVYNGMDVW;	CGTWDSSLSAVVF ; FAFSSYGMH; AVIWDGINKYYA; SGSSSNIGNNYIS; DNNKRPS; IGLV1-51

5 IGHV3-30/33rn; IGKV3-15

IGHV3-30/33rn; CAKDPVNYGLWSGDYTGVDYDFDYW;	CQOYNWVPLTF; ITFSNFGIH; AGISYHGVYQYYA; RASQTLNMMNLA; GASTRAT; IGKV3-15
IGHV3-30/33rn; CARDAGVVRGVTNSPRTFRNYYYIMDWW;	CQOYNWVPTF; FTFSLYGLH; AVISYDGRHKYYA; GASESITMMLA; GSSSRAT; IGKV3-15
IGHV3-30/33rn; CARDNPPSHILTDKGSYYYYYAMDVW;	CQOYNSRPLTF; FTLDNYAMH; AVISFDGGHQYSA; RASQSVSRNLA; GASTRAT; IGKV3-15
IGHV3-30/33rn; CARDSFPQVAREFGDPWESKTLDTW;	CQOYRSWPRTE; SDFSDYALH; AVIWDGSKNYA; RASQSVNDYLA; GASTRAL; IGKV3-15
IGHV3-30/33rn; CARDSFSQVSAREFGDPWESKTLDLW;	CQOYNWVPRTE; SDFSDYALH; AVIWDGSKNYA; RASQSVSDYLA; GASSRAP; IGKV3-15

5 IGHV4-59/61; IGKV1-39

IGHV4-59/61; CARDNLLDIVVVPAAAMEPDANYYYYHGMDVW;	CQOSYNTPSLTF; GSINSGSYWYWN; GRIYTSGDTNYN; RASQSIINNYLN; AASSLQS; IGKV1-39
IGHV4-59/61; CARDSLITDIVVVPADQRYYYYYHGMDVW;	CQOSYSTPRAF; GSISSYYWS; GYIYYSGSTNYN; RASQSISSYLN; AASSLQS; IGKV1-39
IGHV4-59/61; CARGAKERVVMVRGVIKSDYGLDVM;	CQQSHSIPRTF; ASIRTYFWS; GYIYYGGKTDYN; RASQSVSNYLN; AALSLOS; IGKV1-39
IGHV4-59/61; CARIHYTSGSYLSWAGESIIDYYHGMDVW;	CQQSSVLPLTF; GSVTSDSYWWS; GHIYYSGNTNYH; RASESISDYLH; GAYRLOS; IGKV1-39
IGHV4-59/61; CARSWGKLAGELTPYSTSAIGHDFDW;	CQQNFSLYVTF; GSVSSGNFYWS; GYIHHTGTTKFN; RSSQKISESLN; GTSNLQR; IGKV1-39

5 IGHV4-39; IGKV3-20

IGHV4-39; CARHNIFFWASVTGTVAGTGLIDFW;	CQQYGSSPLTF; GSISSNMYFWG; GTIYYTGRTYYN; RASQSVSGSYLA; GASSRAT; IGKV3-20
IGHV4-39; CARVGVNSANQPVGVWFGQFYGLDVM;	CQQYGTSWWTF; GSITDNSYYWG; GRINNHGRTYYN; RASQSVNSNSLA; GASNRAT; IGKV3-20
IGHV4-39; CEKHASAYYDFWSSSYNNNYHGMDVW;	CQQYGSSPLYSF; GSISSSNYYWG; GCIYYSGTYYN; RASQSISSNLA; GASSRAT; IGKV3-20
IGHV4-39; CVRHPVGNYYDFFSGAQRGOYVDFW;	CQQYSDSPPTF; GSVSSRGYQWG; GTVYYSGNTYYT; RASQRISSSYLA; GASSRAT; IGKV3-20
IGHV4-39; CVRSRKAGVILGFISGGVNYIPGYFDYW;	CQHYGTSSRDTF; GSISSPNFHWG; GSVHYNGNTHYN; RASQSVTSRYLA; GASTRAT; IGKV3-20

5 IGHV3-23; IGKV3-11

IGHV3-23; CAKEGFVLTLSGGITLLRGPISAYYDYGMDVW;	CQQRSE ; ITFNSYAMG; SSISSPSDSTYYA; RASQSVGSDLA; DASSRAT; IGKV3-11
IGHV3-23; CAKNLEKSKIVVVPAAANPPHAFDIW;	CQQRSNWPTF ; FTFSSYAMS; SAISGSGGSTYYA; RASQSVSSSYLA; DASNRAT; IGKV3-11
IGHV3-23; CAKSVSYSHLSPTDPSNFWYSDLW;	CQORTSWLF ; FTFSGYAMS; SGISASGDKAISA; RTSQTINNYLA; DAFNRAT; IGKV3-11
IGHV3-23; CARDILGKEMPKVKVYVGYWYFDLW;	CQHRNWPVRVTF; FTFESNHAMY; SAIAGFSSSSYYA; RASQSVGTFLA; DASSRAT; IGKV3-11
IGHV3-23; CVGWPFQSNESQOSVWLGYSLDVM;	CQQRDWPTE ; ISFNTRYMT; AGVSSGATGTYA; RASRNI DFHLA; DASNRAT; IGKV3-11

5 IGHV3-23; IGKV2-28

IGHV3-23; CAKAAGCSGGSCHREMGFYYYAMAVW;	CMQALQTPYTF ; FMFGSYAMS; SVISGSGDGTYYA; RSSQSLLYNNNEYNYLD; LSSTRAS; IGKV2-28
IGHV3-23; CAKGGDYPGYCSPSQYRHFYYYGLDVM;	CMQALQTPRTE ; FTFSSYAMS; SGISGSGGSTYYA; RSSQSLLHSNGYNYLD; LGSNRAS; IGKV2-28
IGHV3-23; CAKGSGMLTLIRGPATVNVYGYGMDVW;	CMQALQIPPAF ; FAFSSDPMMS; SVISSSGSSTDYA; RSSQSLLHSNGYNYLD; LGFNRRAS; IGKV2-28
IGHV3-23; CAKLGATRPNIAVIPVAIDMKKATVKVRGSEATDYW;	CMOGLQTPYTF; FTLTSDAMS; SSISDSGDKTDYA; RSSQSLLHDLGNYNYLD; LGSNRAS; IGKV2-28
IGHV3-23; CARSSPVNYSTTYQHYYYNGLDVM;	CMQALQTPHTE ; FIFSNYPMS; SAISSNSGTPYYA; RSSQSLLHINEYNYLD; LASNRAS; IGKV2-28

5 IGHV3-21; IGKV3-20

IGHV3-21; CARDLFTPGRLFSGVVEVYVYHGMDVW;	CQOFRISPPTF ; FSLSNYSLN; SSSSAGRYHIDYA; RASQSVGSTYFA; GASSRAT; IGKV3-20
IGHV3-21; CARDLKSWYTTSGVRVVPNYDFAMDVM;	CQHYDTSLGYTE ; FTFNDFSKN; SSISSGSTFIYYA; RASESVSSTYLA; GSSRRAT; IGKV3-20
IGHV3-21; CARDNVGHYYDSSGYHAYYYYGMDVW;	CQQYGSSF ; FTFSSYSMN; SSISSSSSYIYYA; RASQSVSSSYLA; GASSRAT; IGKV3-20
IGHV3-21; CARENFLISGASGLLVRDYNGLDVM;	CQHYGTSPRITE ; FIFSSYGMT; SSINDTSTYINYA; RASQSVSRAYLA; GASNRAT; IGKV3-20
IGHV3-21; CASLFITMVRVIEPQGHYYYGMDVW;	CQQYGSSPFLTE ; FTFSSYSMN; SSISSSSSYIYYA; RASQSVSSSYLA; GASSRAT; IGKV3-20

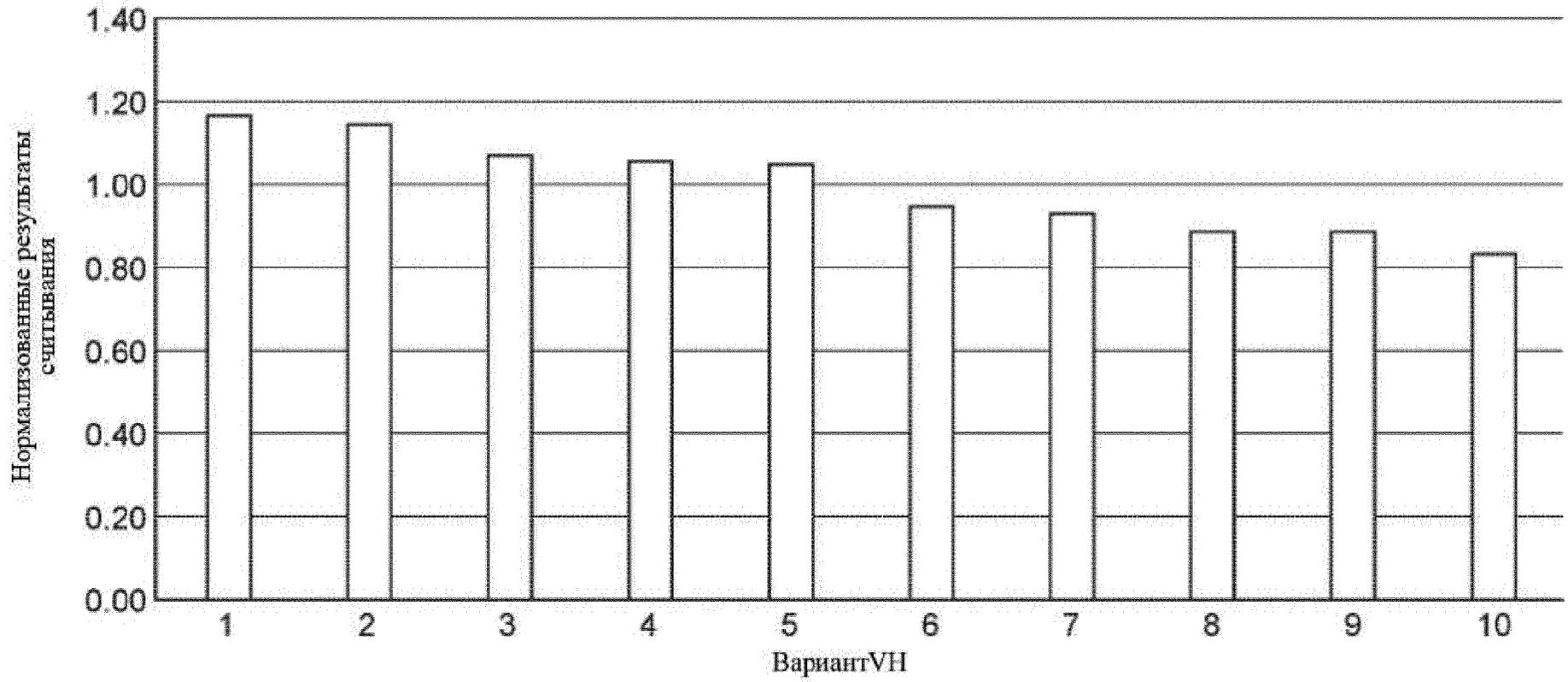
4 IGHV4-39; IGLV2-14

IGHV4-39; CARPLASIFGVVLMGHGGYYFDHW;	CSSYTSNTVVE ; DSISNNHYYWG; GSLSSTGSTYHK; TGTSIDLNNYQFVS; EVSSRPS; IGLV2-14
IGHV4-39; CARDLGGDFWSSGHNLIFYYYAMDVM;	CSSYTSNTVLE ; GSITSTINFWA; GHIYYDGTTHYN; TGTSSDVGGFNYS; EVSHRPS; IGLV2-14
IGHV4-39; CARRVFGYDYDRSGHYRVVHDAFDIW;	CSSYTRSSVVE ; GSISSSSFWG; GNI FYTGTSTYIN; TGTSSDVGGNYVS; DVSNRPS; IGLV2-14
IGHV4-39; CASGHSAYYDFWSSGYLELNDAFDIW;	CSSYTSASTHVIF; ASISTSSYYWG; GSVYYSGSTYIN; TGTSSDIGGYKYS; AVNNRPS; IGLV2-14

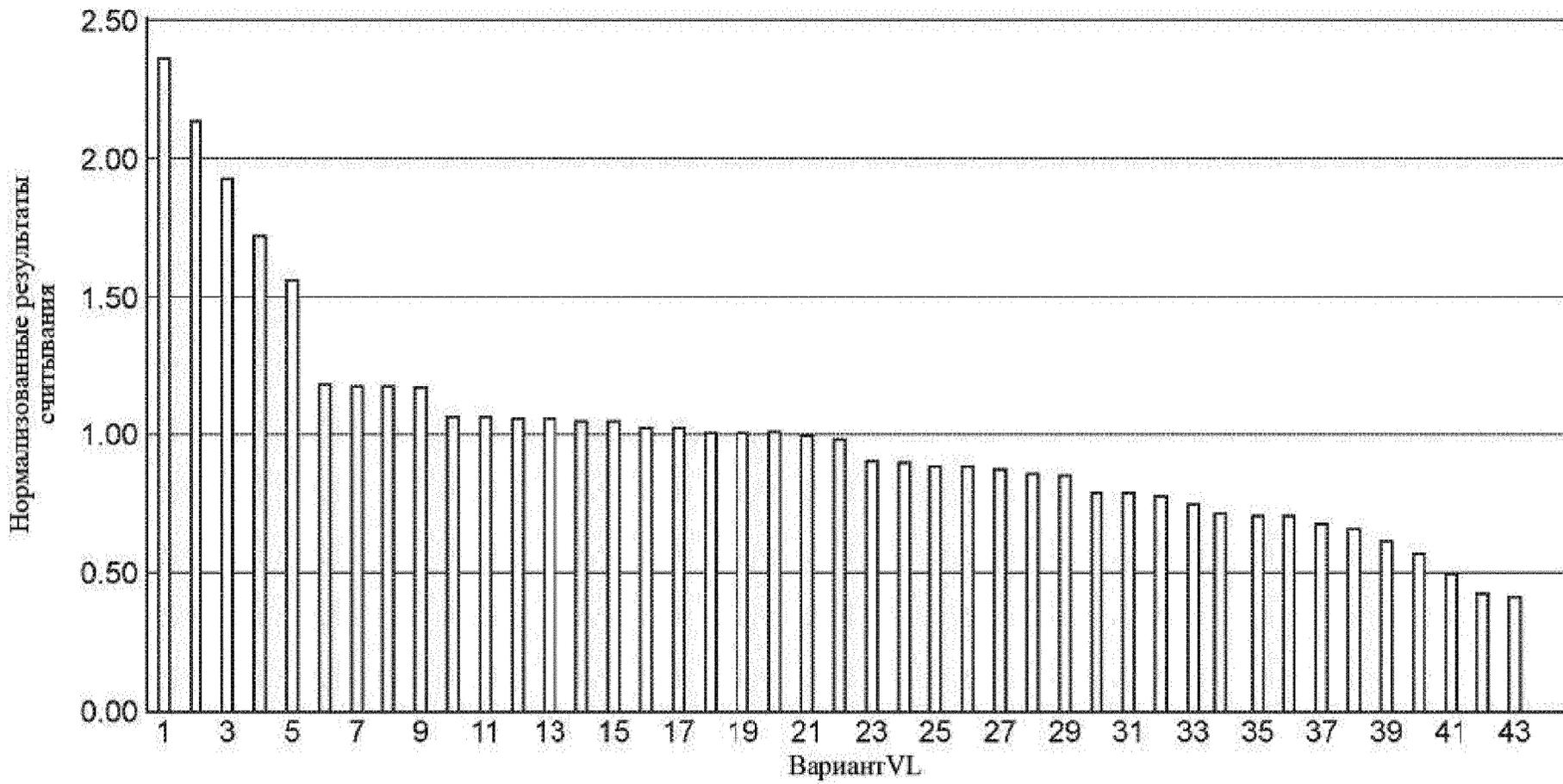
Фиг. 12С

<u>PDB ID</u>	<u>CDR-H3</u>	<u>Длина CDR-H3</u>
4ocw_H_mol	CARDLRELECEEWTLNYYDFGSRGPCVDPRGVAGSFDVW	40 VRC26
4od3_H_mol	CAKDLREDECEEWWSDYDFGKLLPCRKSRGVAGVFDKW	39 VRC26
4ocs_H_mol	CAKDMREYECEYWTSDYYDFGRPQPCIDRRGVVGI FDMW	39 VRC26
4od1_H_mol	CAKDLREDECEEWWSDYDFGKQLPCRKSRGVAGIFDGW	39 VRC26
4odh_H_mol	CAKDLGESENEEWATDYDFSIGYPGQDPRGVVGAFDIW	39 VRC26
3uls_H_mol	CLTGSKHRLRDYFLYNEYGPNYEEWGDYLATLDVW	35 VRC26
3mug_H_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
3lrs_E_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
3mme_C_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
3u36_H_mol	CVREAGGPDYRNGYNYDFYDGYNYHYMDVW	32 PG16
3u36_C_mol	CVREAGGPDYRNGYNYDFYDGYNYHYMDVW	32 PG16
3mme_A_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
3lrs_C_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
3mug_J_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
3mug_D_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
3mug_B_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
3u4e_A_mol	CVREAGGPDYRNGYNYDFYDGYNYHYMDVW	32 PG16
3u2s_H_mol	CVREAGGPDYRNGYNYDFYDGYNYHYMDVW	32 PG16
3u4e_H_mol	CVREAGGPDYRNGYNYDFYDGYNYHYMDVW	32 PG16
3lrs_H_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
3mme_H_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
3mug_F_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
3u2s_A_mol	CVREAGGPDYRNGYNYDFYDGYNYHYMDVW	32 PG16
3lrs_A_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
3u36_A_mol	CVREAGGPDYRNGYNYDFYDGYNYHYMDVW	32 PG16
3u36_E_mol	CVREAGGPDYRNGYNYDFYDGYNYHYMDVW	32 PG16
3mug_L_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
4dgo_H_mol	CAREAGGPIWHDDVKYYDFNDGYNYHYMDVW	32 PG16
4nuj_B_mol	CARMFQESGPPRFDSWSGRNYYYYSGMDVW	30 PGT152
4nug_H_mol	CARMFQESGPPRLDRWSGRNYYYYSGMDVW	30 PGT152
3q6g_l_mol	CARDVVITSHEDDFGDYTTGEYYGLDSW	29 rmAb
3q6g_H_mol	CARDVVITSHEDDFGDYTTGEYYGLDSW	29 rmAb
lrzf_H_mol	CASN SIAGVAAAGDYADYDGGYYYDMDVW	29 loop-d-loop
ljto_A_mol	CAADSTIYASYECGBGLSTGGYGYDSW	28
4fqr_c_mol	CAKHMSMQQVVSAGWERADLVGD AFDVW	28
4fqr_m_mol	CAKHMSMQQVVSAGWERADLVGD AFDVW	28
4fqr_i_mol	CAKHMSMQQVVSAGWERADLVGD AFDVW	28
4fqr_k_mol	CAKHMSMQQVVSAGWERADLVGD AFDVW	28
4fqr_u_mol	CAKHMSMQQVVSAGWERADLVGD AFDVW	28
4fnl_l_mol	CAKHMSMQQVVSAGWERADLVGD AFDVW	28
4fqr_w_mol	CAKHMSMQQVVSAGWERADLVGD AFDVW	28

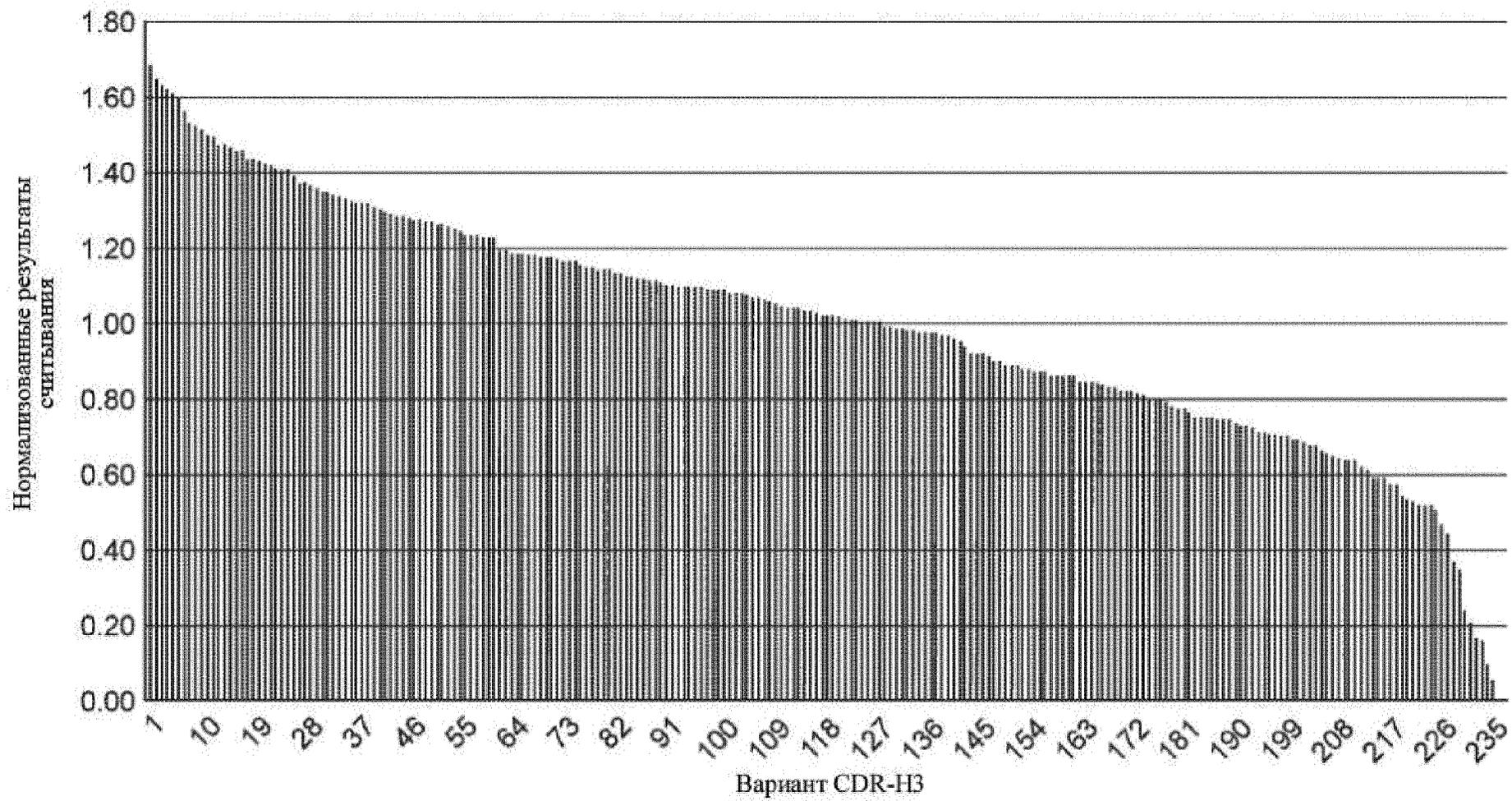
Фиг. 13



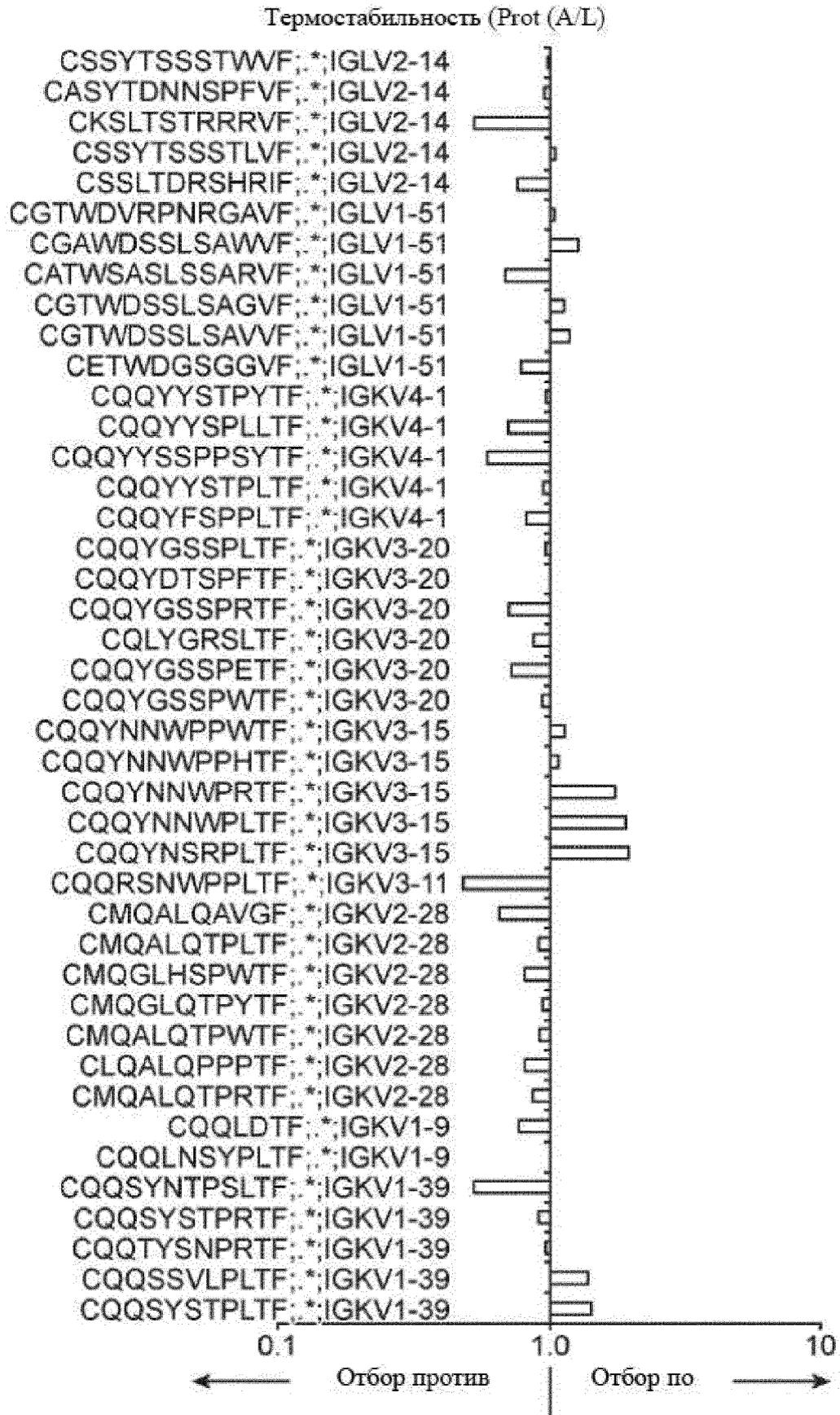
Фиг. 14



Фиг. 15

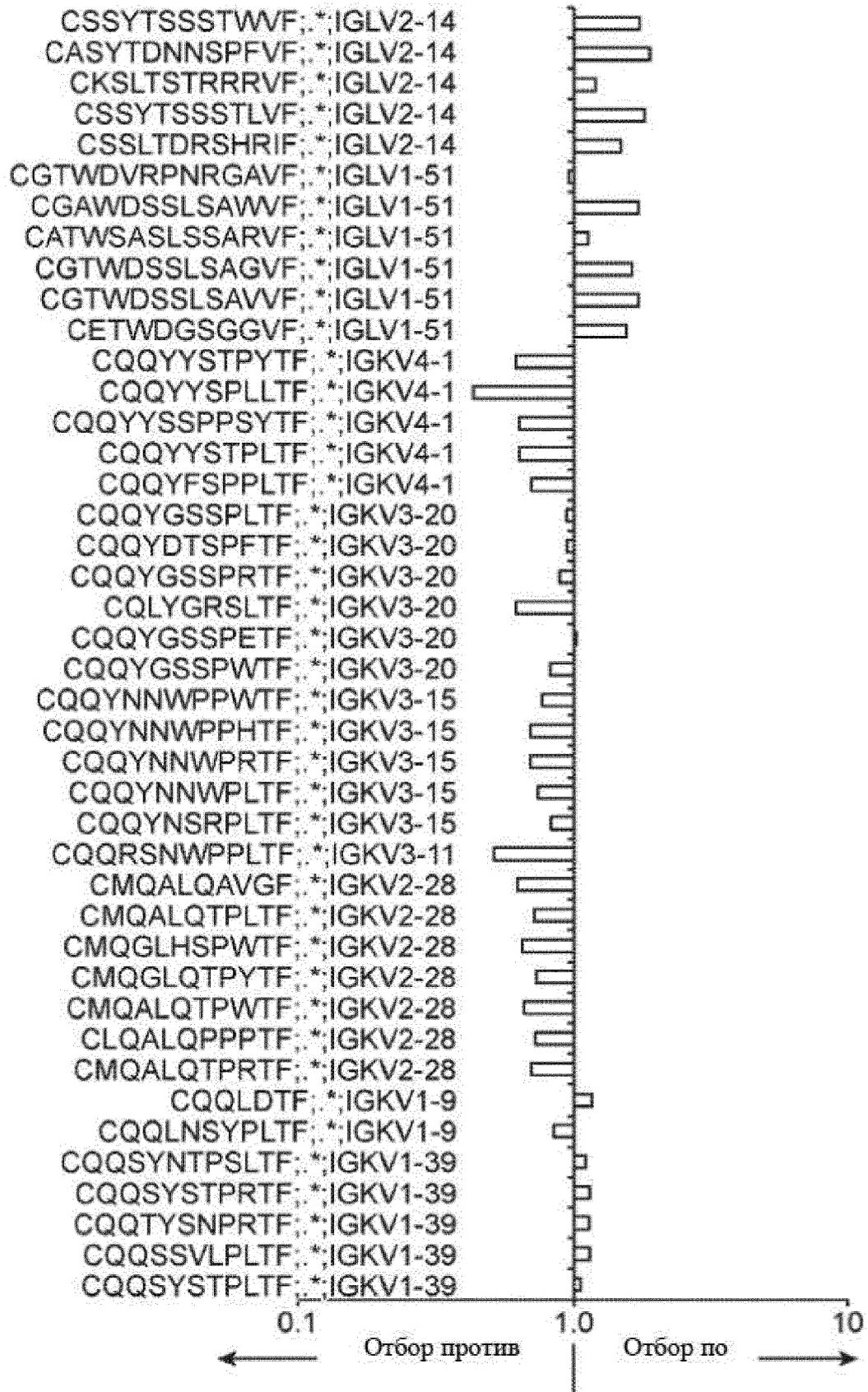


Фиг. 16

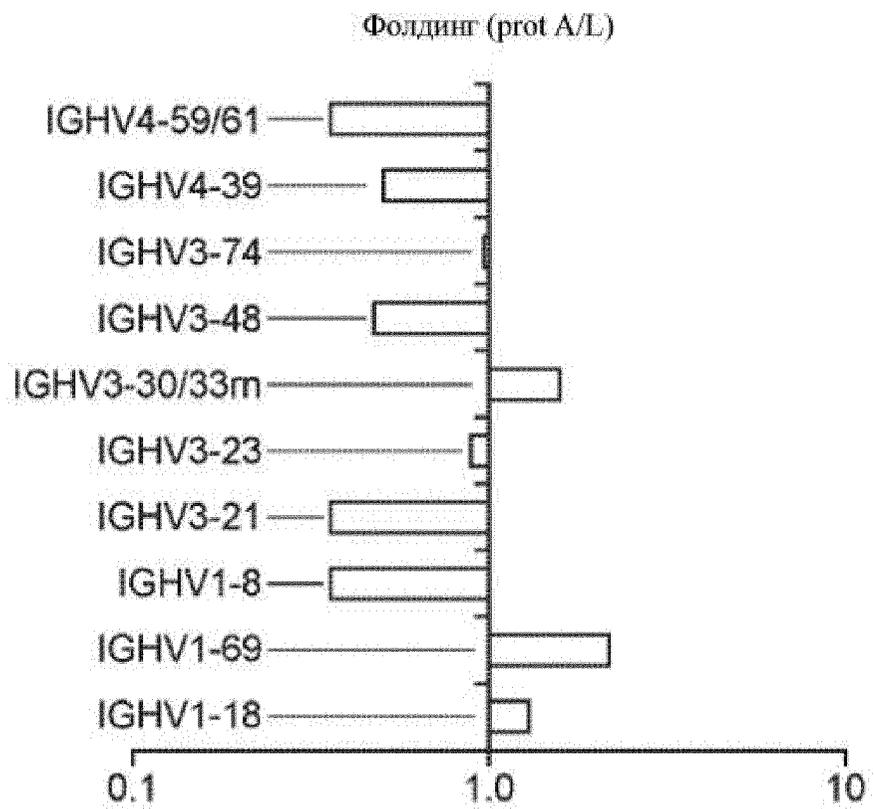


Фиг. 17В

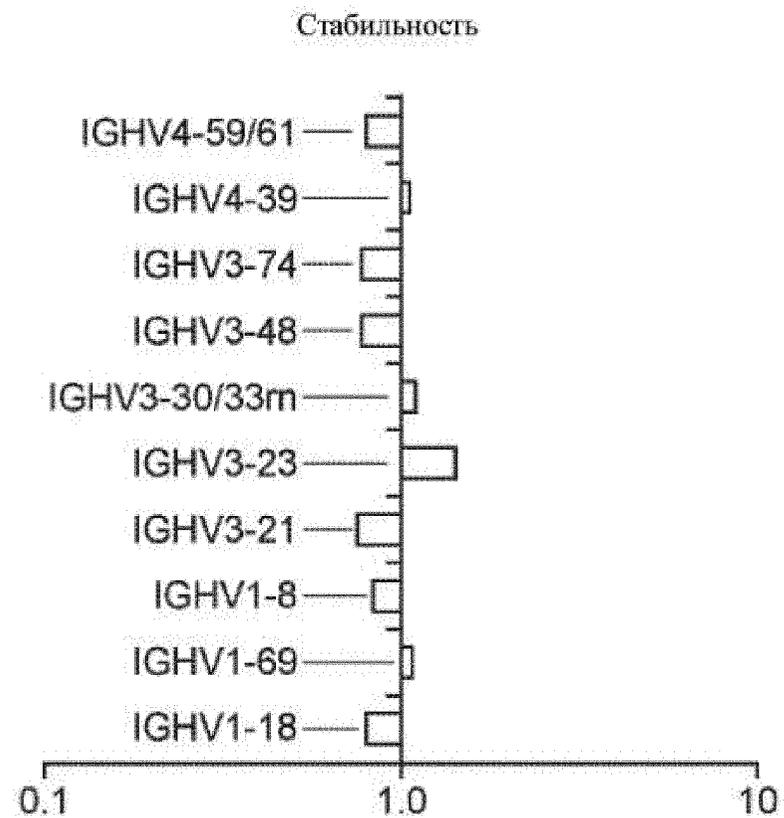
Экспонирование мотива (His)



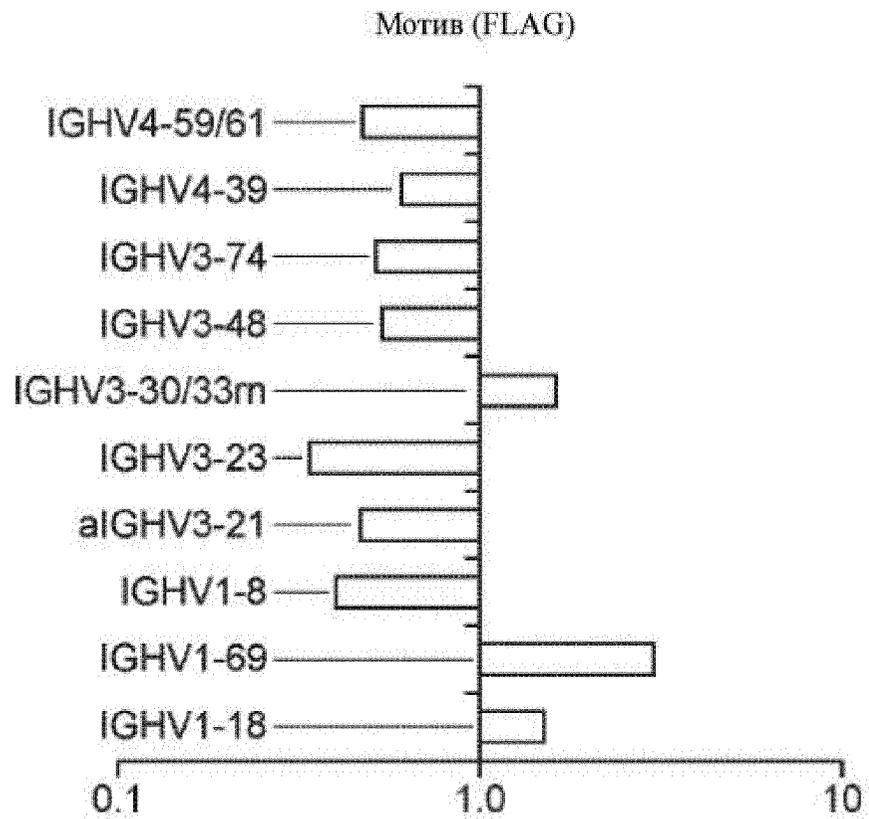
Фиг. 17D



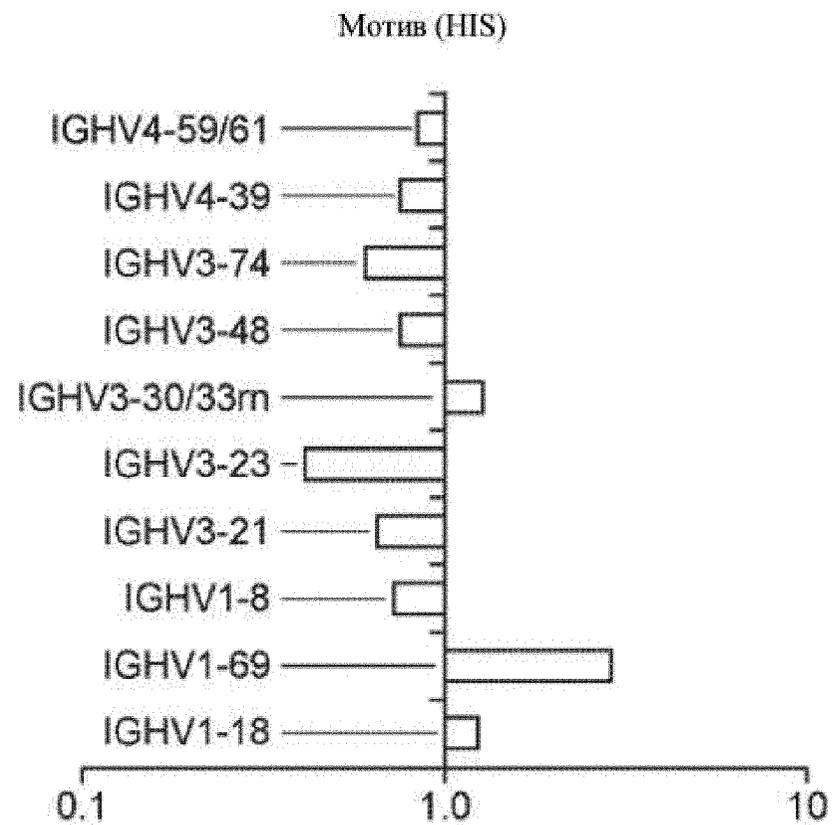
Фиг. 18А



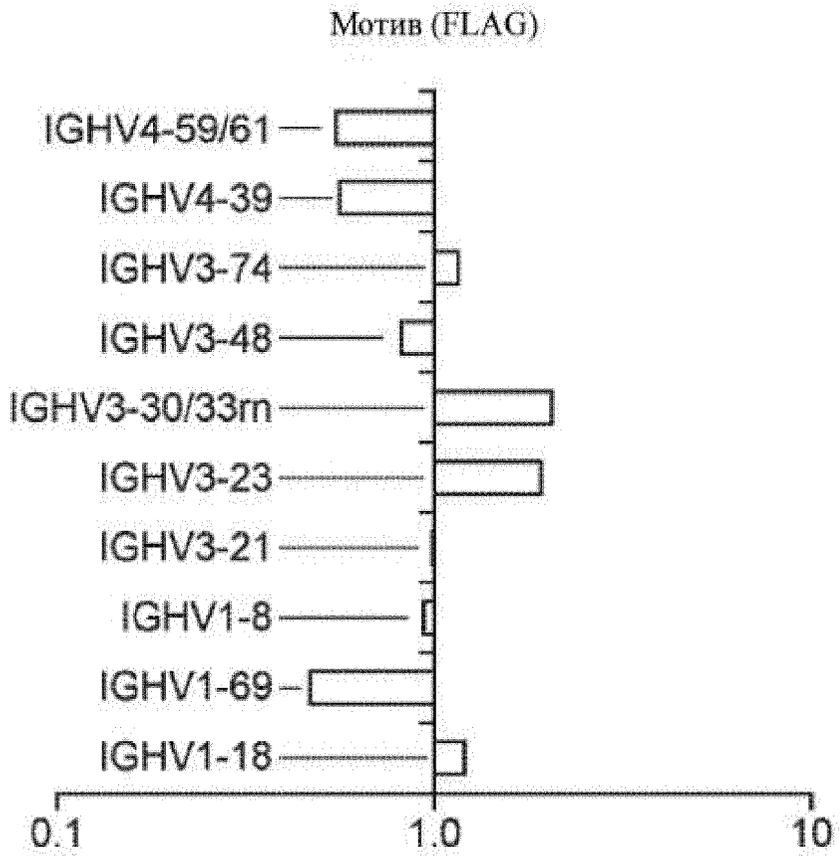
Фиг. 18В



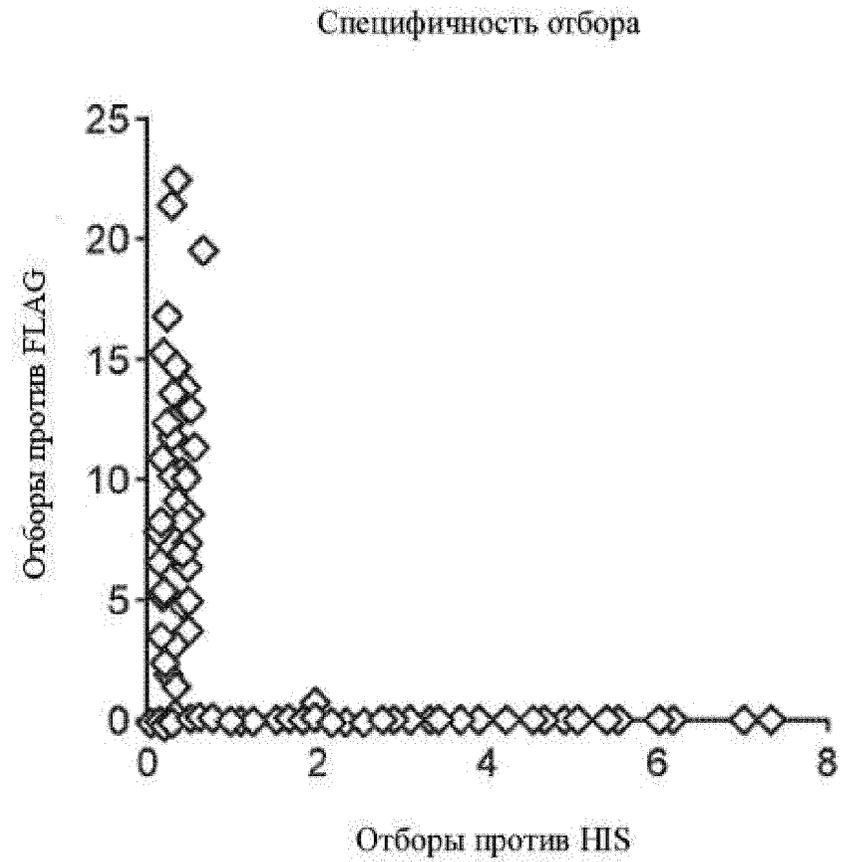
Фиг. 18С



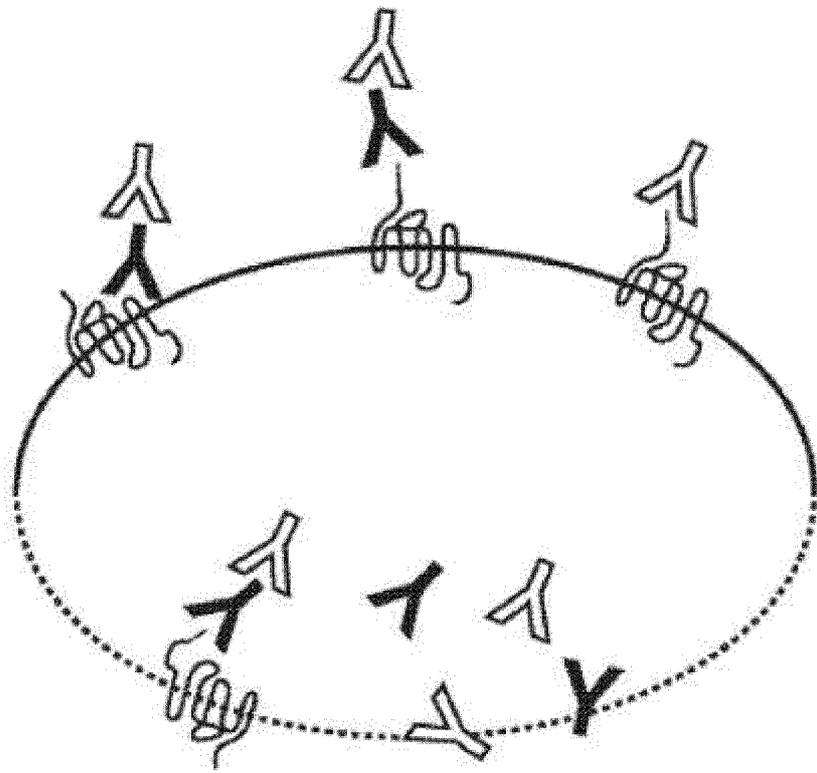
Фиг. 18D



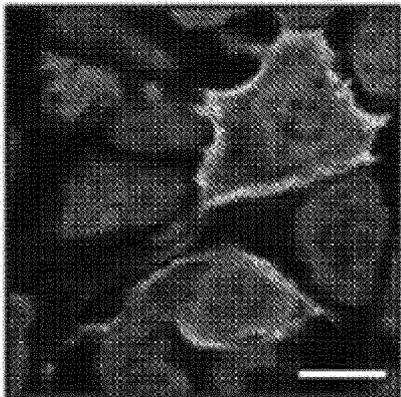
Фиг. 18E



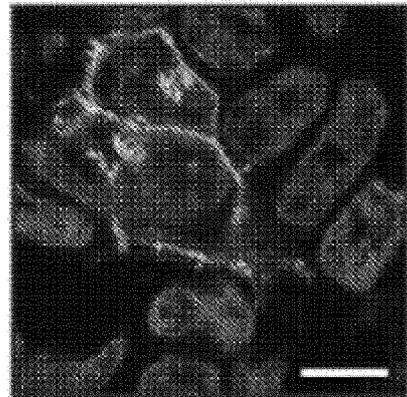
Фиг. 18F



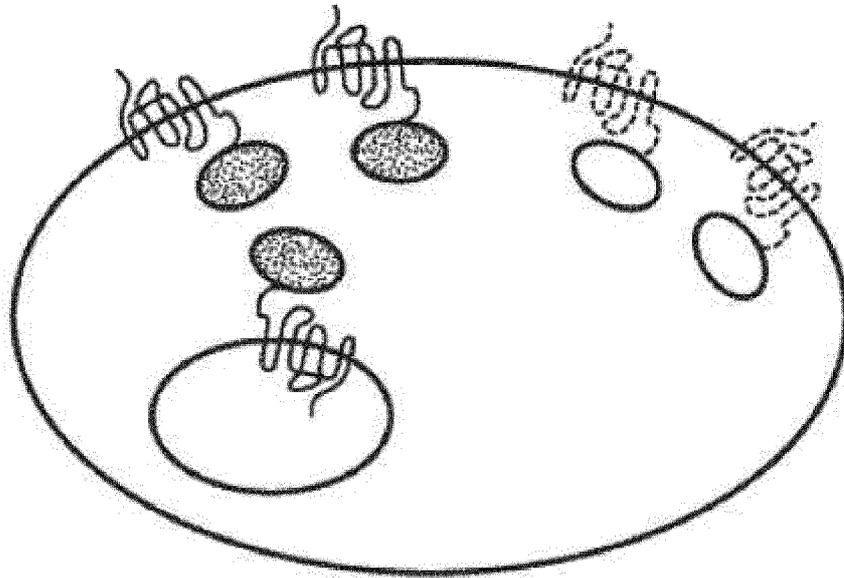
Фиг. 19А



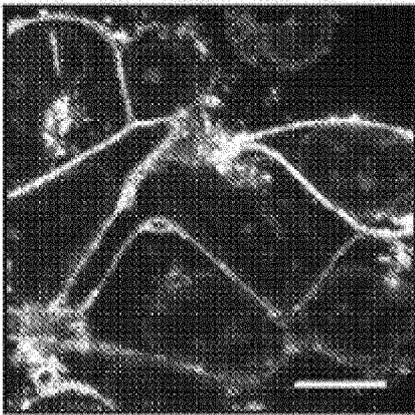
Фиг. 19В



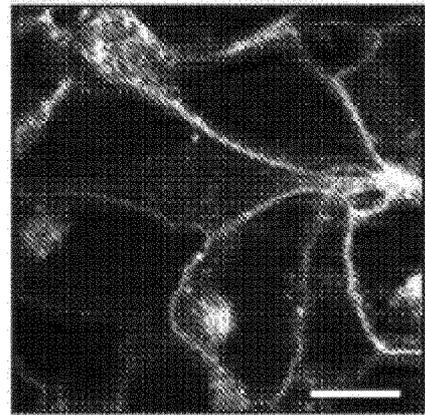
Фиг. 19С



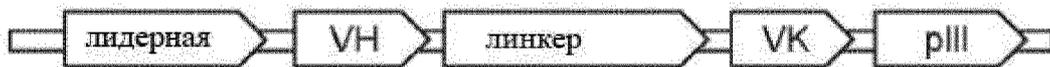
ФИГ. 20А

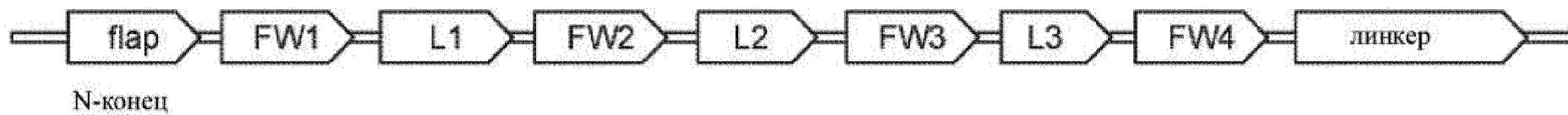


ФИГ. 20В



ФИГ. 20С

**Фиг. 21А****Фиг. 21В**



Фиг. 21С