

(19)



**Евразийское
патентное
ведомство**

(11) **048071**

(13) **B1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2024.10.23

(21) Номер заявки
201891619

(22) Дата подачи заявки
2017.01.10

(51) Int. Cl. **C12N 15/11** (2006.01)
C12N 15/90 (2006.01)
C12N 9/22 (2006.01)

(54) ХИМЕРНЫЕ БЕЛКИ И СПОСОБЫ РЕГУЛИРОВАНИЯ ЭКСПРЕССИИ ГЕНОВ

(31) 62/277,322; 62/351,522; 62/399,902

(32) 2016.01.11; 2016.06.17; 2016.09.26

(33) US

(43) 2019.02.28

(86) PCT/US2017/012885

(87) WO 2017/123559 2017.07.20

(71)(73) Заявитель и патентовладелец:

**ТЕ БОРД ОФ ТРАСТИЗ ОФ ТЕ
ЛИЛЭНД СТЭНФОРД ДЖУНИОР
ЮНИВЕРСИТИ (US)**

(72) Изобретатель:

Ци Лэй С., Дингал П. С. Дейв П. (US)

(74) Представитель:

Фелицына С.Б. (RU)

(56) US-A1-20150232881

WO-A2-2014018423

WO-A1-2015139139

US-A1-20040197346

WO-A1-2015150771

US-A1-20140068797

WO-A1-2014196932

Kopan et al., "The Canonical Notch Signaling Pathway: Unfolding The Activation Mechanism", Cell, 2009, Vol. 137, No.2, pp. 216-233; page 216, second column, first paragraph

UniProtKB Q01705 (NOTC1_09 December 2015; [retrieved 08 June 2017]. Retrieved from the Internet: <URL: <http://www.uniprot.org/uniprot/Q01705#sequences>>; pp. 1-29; page 21.

(57) В настоящем изобретении представлены системы, композиции и способы для регуляции экспрессии целевых полинуклеотидов в клетках. Системы, композиции и способы включают химерный рецепторный полипептид, химерный адаптерный полипептид, по меньшей мере один исполнительный элемент и расщепляющий элемент.

048071
B1

048071
B1

Перекрестные ссылки

Настоящая заявка претендует на приоритет от U.S. Provisional Application № 62/277322, поданной 11 января 2016 г., U.S. Provisional Application № 62/351522, поданной 17 июня 2016 г., и U.S. Provisional Application № 62/399902, поданной 26 сентября 2016 г., каждая из которых включена сюда во всей полноте путем ссылки.

Уровень техники

Регуляция клеточной активности может включать связывание лиганда с мембраносвязанным рецептором, содержащим внеклеточный лиганд-связывающий домен и внутриклеточный (например, цитоплазматический) домен сигнализации. Образование комплекса между лигандом и связывающим лиганд доменом может вызвать конформационную и/или химическую модификацию рецептора, что приведет к передаче сигнала внутрь клетки. В некоторых ситуациях цитоплазматическая часть рецептора подвергается фосфорилированию (например, транс- и/или автофосфорилированию), вызывая изменение его активности. Эти события могут быть сопряжены с вторичными мессенджерами и/или с привлечением кофакторных белков. В некоторых случаях изменения в цитоплазматической части приводят к связыванию с другими белками (например, кофакторными белками и/или другими рецепторами). Эти другие белки могут активироваться, а затем выполнять различные функции внутри клетки.

Системы условной экспрессии генов позволяют условно регулировать один или несколько целевых генов. Системы условной экспрессии генов типа индуцируемых лекарствами систем экспрессии генов позволяют активировать и/или деактивировать экспрессию гена в ответ на стимул типа присутствия лекарственного средства. Однако существующие системы могут быть ограниченными из-за неточного контроля, недостаточного уровня индукции (например, активации и/или деактивации экспрессии генов) и отсутствия специфичности.

Сущность изобретения

Ввиду вышесказанного, существует значительная потребность в альтернативных композициях и способах проведения условной регуляции экспрессии генов, к примеру, путем регуляции экспрессии целевого полинуклеотида. В одном аспекте настоящего изобретения предусмотрена система регуляции экспрессии целевого полинуклеотида в клетках. В некоторых воплощениях система включает: (a) химерный рецепторный полипептид, который подвергается модификации при связывании антигена, причем модификация рецептора включает конформационное изменение или химическую модификацию; (b) химерный адаптерный полипептид, который связывается с рецептором в ответ на модификацию рецептора; (c) полипептид, модулирующий ген (GMP), содержащий исполнительный элемент, соединенный с сайтом распознавания расщепления, причем при расщеплении сайта распознавания расщепления исполнительный элемент активируется и образует комплекс с целевым полинуклеотидом; и (d) расщепляющий элемент, который расщепляет сайт распознавания расщепления, когда находится вблизи сайта распознавания расщепления; причем: (i) GMP входит в состав внутриклеточной области химерного рецепторного полипептида, а расщепляющий элемент входит в состав химерного адаптерного полипептида; (ii) GMP входит в состав химерного адаптерного полипептида, а расщепляющий элемент входит в состав внутриклеточной области химерного рецепторного полипептида; или (iii) расщепляющий элемент образует комплекс со вторым адаптерным полипептидом, который связывает химерный рецепторный полипептид в ответ на модификацию рецептора, а GMP входит в состав химерного адаптерного полипептида. В некоторых воплощениях рецептор не содержит SEQ ID NO: 39.

В некоторых воплощениях целевой полинуклеотид представлен геномной ДНК. В некоторых воплощениях целевой полинуклеотид представлен РНК. В некоторых воплощениях модификация представляет собой фосфорилирование.

В некоторых воплощениях исполнительный элемент представлен Cas-белком, а система дополнительно содержит направляющую РНК, которая действует, образуя комплекс с Cas-белком. В некоторых воплощениях (i) исполнительный элемент представлен белком, связывающим РНК, (RBP), необязательно в комплексе с направляющей РНК, а (ii) система дополнительно содержит Cas-белок, который способен образовывать комплекс с направляющей РНК. В некоторых воплощениях Cas-белок практически лишен активности расщепления ДНК.

В некоторых воплощениях (i) GMP входит в состав химерного адаптерного полипептида, (ii) при расщеплении сайта распознавания расщепления химерный адаптерный полипептид эффективно высвобождается из рецептора, а (iii) система содержит дополнительный химерный адаптерный полипептид, содержащий GMP, который связывается с модифицированным рецептором.

В некоторых воплощениях модификация рецептора включает модификации по нескольким сайтам модификации, причем каждый сайт модификации эффективно связывает адаптерный полипептид.

В некоторых воплощениях сайт распознавания расщепления включает полипептидную последовательность, а расщепляющий элемент обладает протеазной активностью. В некоторых воплощениях сайт распознавания расщепления содержит дисульфидную связь, а расщепляющий элемент обладает оксидоредуктазной активностью. В некоторых воплощениях сайт распознавания расщепления содержит первую часть последовательности интеина, которая реагирует со второй частью последовательности интеина с высвобождением исполнительного элемента.

В некоторых воплощениях рецептор является трансмембранным рецептором. В некоторых воплощениях рецептор является ядерным рецептором.

В некоторых воплощениях исполнительный элемент регулирует экспрессию целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит активатор для эффективного повышения экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент соединяется по меньшей мере с одним сигналом ядерной локализации (NLS).

В некоторых воплощениях химерный рецепторный полипептид связан по меньшей мере с одной нацеливающей (targeting) последовательностью, которая направляет транспортировку рецептора в определенный район клетки. В некоторых воплощениях нацеливающая последовательность направляет транспорт рецептора в ядро, цитоплазму, митохондрии, эндоплазматический ретикулум (ER), хлоропласты, апопласты, пероксисомы или плазматическую мембрану. В некоторых воплощениях нацеливающая последовательность содержит сигнал ядерного экспорта (NES). В некоторых воплощениях нацеливающая последовательность включает нацеливающий на плазматическую мембрану пептид.

В некоторых воплощениях химерный адаптерный полипептид связан по меньшей мере с одной нацеливающей последовательностью, которая направляет транспортировку адаптера в определенный район клетки. В некоторых воплощениях нацеливающая последовательность направляет транспорт химерного адаптерного полипептида в ядро, цитоплазму, митохондрии, эндоплазматический ретикулум (ER), хлоропласты, апопласты, пероксисомы или плазматическую мембрану. В некоторых воплощениях нацеливающая последовательность содержит сигнал ядерного экспорта (NES). В некоторых воплощениях нацеливающая последовательность включает нацеливающий пептид для плазматической мембраны.

В некоторых воплощениях рецептор связан с доменом фолдинга полипептидов. В некоторых воплощениях с доменом фолдинга полипептидов связан химерный адаптерный полипептид.

В одном аспекте настоящего изобретения предусмотрен способ регулирования экспрессии целевого полинуклеотида в клетках. В некоторых воплощениях способ включает: (a) экспонирование химерного рецепторного полипептида с антигеном, причем (i) рецептор подвергается модификации при экспозиции с антигеном, а (ii) модификация рецептора включает конформационное изменение или химическую модификацию; (b) связывание химерного адаптерного полипептида с химерным рецепторным полипептидом в ответ на модификацию рецептора с образованием комплекса между ген-модулирующим полипептидом (GMP) и расщепляющим элементом, причем GMP содержит исполнительный элемент, соединенный с сайтом распознавания расщепления; и (c) расщепление сайта распознавания расщепления расщепляющим элементом, причем при расщеплении сайта распознавания расщепления исполнительный элемент образует комплекс с целевым полинуклеотидом, тем самым регулируя экспрессию целевого полинуклеотида в клетке; при этом: (i) GMP входит в состав внутриклеточной области химерного рецепторного полипептида, а расщепляющий элемент входит в состав химерного адаптерного полипептида; (ii) расщепляющий элемент входит в состав химерного адаптерного полипептида, а GMP входит в состав внутриклеточной области химерного рецептора; или (iii) расщепляющий элемент образует комплекс со вторым адаптерным полипептидом, который связывается с рецептором в ответ на модификацию рецептора, а GMP входит в состав химерного адаптерного полипептида. В некоторых воплощениях рецептор не содержит SEQ ID NO: 39.

В некоторых воплощениях целевой полинуклеотид представлен геномной ДНК. В некоторых воплощениях целевой полинуклеотид представлен РНК. В некоторых воплощениях модификация представляет собой фосфорилирование.

В некоторых воплощениях исполнительный элемент представлен Cas-белком, который образует комплекс с направляющей РНК. В некоторых воплощениях исполнительный элемент представлен белком, связывающим РНК, (RBP) в комплексе с направляющей РНК, которая образует комплекс с белком Cas. В некоторых воплощениях белок Cas практически лишен активности расщепления ДНК.

В некоторых воплощениях (i) GMP входит в состав химерного адаптерного полипептида, (ii) химерный адаптерный полипептид высвобождается из рецептора после расщепления сайта распознавания расщепления, а (iii) дополнительный химерный адаптерный полипептид, содержащий GMP, связывается с модифицированным рецептором.

В некоторых воплощениях модификация рецептора включает модификации по нескольким сайтам модификации, причем каждый сайт модификации эффективно связывает адаптерный полипептид.

В некоторых воплощениях сайт распознавания расщепления включает полипептидную последовательность, а расщепляющий элемент обладает протеазной активностью. В некоторых воплощениях сайт распознавания расщепления содержит дисульфидную связь, а расщепляющий элемент обладает оксидоредуктазной активностью. В некоторых воплощениях сайт распознавания расщепления содержит первую часть последовательности интеина, которая реагирует со второй частью последовательности интеина с высвобождением исполнительного элемента.

В некоторых воплощениях рецептор является трансмембранным рецептором. В некоторых воплощениях рецептор является ядерным рецептором.

В некоторых воплощениях исполнительный элемент регулирует экспрессию целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит активатор для эффективного повышения экспрессии целевого полинуклеотида.

В одном аспекте настоящего изобретения предусмотрен химерный внутриклеточный рецептор. В некоторых воплощениях рецептор включает: (a) взаимодействующий с антигеном домен, который специфически связывает антиген; и (b) исполнительный элемент, связанный со взаимодействующим с антигеном доменом; причем: (i) химерный внутриклеточный рецептор подвергается модификации в ответ на связывание антигена; (ii) химерный рецепторный полипептид транслоцируется в ядро клетки в ответ на модификацию; и (iii) исполнительный элемент образует комплекс с целевым полинуклеотидом в ядре.

В некоторых воплощениях исполнительный элемент представлен белком Cas, который образует комплекс с направляющей РНК. В некоторых воплощениях белок Cas практически лишен активности расщепления ДНК.

В некоторых воплощениях исполнительный элемент регулирует экспрессию целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит активатор для эффективного повышения экспрессии целевого полинуклеотида.

В некоторых воплощениях антиген представлен гормоном.

В некоторых воплощениях исполнительный элемент соединяется по меньшей мере с одним сигналом ядерной локализации (NLS).

В некоторых воплощениях рецептор связан по меньшей мере с одной нацеливающей последовательностью, которая направляет транспортировку рецептора в определенный район клетки. В некоторых воплощениях нацеливающая последовательность направляет транспорт рецептора в ядро, цитоплазму, митохондрии, эндоплазматический ретикулум (ER), хлоропласты, апопласты или пероксисомы. В некоторых воплощениях нацеливающая последовательность содержит сигнал ядерного экспорта (NES). В некоторых воплощениях нацеливающая последовательность включает нацеливающий пептид для плазматической мембраны. В некоторых воплощениях рецептор связан с доменом укладки полипептидов.

В одном аспекте настоящего изобретения предусмотрен способ регулирования экспрессии целевого полинуклеотида в клетках, содержащих ядро. В некоторых воплощениях способ включает: (a) экспонирование химерного внутриклеточного рецептора с антигеном, причем (i) рецептор содержит взаимодействующий с антигеном домен и исполнительный элемент, и (ii) рецептор подвергается модификации при экспозиции с антигеном; (b) транслокацию модифицированного рецептора в ядро; и (c) образование комплекса между исполнительным элементом и целевым полинуклеотидом.

В некоторых воплощениях исполнительный элемент представлен белком Cas, который образует комплекс с направляющей РНК. В некоторых воплощениях белок Cas практически лишен активности расщепления ДНК.

В некоторых воплощениях исполнительный элемент регулирует экспрессию целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит активатор для эффективного повышения экспрессии целевого полинуклеотида.

В некоторых воплощениях антиген представлен гормоном.

В одном аспекте настоящего изобретения предусмотрен химерный рецепторный полипептид. В некоторых воплощениях химерный рецепторный полипептид включает: (a) взаимодействующий с антигеном домен; и (b) полипептид, модулирующий ген (GMP), включающий исполнительный элемент, соединенный с сайтом распознавания расщепления; причем (i) химерный рецепторный полипептид подвергается модификации в ответ на связывание антигена; (ii) сайт распознавания расщепления расщепляется расщепляющим элементом в ответ на модификацию химерного рецепторного полипептида; (iii) исполнительный элемент образует комплекс с целевым полинуклеотидом после отщепления от химерного рецепторного полипептида по сайту распознавания расщепления; и (iv) химерный рецепторный полипептид не содержит SEQ ID NO: 39.

В некоторых воплощениях сайт распознавания расщепления фланкирован взаимодействующим с антигеном доменом и исполнительным элементом.

В некоторых воплощениях взаимодействующий с антигеном домен входит в состав внеклеточной области химерного рецепторного полипептида, а GMP входит в состав внутриклеточной области химерного рецепторного полипептида.

В некоторых воплощениях исполнительный фрагмент транслоцируется в ядро клетки после расщепления последовательности распознавания расщепления. В некоторых воплощениях химерный рецепторный полипептид представлен ядерным рецептором, который транслоцируется в ядро клетки в ответ на связывание антигена.

В некоторых воплощениях исполнительный элемент представлен белком Cas, который образует комплекс с направляющей РНК. В некоторых воплощениях исполнительный элемент представлен белком, связывающим РНК, (RBP), необязательно в комплексе с направляющей РНК, которая способна образовывать комплекс с белком Cas. В некоторых воплощениях белок Cas практически лишен активности расщепления ДНК.

В некоторых воплощениях сайт распознавания расщепления включает полипептидную последовательность, которая является распознавательной последовательностью протеазы. В некоторых воплощениях сайт распознавания расщепления содержит первую часть последовательности интеина, которая реагирует со второй частью последовательности интеина с высвобождением исполнительного элемента. В некоторых воплощениях сайт распознавания расщепления содержит дисульфидную связь.

В некоторых воплощениях рецептор является трансмембранным рецептором. В некоторых воплощениях рецептор является ядерным рецептором.

В некоторых воплощениях исполнительный элемент регулирует экспрессию целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит активатор для эффективного повышения экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент соединяется по меньшей мере с одним сигналом ядерной локализации (NLS).

В некоторых воплощениях рецептор связан по меньшей мере с одной нацеливающей последовательностью, которая направляет транспортировку рецептора в определенный район клетки. В некоторых воплощениях нацеливающая последовательность направляет транспорт рецептора в ядро, цитоплазму, митохондрии, эндоплазматический ретикулум (ER), хлоропласты, апопласты, пероксисомы или плазматическую мембрану. В некоторых воплощениях нацеливающая последовательность содержит сигнал ядерного экспорта (NES). В некоторых воплощениях нацеливающая последовательность включает нацеливающий пептид для плазматической мембраны. В некоторых воплощениях рецептор связан с доменом фолдинга полипептидов.

В одном аспекте настоящего изобретения предусмотрен химерный адаптерный полипептид. В некоторых воплощениях химерный адаптерный полипептид включает: (a) участок связывания рецептора, который связывает рецептор, подвергшийся модификации при связывании антигена; и (b) полипептид, модулирующий ген (GMP), соединенный с участком рецепторного связывания, причем: (i) сайт распознавания расщепления расщепляется расщепляющим элементом в ответ на связывание рецептора; а (ii) исполнительный элемент образует комплекс с целевым полинуклеотидом в ответ на расщепление сайта распознавания расщепления. В некоторых воплощениях исполнительный элемент способен транслоцироваться в ядро клетки после расщепления последовательности распознавания расщепления.

В некоторых воплощениях исполнительный элемент представлен белком Cas, который образует комплекс с направляющей РНК. В некоторых воплощениях исполнительный элемент представлен белком, связывающим РНК (RBP), необязательно в комплексе с направляющей РНК, которая способна образовывать комплекс с белком Cas. В некоторых воплощениях белок Cas практически лишен активности расщепления ДНК.

В некоторых воплощениях сайт распознавания расщепления включает полипептидную последовательность, которая является распознавательной последовательностью протеазы. В некоторых воплощениях сайт распознавания расщепления содержит первую часть последовательности интеина, которая реагирует со второй частью последовательности интеина с высвобождением исполнительного элемента. В некоторых воплощениях сайт распознавания расщепления содержит дисульфидную связь.

В некоторых воплощениях исполнительный элемент регулирует экспрессию целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит активатор для эффективного повышения экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент соединяется по меньшей мере с одним сигналом ядерной локализации (NLS).

В некоторых воплощениях адаптерный полипептид связан по меньшей мере с одной нацеливающей последовательностью, которая направляет транспортировку адаптера в определенный район клетки. В некоторых воплощениях нацеливающая последовательность направляет транспорт адаптера в ядро, цитоплазму, митохондрии, эндоплазматический ретикулум (ER), хлоропласты, апопласты, пероксисомы или плазматическую мембрану. В некоторых воплощениях нацеливающая последовательность содержит сигнал ядерного экспорта (NES). В некоторых воплощениях нацеливающая последовательность включает пептид, нацеливающий на плазматическую мембрану.

В одном аспекте настоящего изобретения предусмотрена система для регуляции экспрессии целевого полинуклеотида в клетках. В некоторых воплощениях система включает: (a) химерный рецепторный полипептид, который подвергается модификации при связывании антигена, причем модификация рецептора включает конформационное изменение или химическую модификацию; (b) химерный адаптерный полипептид, который связывается с рецептором в ответ на модификацию рецептора; (c) исполни-

тельный элемент, соединенный с доменом расщепления пептидов, причем при расщеплении домена расщепления пептидов исполнительный элемент активируется и образует комплекс с целевым полинуклеотидом; и (d) расщепляющий элемент, который расщепляет домен расщепления пептидов, когда находится вблизи домена расщепления пептидов; причем: (i) расщепляющий элемент входит в состав внутриклеточной части рецептора, а исполнительный элемент, связанный с доменом расщепления пептидов, входит в состав химерного адаптерного полипептида; (ii) расщепляющий элемент образует комплекс со вторым адаптерным полипептидом, который связывается с рецептором в ответ на модификацию рецептора, а исполнительный элемент, связанный с доменом расщепления пептидов, входит в состав химерного адаптерного полипептида; или же (iii) расщепляющий элемент входит в состав адаптерного полипептида, а исполнительный элемент, связанный с доменом расщепления пептидов, входит в состав внутриклеточной части рецептора.

В одном аспекте настоящего изобретения предусмотрена система для регуляции экспрессии целевого полинуклеотида в клетках. В некоторых воплощениях система включает: (a) химерный рецепторный полипептид, который подвергается модификации при связывании антигена, причем модификация рецептора включает конформационное изменение или химическую модификацию; (b) химерный адаптерный полипептид, который связывается с рецептором в ответ на модификацию рецептора; (c) исполнительный элемент, соединенный с доменом расщепления пептидов, причем при расщеплении домена расщепления пептидов исполнительный элемент активируется и образует комплекс с целевым полинуклеотидом; и (d) рекомбинантный протеазный домен, который расщепляет домен расщепления пептидов, когда находится вблизи домена расщепления пептидов; причем: (i) рекомбинантный протеазный домен входит в состав внутриклеточной части рецептора, а исполнительный элемент, связанный с доменом расщепления пептидов, входит в состав химерного адаптерного полипептида; (ii) рекомбинантный протеазный домен образует комплекс со вторым адаптерным полипептидом, который связывается с рецептором в ответ на модификацию рецептора, а исполнительный элемент, связанный с доменом расщепления пептидов, входит в состав химерного адаптерного полипептида; или же (iii) рекомбинантный протеазный домен входит в состав адаптерного полипептида, а исполнительный элемент, связанный с доменом расщепления пептидов, входит в состав внутриклеточной части рецептора.

В одном аспекте настоящего изобретения предусмотрен химерный рецепторный полипептид. Химерный рецепторный полипептид включает: внеклеточный взаимодействующий с антигеном домен, который связывает антиген; трансмембранный домен; и внутриклеточный домен, модулирующий ген, включающий Cas-белок, причем домен расщепления пептидов располагается на N-конце ген-модулирующего домена; причем при связывании внеклеточного взаимодействующего с антигеном домена с антигеном домен, модулирующий ген, высвобождается из химерного рецепторного полипептида путем расщепления домена расщепления пептидов. В некоторых воплощениях химерный рецепторный полипептид подвергается модификации при связывании рецептора с антигеном. В некоторых воплощениях трансмембранный домен содержит часть белка рецептора Notch либо его производное, вариант или фрагмент. В некоторых воплощениях трансмембранный домен имеет аминокислотную последовательность, которая по меньшей мере на 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 или 100% идентична SEQ ID NO: 39 или её фрагменту. В некоторых воплощениях белок Cas практически лишен активности расщепления ДНК. В некоторых воплощениях белок Cas представлен белком Cas9.

В некоторых воплощениях домен, модулирующий ген, дополнительно содержит активаторный домен для эффективного повышения экспрессии целевого полинуклеотида. В некоторых воплощениях домен, модулирующий ген, дополнительно содержит репрессорный домен для эффективного снижения экспрессии целевого полинуклеотида.

В некоторых воплощениях домен, модулирующий ген, дополнительно содержит по меньшей мере одну нацеливающую последовательность, которая направляет транспортировку домена, модулирующего ген, в определенный район клетки после высвобождения домена, модулирующего ген, из рецептора. В некоторых воплощениях по меньшей мере одна нацеливающая последовательность содержит сигнал ядерной локализации (NLS).

В некоторых воплощениях рецептор связан по меньшей мере с одной нацеливающей последовательностью, которая направляет транспортировку рецептора в определенный район клетки. В некоторых воплощениях рецептор связан с доменом укладки полипептидов.

В некоторых воплощениях домен расщепления пептидов включает распознавательную последовательность протеазы.

Включение путем ссылки

Все публикации, патенты и патентные заявки, приведенные в этом описании, включены сюда путем ссылки в такой же степени, как если бы каждая отдельная публикация, патент или патентная заявка была конкретно и индивидуально указана как включенная путем ссылки.

Краткое описание фигур

Привносящие новизну признаки изобретения изложены, в частности, в прилагаемой формуле изобретения. Для лучшего понимания признаков и преимуществ настоящего изобретения следует обратиться к нижеследующему подробному описанию, в котором представлены иллюстративные воплощения, в

которых применяются принципы изобретения, и к прилагаемым чертежам.

На фиг. 1 представлен типичный химерный рецепторный полипептид, содержащий взаимодействующий с антигеном домен и полипептид, модулирующий ген (GMP).

На фиг. 2 представлен типичный химерный полипептид трансмембранного рецептора.

На фиг. 3А представлен типичный химерный рецепторный полипептид, включающий исполнительный элемент, содержащий связывающий белок, связывающий РНК, необязательно в комплексе с направляющей нуклеиновой кислотой (например, sgRNA). На фиг. 3В представлена типичная система, включающая химерный рецепторный полипептид и химерный адаптерный полипептид, содержащий расщепляющий элемент.

На фиг. 4А-Д схематически представлено высвобождение исполнительного элемента из GMP в системе, включающей рецептор, который подвергается фосфорилированию. На фиг. 4Е-Н схематически представлено высвобождение исполнительного элемента из GMP в системе, включающей рецептор, который подвергается конформационному изменению.

На фиг. 5 представлен типичный химерный рецепторный полипептид, включающий по меньшей мере одну нацеливающую последовательность.

На фиг. 6А представлен типичный химерный адаптерный полипептид, включающий компонент рецепторного связывания и полипептид, модулирующий ген (GMP). На фиг. 6В представлен типичный химерный адаптерный полипептид, включающий исполнительный элемент, содержащий связывающий белок, связывающий РНК, необязательно в комплексе с направляющей нуклеиновой кислотой (например, sgRNA).

На фиг. 7 представлена типичная система, включающая химерный рецепторный полипептид, содержащий расщепляющий элемент, и химерный адаптерный полипептид, содержащий GMP.

На фиг. 8А-Д схематически представлено высвобождение исполнительного элемента из GMP в системе, включающей рецептор, который подвергается фосфорилированию. На фиг. 8Е-Н схематически представлено высвобождение исполнительного элемента из GMP в системе, включающей рецептор, который подвергается конформационному изменению.

На фиг. 9 представлена типичная система, включающая химерный рецепторный полипептид, химерный адаптерный полипептид, содержащий GMP, и второй адаптерный полипептид, содержащий расщепляющий элемент.

На фиг. 10А-Д схематически представлено высвобождение исполнительного элемента из GMP в системе, включающей по меньшей мере два адаптерных полипептида и рецептор, который подвергается фосфорилированию. На фиг. 10Е-Н схематически представлено высвобождение исполнительного элемента из GMP в системе, включающей по меньшей мере два адаптерных полипептида и рецептор, который подвергается конформационному изменению.

На фиг. 11 представлен типичный химерный адаптерный полипептид, включающий по меньшей мере одну нацеливающую последовательность.

На фиг. 12А-С схематически представлена система, включающая типичный внутриклеточный рецептор.

На фиг. 13А-Д схематически представлена система, в которой сайт распознавания расщепления содержит последовательность интеина. На фиг. 13Е-Н представлено альтернативное устройство системы, в которой сайт распознавания расщепления содержит последовательность интеина.

На фиг. 14А-Д схематически представлена система, в которой сайт распознавания расщепления содержит дисульфидную связь. На фиг. 14Е-Н представлено альтернативное устройство системы, в которой сайт распознавания расщепления содержит дисульфидную связь.

На фиг. 15 представлена иллюстрация, адаптированная из фиг. 2 в Makarova K.S. et al. "An updated evolutionary classification of CRISPR-Cas systems" *Nat Rev Microbiol* (2015) 13:722-736, в которой представлена архитектура геномных локусов для подтипов системы CRISPR-Cas.

На фиг. 16А-Д схематически представлено высвобождение исполнительного элемента из GMP в системе, включающей по меньшей мере два адаптерных полипептида.

На фиг. 17 схематически представлены генно-инженерные (преобразованные) химерные рецепторы антигенов по настоящему изобретению для модулирования генов типа редактирования генома и регуляции генов.

На фиг. 18 представлены варианты линкеров, расположенных между трансмембранным доменом и ген-модулирующим доменом преобразованных химерных рецепторов антигенов по настоящему изобретению.

На фиг. 19А-19F представлены преобразованные химерные рецепторы антигенов с доменами, модулирующими ген, и, в некоторых случаях, связанные с ними адаптеры-протеазы. На фиг. 19А представлены такие рекомбинантные рецепторы, которые связываются с поверхностными антигенами клетки. На фиг. 19В представлены рекомбинантные рецепторы для модулирования генов, которые могут связывать растворимые антигены. На фиг. 19С представлены модулирующие ген преобразованные рецепторы, которые могут связываться с сигнальными молекулами внеклеточного матрикса (ECM). На фиг. 19D представлены димеризующиеся рецепторы. Один из рецепторов содержит внеклеточный домен (ECD),

трансмембранный домен (TM), внутриклеточный домен (ICD), последовательность расщепления пептидов и модулирующий ген эффекторный домен. Другой рецептор димера содержит ECD, TM, ICD и протеазу. На фиг. 19E представлен другой пример димеризующихся рецепторов, которые могут модулировать экспрессию генов или редактировать гены. Один из димеризующихся рецепторов содержит ECD, TM, ICD, последовательность расщепления пептидов и ген-модулирующий эффекторный домен. Другой рецептор димера содержит ECD, TM и ICD, но без протеазы. Протеаза, которая расщепляет этот димеризующийся рецептор, может быть пришита к адаптерному белку, который ассоциируется с активированным димеризующимся рецептором. На фиг. 19F представлен пример олигомеризующегося рецептора, который включает преобразованные химерные рецепторы антигенов, слитые с доменами, модулирующими ген.

На фиг. 20A и 20B представлены различные химерные рецепторы антигенов и представлено связывание активаторного домена dCas9, наводимого на целевой ген при помощи sgRNA. На фиг. 20A представлены рекомбинантные полипептиды химерных рецепторов антигенов и, в некоторых случаях, связанные с ними полипептиды адаптеров-протеаз типа Notch и протеазы пресенилина, GPCR и протеазы β_2 -аррестина, интегринов и протеаз паксиллина, кадгеринов и протеаз β -катенина, рецепторов смерти и протеаз FADD, а также химерные рецепторы антигенов.

На фиг. 21A представлено применение химерных антигенов GPCR, конъюгированных с активаторами dCas9. На фиг. 20B представлена альтернативная конфигурация, в которой протеазная часть конъюгирована с GPCR, а активаторный домен dCas9 конъюгирован с адаптерным белком, который привлекается к активированному GPCR.

На фиг. 22A и 22B представлены модулирующие ген полипептиды интегрин-dCas9 и их реакции на лиганды интегрин типа фибронектина. На фиг. 22A представлена схема активатора химерного рецептора антигенов типа интегрин-dCas9. На фиг. 22B представлен комплекс интегрин-dCas9, который реагирует на фибронектин. При связывании sgRNA, специфичной к репортерному гену, рекомбинантный комплекс активирует транскрипцию репортера (H2B-GFP). На фиг. 22C представлено действие комплекса интегрин-dCas9 в адгезированных клетках в сравнении с клетками в суспензии. На фиг. 22D и 22E представлена специфичность связывания паксиллина-TEV для β -субъединицы интегрин в сравнении с α -субъединицей.

На фиг. 23A и 23B представлено типичное воплощение химерного полипептида GPCR с ген-модулирующим доменом. На фиг. 23A представлена схема регуляции целевого гена при помощи активаторов химерных рецепторов антигенов на основе GPCR с dCas9. На фиг. 23B показано, что полипептид CXCR4-dCas9 реагирует на CXCL12 и активирует люминесцентный репортер. На фиг. 23C и 23D представлены химерные рецепторы GPCR, содержащие LPAR1, CXCR4 и hM3D, и соответствующая активация флуоресцентного репортера (GFP) в присутствии лиганда, протеазы β -аррестина и sgRNA. На фиг. 23E и 23F представлено сравнение уровня регуляции транскрипции гена-репортера в результате наводки sgRNA на dCas9-VPR после высвобождения из химерного рецептора и TetR-VPR, который связывается непосредственно с промотором гена-репортера.

На фиг. 24A-24G представлены типичные воплощения модульных искусственных химерных рецепторов Notch по настоящему изобретению. На фиг. 24A представлен Notch дикого типа, связанный со своим лигандом Delta. После того как рецептор активируется при связывании Delta, ICD отщепляется протеазой и транслоцируется в ядро для регуляции целевых генов. На фиг. 24B представлен химерный искусственный рецептор Notch, у которого ICD Notch был заменен на слитый с dCas9 белок. Слитый белок dCas9 может содержать эффекторный домен типа домена-активатора, например, домена VP64, или домена-репрессора, например, домена KRAB. На фиг. 24C представлен другой химерный искусственный рецептор Notch, содержащий слитый с dCas9 белок, у которого ECD Notch был заменен на scFv, связывающий CD47. На фиг. 24D представлен типичный модульный искусственный химерный рецептор Notch и слитый белок адаптер-протеаза (пресенилин-протеаза TEV), экспрессируемые на поверхности клеток типа иммунных клеток. Искусственный модульный химерный рецептор Notch может содержать слитый полипептид dCas9, линкер и эффекторный домен. При связывании Delta с Notch пресенилин может связываться с химерным искусственным рецептором Notch. Затем протеаза TEV может расщеплять домен расщепления пептидов химерного искусственного рецептора Notch. Активация флуоресцентного репортера в клетках, экспрессирующих Notch с активатором dCas9, представлена для клеток HEK293 на фиг. 24E, для клеток Jurkat - на фиг. 24F, и в макрофагах THP-1 - на фиг. 24G.

На фиг. 25A-25C представлены теоретические модели перестройки эндогенной реакции рецептора Notch с использованием описанных здесь химерных антигеновых рецепторов Notch. На фиг. 25A представлена репрессия эндогенной фагоцитарной реакции у принимающей клетки, экспрессирующей эндогенный Notch, при связывании с Delta, экспрессированным на сигнализирующей клетке. На фиг. 25B показано, что можно создать такой генно-инженерный caN-рецептор, у которого эндогенная фагоцитарная реакция будет перенастроена на внешний сигнал Delta. На фиг. 25C показано, что можно получить такой генно-инженерный caN-рецептор, у которого репрессия эндогенной фагоцитарной реакции клетки будет сменяться активацией при связывании Delta.

На фиг. 26A-26D показано, что Notch с активатором dCas9 при связывании Delta может активировать целевые гены типа генов, контролирующих апоптоз клеток или клеточный цикл. В клетках, экспрессирующих полипептид Notch-dCas9, целевые гены активируются в присутствии Delta (фиг. 26B и 26D). Химерный антигеновый рецептор Notch не активирует транскрипцию целевых генов в отсутствие Delta (фиг. 26A и 26C).

На фиг. 27A и 27B показано, что Notch с активатором dCas9 направляется sgUAS (SEQ ID NO: 1; gtactccgacctctagtgt) на промотор UAS и активирует транскрипцию гена-репортера (H2B-цитрина). На фиг. 27A представлена схема этого процесса. На фиг. 27B показано, что Notch с активатором dCas9 реагирует на Delta.

На фиг. 28A и 28B показано, что полипептид CXCR4-dCas9-VPR реагирует на лиганд CXCL12 и активирует транскрипцию гена-репортера (люциферазы). На фиг. 28A схематически представлено связывание лиганда с полипептидом CXCR4-dCas9-VPR, который образует комплекс с sgRNA (sgTET; SEQ ID NO:2; gtactgtctctatcactgata). На фиг. 28A также показан слитый белок β_2 -аррестин-протеаза, который может связываться со сконструированным химерным антигеновым рецептором. На схеме также показана (1) транслокация свободного dCas9-VPR в ядро, (2) связывание комплекса sgTET-dCas9-VPR с промотором TetO, который регулирует транскрипцию гена люциферазы, и (3) транскрипция репортера. На фиг. 28B показано, что транскрипция гена люциферазы регулируется связыванием CXCL2 с полипептидом CXCR4-dCas9-VPR.

На фиг. 29A и 29B показано, что полипептид интегрин-dCas9-VPR реагирует на лиганд внеклеточного матрикса и активирует транскрипцию репортерного гена (H2B-цитрина). На фиг. 29A схематически представлена активация транскрипции репортера при связывании лиганда с полипептидом интегрин-dCas9-VPR. На фиг. 29B показано, что комплекс сконструированного химерного антигенового рецептора на основе интегрина с dCas9 индуцирует экспрессию репортера в ответ на сигналы от ECM.

На фиг. 30A и 30B представлены типичные воплощения описанных здесь полипептидов типа химерный антигеновый рецептор-эффектор, которые основываются на логической схеме "расщепленного AND" или "каскадного AND". На фиг. 30A схематически представлен отщепленный эффектор dCas9, пришитый к отдельным преобразованным рецепторам. На фиг. 30B схематически представлен полипептид типа химерный рецептор-tTa, который при связывании со своим лигандом индуцирует управляемую TetO экспрессию полипептида химерный рецептор-dCas9.

На фиг. 31A-31I представлены воплощения стратегии на основе рецепторов для мобилизации Cas9 в ответ на внеклеточные сигналы. На фиг. 31A представлена активация рецепторов Notch1, включающая отщепление и транслокацию в ядро внутриклеточного домена Notch, который может быть заменен или подвергнут инженерии для усиления экспрессии производных Cas9. Слияние эффекторных доменов с Cas9 и заданные пользователями последовательности одинарной направляющей РНК (sgRNA) позволяют проводить целенаправленную регуляцию генов. На фиг. 31B представлены схемы конструкций химерных рецепторов с меткой mCherry, которые вначале тестировали на клеточную локализацию и Delta-зависимую активацию репортера. Оптимизированный по кодонам человека Cas9 с "мертвой" нуклеазой (dCas9) и состоящий из трех доменов (VP64, р65 и Rta) активатор VPR вставляли сразу после внеклеточного домена Notch1 (hNECD) и трансмембранного домена. Конструкция NC5 включает сигналы созревания из известного сигнала экспорта в ER. На фиг. 31C представлены схемы встроенных в клетки линии яичников китайского хомячка (CHO) вышележащих активирующих последовательностей (UAS) или связывающих CSL (не показано) промоторов, управляющих геном репортера гистона 2B (H2B)-цитрина, и наводящие на стабильно встроенные промоторы sgRNA (например, sgUAS или sasgCSL, соответственно), используемые для проверки эффективности активации генов у химерных рецепторов при культивировании с иммобилизованным на поверхности Delta или без него. Активация рецепторов NC5 иммобилизованными лигандами Delta ведет к отщеплению и транслокации в ядро dCas9-VPR. dCas9-VPR в комплексе со специфичной к последовательности sgRNA (например, sgUAS или sasgCSL) способствует связыванию комплекса с промотором и активации гена H2B-цитрина. На фиг. 31D представлены примеры снимков микроскопии клеток CHO, трансфицированных химерой NC5, что приводит к экспрессии H2B при обработке иммобилизованным Delta в течение 4 дней. Отметка масштаба - 20 мкм. На фиг. 31E представлены относительные уровни H2B (нормированная экспрессия) в клонах CHO UAS-H2B, отобранных по устойчивой экспрессии NC5 (dCas9 из *S.pyogenes* и sgUAS), или в клонах CHO CSL-H2B с Notch1 или NC5 дикого типа человека (dCas9 из *S.aureus* и sasgCSL) при культивировании на голой поверхности или с иммобилизованным Delta в течение 4 дней (n=3). Среднее \pm SEM. На фиг. 31F представлены проценты клеток с активированным H2B-цитрином в клонах CHO UAS-H2B, отобранных по устойчивой экспрессии NC5 (dCas9 из *S.pyogenes* и sgUAS), или в клонах CHO 12xCSL-H2B с Notch1 или NC5 дикого типа человека (dCas9 из *S.aureus* и sasgCSL) при культивировании на голой поверхности или с иммобилизованным Delta в течение 4 дней (n=3). Среднее \pm SEM, **p<0.01 по сравнению с контролями без Delta. На фиг. 31G представлена активация рецепторов NC5 иммобилизованным лигандом Delta, приводящая к отщеплению и транслокации в ядро dCas9-VPR. На фиг. 31H представлены гистограммы интенсивности репортера EGFP в репортерных клетках HEK293T, стабильно экспрессирующих tet-индуцибельный ген

EGFP и наводящую sgRNA (sgTET), в присутствии dCas9-VPR, NC5 + Delta + DAPT, NC5 + Delta, NC5 - Delta или без конструкции. На фиг. 31I представлены контурные графики (на которых между каждой парой контурных линий приходится одно и то же количество клеток) активации EGFP в репортерных клетках HEK293T, трансфицированных рецептором NC5 и культивируемых при различных концентрациях иммобилизованного Delta в течение 3 дней.

На фиг. 32A схематически представлена система типа синтетический Cas9-рецептор в соответствии с одним из описанных здесь воплощений. Модуляция целевых эндогенных генов реагирует на внеклеточные сигналы типа Delta. В зависимости от конструкции системы, нативные внеклеточные сигналы могут запускать различные варианты исходящего поведения клеток. На фиг. 32B слева представлена активация CDKN1B под действием hNECD-dCas9-VPR (конструкция NC5), которая вызывает Delta-зависимую остановку клеточного цикла в фазе G0/G1. Справа схематически представлено Delta-зависимое отщепление dCas9-VPR от NC5, которое ведет к появлению клеток с высоким содержанием CDKN1B (CDKN1B^{higt}) и с остановкой в фазе G0/G1 (2n DNA), в соответствии с одним воплощением. На фиг. 32C представлены примеры графиков проточной цитометрии и количественное содержание (заштриховано) клеток CDKN1B^{higt} и с остановкой в фазе G0/G1 при различных условиях культивирования: клетки HEK293 только с sgRNA, sgRNA + Delta, sgRNA + NC5 или sgRNA + NC5 + Delta. NC5 вводили путем краткосрочной трансфекции клеток со встроенной sgCDKN1B, которые затем культивировали на покрытой Delta или голой поверхности в течение 4 дней. Среднее \pm SEM, n=3. Клетки с остановкой в фазе G0/G1 представлены в левом верхнем заштрихованном углу графика и соответствуют верхней планке в каждой паре на гистограмме. Клетки с остановкой в фазе G2 представлены в левом верхнем заштрихованном углу графика и соответствуют нижней планке в каждой паре на гистограмме.

На фиг. 33 представлен пример снимка конфокальной флуоресцентной микроскопии клетки, трансфицированной конструкцией NC1 (красная) и sgUAS (синяя). Наблюдается преждевременная активация H2B (зеленая, ядро). На схеме справа представлен баланс между силой сигналов ядерной локализации и сигналов мембранного созревания при Delta-зависимом расщеплении химер Notch-Cas9.

На фиг. 34 представлены снимки флуоресцентной микроскопии клеток CHO, трансфицированных конструкциями NC1-NC4 (сверху, см. фиг. 31B) и проявляющих экспрессию H2B (внизу) в отсутствие Delta. Отметка масштаба - 20 мкм.

На фиг. 35A представлена природная последовательность NLS в Cas9 из *S.pyogenes*, которая приходится на аминокислотные остатки 647-670. Из иллюстрации видно, что этот "истинный" NLS (iNLS) может образовывать выходящую на поверхность структуру типа спираль-линкер-спираль (PDB ID: 4UN3). На фиг. 35A показана SEQ ID NO: 31. На фиг. 35B представлены типичные результаты для клеток HEK293 со стабильно встроенным репортером рTET-EGFP и нацеливающей sgRNA (sgTET), которые трансфицировали dCas9-VPR (сверху) или dCas9-VPR с мутантным iNLS (в центре). Представлены репрезентативные графики активации EGFP в присутствии или в отсутствие (внизу) конструкций и количественные гистограммы активации EGFP (нормирована по нетрансфицированным клеткам). На фиг. 35B раскрыты SEQ ID NO: 31-32 соответственно порядку их представления. На фиг. 35C показано по микроскопии нарушение ядерной локализации и активации репортера EGFP в клетках HEK293T при разрушении мотива iNLS. На фиг. 35C раскрыты SEQ ID NO: 31, 32, и 17 соответственно порядку их представления. На фиг. 35C и 35D методами оптической микроскопии и количественного определения показано, что при добавлении синтетического NLS к N-концу dCas9-VPR с мутантным iNLS частично восстанавливается активация EGFP.

На фиг. 36A-36E представлено Delta-зависимое редактирование гена. На фиг. 36A схематически представлено Delta-зависимое разрезание ДНК с помощью варианта NC5: hNECD слитый с Cas9 дикого типа *S.pyogenes* с активной нуклеазой (hNECD-Cas9). На фиг. 36B показана эффективность редактирования/разрезания гена с помощью hNECD-Cas9 в клетках CHO со стабильно встроенным геном EGFP и нацеливающей sgRNA (sgEGFP). На фиг. 36C представлена типичная схема для двух sgRNAs (короткие прямоугольники слева и справа от "sgCXCR4"), которые стабильно экспрессируются в клетках HEK293T и служат для наведения на 5'-нетранслируемый участок (UTR) и интрон-1 у CXCR4. Отметка масштаба - 1000 п.о. На фиг. 36D сверху представлены типичные результаты по эндонуклеазе T7E1 для анализа степени индуцируемой Delta опосредованной hNECD-Cas9 модификации гена CXCR4 в клетках HEK293T при детектировании по количеству расщепленных T7E1 продуктов в гелях SDS-PAGE. Внизу представлены типичные результаты по частоте мутаций indel у CXCR4. На фиг. 36E представлены типичные результаты по количественному определению экспрессии белка CXCR4 в клетках HEK293T методом иммунофлуоресцентного окрашивания на основе проточной цитометрии.

На фиг. 37A представлено совмещение повторов EGF-11 и -12 у человека, шпорцевой лягушки *Xenopus*, полосатого данио и дрозофилы (SEQ ID NO: 57-60). Идентичные остатки заключены в серые прямоугольники. Звездочкой (*) отмечены консервативные остатки цистеина и консенсусные связывающие Ca²⁺ остатки. На фиг. 37B представлены варибельные уровни активации у различных вариантов с делециями EGF (при оценке по анализу репортера).

На фиг. 38 представлена одновременная активация генов CXCR4 и CD95 при наводке на них геноспецифичных sgRNAs (2 на 1 ген). Приведены относительные уровни активации CXCR4 и CD95 при ука-

занных условиях в 4-дневных культурах (а.у., условные единицы). Данные представлены в виде средней интенсивности флуоресценции \pm SEM (n=3 независимых эксперимента). #, ##: p<0,01, 0,001; *, **: p<0,01, 0,001 по сравнению с уровнями CXCR4 и CD95, соответственно, у отрицательного контроля dCas9-VPR + sgNT (без наводки).

На фиг. 39A-39F показано преобразование сигналов Delta в фазоспецифичную остановку клеточного цикла. На фиг. 39A схематически представлено применение минимального варианта рецептора NC5 для запуска Delta-зависимой остановки клеточного цикла. На фиг. 39B показана остановка клеток в фазе G0/G1 в результате гиперэкспрессии CDKN1B. На фиг. 39C и 39D показано, что в клетках с dCas9-VPR и sgCDKN1B повышающая регуляция CDKN1B сопровождается наращиванием G0/G1, а в клетках с dCas9-VPR и нецелевой gsRNA наблюдается минимальное повышение CDKN1B. На фиг. 39E и 39F показано устранение Delta-индуцированной повышающей регуляции CDKN1B и остановки в фазе G0/G1 у клеток с DAPT.

Раскрытие сущности изобретения

При практическом применении некоторых изложенных здесь способов, если не указано иначе, применяются стандартные методы иммунологии, биохимии, химии, молекулярной биологии, микробиологии, клеточной биологии, геномики и рекомбинантной ДНК, которые входят в компетенцию специалистов в данной области. Например, см. Sambrook and Green, *Molecular Cloning: A Laboratory Manual*, 4th Edition (2012); серия *Current Protocols in Molecular Biology* (F.M. Ausubel et al., eds.); серия *Methods in Enzymology* (Academic Press, Inc.), PCR 2: A Practical Approach (M.J. MacPherson, B.D. Hames and G.R. Taylor, eds. (1995)); Harlow and Lane, eds. (1988) *Antibodies, A Laboratory Manual*; и *Culture of Animal Cells: A Manual of Basic Technique and Specialized Applications*, 6th Edition (R.I. Freshney, ed. (2010)).

В настоящем описании и формуле изобретения формы единственного числа включают и значения множественного числа, если контекстом четко не указано иначе. Например, термин "химерный трансмембранный рецептор" включает в себя и множество химерных трансмембранных рецепторов.

Термин "примерно" или "приблизительно" означает, что конкретное значение при определении рядовым специалистом в данной области находится в допустимом диапазоне погрешностей, что отчасти будет зависеть от того, как измеряется или определяется это значение, то есть от ограничений системы измерения. Например, "примерно" может означать в пределах 1 или более 1 стандартного отклонения, согласно практике в данной области. С другой стороны, "примерно" может означать диапазон до 20%, до 10%, до 5% или до 1% от заданного значения. Альтернативно, особенно в отношении биологических систем или процессов, этот термин может означать порядок величины, предпочтительно в пределах 5 раз, более предпочтительно в пределах 2 раз от значения. Когда в заявке и формуле изобретения приводятся конкретные значения, то следует принимать, если не указано иначе, что термин "примерно" означает в пределах допустимого диапазона погрешностей для конкретного значения.

В настоящем изобретении "клетка" обычно относится к биологическим клеткам. Клетка может быть основной структурной, функциональной и/или биологической единицей живого организма. Клетка может происходить из любого организма, содержащего одну или несколько клеток. Некоторые неограничительные примеры включают: прокариотические клетки, эукариотические клетки, бактериальные клетки, археальные клетки, клетки одноклеточных эукариотических организмов, клетки простейших, клетки растений (например, клетки сельскохозяйственных культур, плодов, овощей, зерновых, соевых бобов, кукурузы, пшеницы, семян, помидоров, риса, маниоки, сахарного тростника, тыквы, сена, картофеля, хлопка, конопли, табака, цветущих растений, хвойных деревьев, голосеменных, папоротников, плаунов, роголистников, печеночников, мхов), клетки водорослей (например, *Botryococcus braunii*, *Chlamydomonas reinhardtii*, *Nannochloropsis gaditana*, *Chlorella pyrenoidosa*, *Sargassum patens* C. Agardh и др.), морских водорослей (например, ламинарии), клетки грибов (например, дрожжевые клетки, клетки шампиньонов), клетки животных, клетки беспозвоночных животных (например, плодовой мушки, кишечнополостных, нематод и др.), клетки позвоночных животных (например, рыб, амфибий, рептилий, птиц, млекопитающих), клетки млекопитающих (например, свиней, коров, коз, овец, грызунов, крыс, мышей, приматов, человека и др.) и т.д. Иногда клетки происходят не из природного организма (например, клетки могут быть получены синтетически, их иногда называют искусственными клетками).

Термин "нуклеотид" в настоящем изобретении в общем относится к комбинации основания-сахарофосфат. К нуклеотидам относятся и синтетические нуклеотиды. К нуклеотидам относятся и синтетические аналоги нуклеотидов. Нуклеотиды могут быть мономерными звеньями последовательности нуклеиновой кислоты (например, дезоксирибонуклеиновой кислоты (ДНК) или рибонуклеиновой кислоты (РНК)). Термин нуклеотид охватывает рибонуклеозидтрифосфаты аденозинтрифосфат (АТФ), уридинтрифосфат (УТФ), цитозинтрифосфат (ЦТФ), гуанозинтрифосфат (ГТФ) и дезоксирибонуклеозидтрифосфаты типа dATP, dCTP, dITP, dUTP, dGTP, dTTP и их производные. К таким производным относятся, к примеру, $[\alpha S]dATP$, 7-деза-dGTP и 7-деза-dATP, и такие производные нуклеотидов, которые придают содержащим их молекулам нуклеиновой кислоты устойчивость к нуклеазам. Термин нуклеотид в настоящем изобретении может относиться и к дидезоксирибонуклеозидтрифосфатам (ddNTPs) и их производным. Типичные примеры дидезоксирибонуклеозидтрифосфатов включают, без ограничения, ddATP,

ddCTP, ddGTP, ddITP и ddTTP. Нуклеотид может быть немеченым или помечен детектируемой меткой хорошо известными методами. Мечение также может проводиться с помощью квантовых точек. К детектируемым меткам относятся, к примеру, радиоактивные изотопы, флуоресцентные метки, хемилюминесцентные метки, биолюминесцентные метки и ферментные метки. К флуоресцентным меткам нуклеотидов относятся, без ограничения, флуоресцеин, 5-карбоксихлорофлуоресцеин (FAM), 2',7'-диметокси-4',5-дихлор-6-карбоксихлорофлуоресцеин (JOE), родамин, 6-карбоксихлородамин (R6G), N,N,N',N'-тетраметил-6-карбоксихлородамин (TAMRA), 6-карбоксихлоро-родамин (ROX), 4-(4'-диметиламинофенилазо)бензойная кислота (DABCYL), Cascade Blue, Oregon Green, Texas Red, цианин и 5-(2'-аминоэтил)аминонафталин-1-сульфоновая кислота (EDANS). Конкретные примеры флуоресцентно меченых нуклеотидов включают [R6G]dUTP, [TAMRA]dUTP, [R110]dCTP, [R6G]dCTP, [TAMRA]dCTP, [JOE]ddATP, [R6G]ddATP, [FAM]ddCTP, [R110]ddCTP, [TAMRA]ddGTP, [ROX]ddTTP, [dR6G]ddATP, [dR110]ddCTP, [dTAMRA]ddGTP и [dROX]ddTTP фирмы Perkin Elmer, Foster City, Calif.; дезоксинуклеотиды FluoroLink, FluoroLink Cy3-dCTP, FluoroLink Cy5-dCTP, FluoroLink Fluor X-dCTP, FluoroLink Cy3-dUTP и FluoroLink Cy5-dUTP фирмы Amersham, Arlington Heights, Ill.; флуоресцеин-15-dATP, флуоресцеин-12-dUTP, тетраметилродамин-6-dUTP, IR770-9-dATP, флуоресцеин-12-ddUTP, флуоресцеин-12-UTP и флуоресцеин-15-2'-dATP фирмы Boehringer Mannheim, Indianapolis, Ind.; и меченые нуклеотиды хромосом, BODIPY-FL-14-UTP, BODIPY-FL-4-UTP, BODIPY-TMR-14-UTP, BODIPY-TMR-14-dUTP, BODIPY-TR-14-UTP, BODIPY-TR-14-dUTP, Cascade Blue-7-UTP, Cascade Blue-7-dUTP, флуоресцеин-12-UTP, флуоресцеин-12-dUTP, Oregon Green 488-5-dUTP, Rhodamine Green-5-UTP, Rhodamine Green-5-dUTP, тетраметилродамин-6-UTP, тетраметилродамин-6-dUTP, Texas Red-5-UTP, Texas Red-5-dUTP и Texas Red-12-dUTP фирмы Molecular Probes, Eugene, Oreg. Нуклеотиды также могут быть помечены или маркированы путем химической модификации. Химически модифицированными нуклеотидами являются биотин-dNTP. Некоторые неограничительные примеры биотинилированных dNTP включают биотин-dATP (например, биотин-N6-ddATP, биотин-14-dATP), биотин-dCTP (например, биотин-11-dCTP, биотин-14-dCTP) и биотин-dUTP (например, биотин-11-dUTP, биотин-16-dUTP, биотин-20-dUTP).

Термины "полинуклеотид", "олигонуклеотид" и "нуклеиновая кислота" применяются взаимозаменяемо для обозначения полимерной формы нуклеотидов любой длины, будь то дезоксирибонуклеотидов, рибонуклеотидов или их аналогов, в одно-, двух- или многоцепочечном виде. Полинуклеотиды могут быть экзогенными или эндогенными для клетки. Полинуклеотиды могут существовать в бесклеточной среде. Полинуклеотид может представлять собой ген или его фрагмент. Полинуклеотид может представлять собой ДНК. Полинуклеотид может представлять собой РНК. Полинуклеотид может иметь любую трехмерную структуру и может выполнять любую функцию, известную или неизвестную. Полинуклеотид может включать в себя один или несколько аналогов (например, с измененным остовом, сахаром или основанием). Если они присутствуют, то модификации структуры нуклеотидов могут происходить до или после сборки полимера. Некоторые неограничительные примеры аналогов включают: 5-бромурацил, пептидонуклеиновые кислоты, ксенонуклеиновые кислоты, морфолиновые олигонуклеотиды, блокированные нуклеиновые кислоты, гликоленуклеиновые кислоты, треозонуклеиновые кислоты, дидезокси-нуклеотиды, кордицепин, 7-деза-ГТФ, флуорофоры (например, связанный с сахаром родамин или флуоресцеин), тиолсодержащие нуклеотиды, связанные с биотином нуклеотиды, флуоресцентные аналоги оснований, островки CpG, метил-7-гуанозин, метилированные нуклеотиды, инозин, тиюридин, псевдодуридин, дигидродуридин, кьюозин и вайозин. Неограничительные примеры полинуклеотидов включают кодирующие или не кодирующие области гена или фрагмента гена, локусы, установленные при анализе сцепления, экзоны, интроны, матричные РНК (мРНК), транспортные РНК (тРНК), рибосомные РНК (рРНК), короткие интерферирующие РНК (siRNA), короткие шпилечные РНК (shRNA), микроРНК (miRNA), рибозимы, кДНК, рекомбинантные полинуклеотиды, разветвленные полинуклеотиды, плазмиды, векторы, выделенные ДНК с любой последовательностью, выделенные РНК с любой последовательностью, бесклеточные полинуклеотиды, включая бесклеточную ДНК (cfDNA) и бесклеточную РНК (cfRNA), зонды из нуклеиновых кислот и праймеры. Последовательность нуклеотидов может прерываться ненуклеотидными компонентами.

Термины "целевой полинуклеотид" и "целевая нуклеиновая кислота" в настоящем изобретении относятся к такой нуклеиновой кислоте или к таким полинуклеотидам, на которые нацелен исполнительный элемент настоящего изобретения. Целевой нуклеиновой кислотой может быть ДНК. Целевой нуклеиновой кислотой может быть РНК. Целевой нуклеиновой кислотой может быть хромосомная последовательность или внехромосомная последовательность (например, эписомная последовательность, последовательность миникольца, митохондриальная последовательность, хлоропластная последовательность и т.д.). Целевой нуклеиновой кислотой может быть последовательность нуклеиновой кислоты, которая может отличаться от какой-либо другой последовательности в образце нуклеиновой кислоты заменой единственного нуклеотида. Целевой нуклеиновой кислотой может быть последовательность нуклеиновой кислоты, которая может отличаться от какой-либо другой последовательности в образце нуклеиновой кислоты заменой 2, 3, 4, 5, 6, 7, 8, 9 или 10 нуклеотидов. В некоторых воплощениях замена не может происходить в пределах 5, 10, 15, 20, 25, 30 или 35 нуклеотидов от 5'-конца целевой нуклеиновой кислоты. В некоторых воплощениях замена не может происходить в пределах 5, 10, 15, 20, 25, 30, 35 нуклео-

тидов от 3'-конца целевой нуклеиновой кислоты. В общем, термин "целевая последовательность" относится к последовательности нуклеиновой кислоты на одной нити целевой нуклеиновой кислоты. Целевой последовательностью может быть часть гена, регуляторной последовательности, геномной ДНК, бесклеточной нуклеиновой кислоты, в том числе cfDNA и/или cfRNA, кДНК, слитого гена или РНК, включая мРНК, miRNA, рРНК и др.

Термин "ген" в настоящем изобретении относится к нуклеиновой кислоте (например, ДНК типа геномной ДНК и кДНК) и соответствующей последовательности нуклеотидов, участвующей в кодировании РНК-транскрипта. Этот термин применительно к геномной ДНК включает промежуточные, некодирующие участки, а также регуляторные участки и может включать 5'- и 3'-концы. В некоторых случаях этот термин охватывает транскрибируемые последовательности, включая 5'- и 3'-нетранслируемые участки (5'-UTR и 3'-UTR), экзоны и интроны. У некоторых генов транскрибируемая область содержит "открытые рамки считывания", кодирующие полипептиды. В некоторых случаях термин "ген" включает только кодирующие последовательности (например, "открытую рамку считывания" или "кодирующую область"), необходимые для кодирования полипептида. В некоторых случаях гены не кодируют полипептиды, к примеру, гены рибосомной РНК (рРНК) и гены транспортной РНК (тРНК). В некоторых случаях термин "ген" включает не только транскрибируемые последовательности, но также и нетранскрибируемые участки, в том числе выше- и нижележащие регуляторные участки, энхансеры и промоторы. Ген может означать "эндогенный ген" или собственный ген в своем естественном положении в геноме организма. Ген может означать "экзогенный ген" или несобственный ген. Несобственный ген может означать ген, который в норме не встречается в организме хозяина, а введен в организм хозяина путем переноса гена. Несобственный ген также может означать ген, который находится не в своем естественном месте в геноме организма. Несобственный ген также может означать природную нуклеиновую кислоту или последовательность полипептида, которая включает мутации, вставки и/или делеции (например, чужеродные последовательности).

Термины "трансфекция" и "трансфецированный" относятся к введению нуклеиновой кислоты в клетки невирусными или вирусными методами. Молекулы нуклеиновой кислоты могут представлять собой последовательности генов, кодирующие полные белки или их функциональные части. Например, см. Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, 18.1-18.88.

Термин "экспрессия" относится к одному или нескольким процессам, посредством которых полинуклеотид транскрибируется из ДНК-матрицы (как-то в мРНК или в другой РНК-транскрипт), и/или процессу, посредством которого транскрибируемая мРНК впоследствии транслируется в пептиды, полипептиды или белки. Транскрипты и кодируемые полипептиды могут собирательно называться "генными продуктами". Если полинуклеотид происходит из геномной ДНК, то экспрессия может включать сплайсинг мРНК в эукариотической клетке. "Повышающая регуляция" в отношении экспрессии обычно означает повышение уровня экспрессии полинуклеотида (например, РНК типа мРНК) и/или последовательности полипептида по сравнению с уровнем экспрессии у дикого типа, а "понижающая регуляция" обычно означает снижение уровня экспрессии полинуклеотида (например, РНК типа мРНК) и/или последовательности полипептида по сравнению с его экспрессией у дикого типа. Экспрессия трансфецированного гена в клетках может быть краткосрочной или стабильной. При "краткосрочной экспрессии" трансфецированный ген не передается дочерней клетке при клеточном делении. Поскольку его экспрессия ограничивается трансфецированными клетками, экспрессия этого гена со временем утрачивается. Напротив, стабильная экспрессия трансфецированного гена происходит тогда, когда ген трансфецируется вместе с другим геном, который придает селекционное преимущество трансфецированной клетке. Таким селекционным преимуществом может быть устойчивость к определенному токсину, который поступает в клетку.

Термины "экспрессионная кассета", "экспрессирующая конструкция" или "экспрессирующий вектор" относятся к такой нуклеиновой кислоте, которая включает в себя последовательность нуклеотидов типа кодирующей последовательности и матричной последовательности, а также последовательности, необходимые для экспрессии кодирующей последовательности. Экспрессионная кассета может быть вирусной или не вирусной. Например, экспрессионная кассета содержит конструкцию из нуклеиновой кислоты, которая при введении в клетки хозяина приводит к транскрипции и/или трансляции РНК или полипептида, соответственно. Антисмысловые и смысловые конструкции, которые не транслируются или не могут транслироваться, также подпадают под это определение. Специалистам должно быть известно, что последовательности вставляемых полинуклеотидов не обязательно должны быть идентичны, а могут быть лишь существенно близки последовательности того гена, из которого они получены.

"Плаزمида" в настоящем изобретении обычно означает невирусный экспрессирующий вектор, например, молекулы нуклеиновой кислоты, кодирующие гены, и/или регуляторные элементы, необходимые для экспрессии генов. "Вирусный вектор" в настоящем изобретении обычно означает происходящую из вируса нуклеиновую кислоту, которая способна транспортировать в клетки другую нуклеиновую кислоту. Вирусный вектор способен направлять экспрессию белка или белков, кодируемых одним или несколькими генами, переносимыми вектором, когда он попадает в соответствующее окружение. Примеры вирусных векторов включают, без ограничения, ретровирусные, аденовирусные, лентивирусные и адено-

ассоциированные вирусные векторы.

Термин "промотор" в настоящем изобретении означает последовательность полинуклеотида, способного управлять транскрипцией кодирующей последовательности в клетке. Так, промоторы, используемые в полинуклеотидных конструкциях по изобретению, включают действующие в цис-положении контрольные элементы транскрипции и регуляторные последовательности, которые участвуют в регуляции или модулировании срока и/или скорости транскрипции гена. Например, промотор может представлять собой действующий в цис-положении контрольный элемент транскрипции, в том числе энхансер, промотор, терминатор транскрипции, начало репликации, последовательность встраивания в хромосомы, 5'- и 3'-нетранслируемые участки или интронные последовательности, которые участвуют в регуляции транскрипции. Эти действующие в цис-положении последовательности обычно взаимодействуют с белками или другими биомолекулами для осуществления (включения/выключения, регуляции, модуляции и т.п.) транскрипции генов. "Конститутивный промотор" способен инициировать транскрипцию почти во всех типах тканей, тогда как "тканеспецифичный промотор" инициирует транскрипцию только в одном или в немногих конкретных типах тканей. "Индукционный промотор" инициирует транскрипцию только при определенных условиях среды, стадии развития или же лекарственных или химических условиях.

Термины "комплементируют", "комплементирует", "комплементарный" и "комплементарность" в настоящем изобретении обычно относятся к последовательности, которая полностью комплементарна и гибридизуется с данной последовательностью. В некоторых случаях последовательность, которая гибридизуется с данной нуклеиновой кислотой, называется "комплементарной" или "обратно комплементарной" данной молекуле, если её последовательность оснований на данном участке способна комплементарно связываться с таковой у партнера по связыванию с тем, что образуются, к примеру, пары оснований А-Т, А-У, G-С и G-U. В общем, первая последовательность, которая гибридизуется со второй последовательностью, гибридизуется специфически или избирательно со второй последовательностью, так как гибридизация со второй последовательностью или набором вторых последовательностей является предпочтительной (например, термодинамически более стабильна при данном наборе условий типа строгих условий, которые обычно применяются в данной области), чем гибридизация с нецелевыми последовательностями при реакции гибридизации. Как правило, у гибридизуемых последовательностей степень комплементарности последовательностей по всей длине или по части соответствующей длины составляет от 25% до 100%, включая комплементарность последовательностей по меньшей мере на 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 и 100%. Идентичность последовательностей, как-то с целью оценки степени комплементарности, можно измерить при помощи любого подходящего алгоритма совмещения, включая, без ограничения, алгоритм Needleman-Wunsch (например, см. EMBOSS Needle aligner, доступный на www.ebi.ac.uk/Tools/psa/emboss_needle/nucleotide.html, необязательно с настройками по умолчанию), алгоритм BLAST (например, см. инструмент совмещения BLAST, доступный на blast.ncbi.nlm.nih.gov/Blast.cgi, необязательно с настройками по умолчанию) или алгоритм Smith-Waterman (например, см. EMBOSS Water aligner, доступный на www.ebi.ac.uk/Tools/psa/emboss_water/nucleotide.html, необязательно с настройками по умолчанию). Оптимальность совмещения можно оценить, используя любые подходящие параметры выбранного алгоритма, включая параметры по умолчанию.

Комплементарность может быть идеальной или существенной/достаточной. Идеальная комплементарность между двумя нуклеиновыми кислотами означает, что две нуклеиновые кислоты могут образовывать дуплекс, в котором каждое основание дуплекса связано с комплементарным основанием путем спаривания Ватсона-Крика. Существенная/достаточная комплементарность — это последовательность в одной нити не полностью и/или не совершенно комплементарна последовательности в противоположной нити, но происходит достаточное связывание между основаниями на двух нитях для образования стабильного гибридного комплекса при заданных условиях гибридизации (например, концентрации солей и температуре). Такие условия могут быть предсказаны с использованием последовательностей и стандартных математических расчетов для прогнозирования T_m гибридизуемых нитей или путем эмпирического определения T_m стандартными методами.

Термины "пептид", "полипептид" и "белок" применяются взаимозаменяемо для обозначения полимеров по меньшей мере из двух аминокислотных остатков, соединенных пептидной связью. Этот термин не подразумевает определенную длину полимера и не означает или не подразумевает, что пептид получен рекомбинантными методами, путем химического или ферментативного синтеза или является естественным. Эти термины применимы к природным аминокислотным полимерам, а также к аминокислотным полимерам, содержащим по меньшей мере одну модифицированную аминокислоту. В некоторых случаях полимер может перемежаться не аминокислотами. Термины включают аминокислотные цепи любой длины, включая полноразмерные белки, и белки с вторичной и/или третичной структурой (например, домены) или без неё. Термины также охватывают аминокислотные полимеры, которые подверглись модификации, к примеру, посредством образования дисульфидной связи, гликозилирования, липидирования, ацетилирования, фосфорилирования, окисления и любых других манипуляций типа конъюгирования с маркирующим компонентом. Термины "аминокислота" и "аминокислоты" в настоящем изобретении обычно относятся к природным и не природным аминокислотам, включая, без ограничения, модифи-

цированные аминокислоты и аналоги аминокислот. Модифицированные аминокислоты могут включать природные аминокислоты и неприродные аминокислоты, которые подвергались химической модификации для включения группы или химического компонента, который в природе не встречается у данной аминокислоты. Аналоги аминокислот могут относиться к производным аминокислот. Термин "аминокислота" включает и D-аминокислоты, и L-аминокислоты.

Термины "производное", "вариант" и "фрагмент" применительно к полипептидам в настоящем изобретении относятся к полипептидам, родственным полипептиду дикого типа, например, по аминокислотной последовательности, структуре (например, вторичной и/или третичной), активности (например, ферментативной активности) и/или функции. Производные, варианты и фрагменты полипептидов могут содержать одно или несколько вариаций (например, мутаций, вставок и делеций), укорочений, модификаций аминокислот либо их комбинаций по сравнению с полипептидом дикого типа.

Термин "степень (%) идентичности" в настоящем изобретении означает, какой процент аминокислотных остатков (или нуклеотидов) в последовательности кандидата идентичен аминокислотным остаткам (или нуклеотидам) эталонной последовательности после совмещения последовательностей и, при необходимости, введения пробелов для достижения максимальной степени идентичности (т.е. пробелы можно вводить в одну или обе последовательности кандидата и эталона для оптимального совмещения, а негомологичные последовательности можно игнорировать в целях сравнения). Совмещение в целях определения степени идентичности может осуществляться различными способами, которые входят в компетенцию специалистов, например, с помощью общедоступных компьютерных программ типа BLAST, ALIGN или Megalign (DNASTAR). Степень идентичности двух последовательностей можно рассчитать путем совмещения исследуемой последовательности со сравнительной последовательностью с помощью BLAST, определения количества аминокислот или нуклеотидов в совмещаемой исследуемой последовательности, которые идентичны аминокислотам или нуклеотидам в таких же положениях у сравнительной последовательности, и деления количества идентичных аминокислот или нуклеотидов на количество аминокислот или нуклеотидов в сравнительной последовательности.

Термин "полипептид, модулирующий ген," или "GMP", в настоящем изобретении относится к полипептидам, содержащим по меньшей мере исполнительный элемент, способный регулировать экспрессию или активность гена и/или редактировать последовательность нуклеиновой кислоты. GMP может содержать дополнительные пептидные последовательности, которые не участвуют в модулировании экспрессии гена, например, сайты распознавания расщепления, линкерные последовательности, наводящие последовательности и пр.

Термин "исполнительный элемент" в настоящем изобретении относится к элементу, который может регулировать экспрессию или активность гена и/или редактировать последовательность нуклеиновой кислоты, будь то экзогенной или эндогенной. Исполнительный элемент может регулировать экспрессию гена на уровне транскрипции и/или на уровне трансляции. Исполнительный элемент может регулировать экспрессию гена на уровне транскрипции, к примеру, путем регуляции получения мРНК из ДНК типа хромосомной ДНК или кДНК. В некоторых воплощениях исполнительный элемент рекрутирует по меньшей мере один фактор транскрипции, который связывается с определенной последовательностью ДНК, тем самым контролируя скорость транскрипции генетической информации из ДНК в мРНК. Исполнительный элемент может и сам связываться с ДНК и регулировать транскрипцию посредством физической преграждения, к примеру, предотвращая сборку белков типа РНК-полимеразы и других ассоциированных белков на ДНК-матрице. Исполнительный элемент может регулировать экспрессию гена на уровне трансляции, к примеру, путем регуляции получения белка из матрицы мРНК. В некоторых воплощениях исполнительный элемент регулирует экспрессию гена путем воздействия на стабильность транскрипта мРНК. В некоторых воплощениях исполнительный элемент регулирует экспрессию гена путем редактирования последовательности нуклеиновой кислоты (например, участка генома). В некоторых воплощениях исполнительный элемент регулирует экспрессию гена путем редактирования мРНК матрицы. Редактирование последовательности нуклеиновой кислоты в некоторых случаях может приводить к изменению базовой матрицы для экспрессии гена.

Упомянутый здесь белок Cas означает тип белка или полипептида. Белок Cas может означать нуклеазу. Белок Cas может означать эндорибонуклеазу. Белок Cas может означать любой модифицированный (например, укороченный, мутированный, удлиненный) полипептид или гомолог белка Cas. Белок Cas может быть оптимизирован по кодонам. Белок Cas может быть оптимизированным по кодонам гомологом Cas-белка. Белок Cas может быть энзиматически неактивным, частично активным, конститутивно активным, полностью активным, индуцибельно активным и/или более активным (например, больше, чем гомолог белка или полипептида дикого типа). Белок Cas может означать Cas9. Белок Cas может означать Cpf1. Белок Cas может означать C2c2. Белок Cas (например, вариант, мутантный, энзиматически неактивный и/или условно энзиматически неактивный сайт-направленный полипептид) может связываться с целевой нуклеиновой кислотой. Белок Cas (например, вариант, мутантная, энзиматически неактивная и/или условно энзиматически неактивная эндорибонуклеаза) может связываться с целевой РНК или ДНК.

Термин "crRNA" в настоящем изобретении в общем относится к такой нуклеиновой кислоте, кото-

рая по последовательности по меньшей мере на 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 или 100% идентична и/или аналогична типичной crRNA дикого типа (например, crRNA из *S.pyogenes*). crRNA обычно означает такую нуклеиновую кислоту, которая по последовательности не более чем на 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 или 100% идентична и/или аналогична типичной crRNA дикого типа (например, crRNA из *S.pyogenes*). crRNA может означать модифицированную форму crRNA, которая может содержать изменения нуклеотидов типа делеции, вставки или замены, вариант, мутантную или химерную форму. crRNA может представлять собой такую нуклеиновую кислоту, которая по последовательности по меньшей мере на 60% идентична последовательности типичной crRNA дикого типа (например, crRNA из *S.pyogenes*) на отрезке из по меньшей мере 6 смежных нуклеотидов. Например, последовательность crRNA может быть по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 70%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 98%, по меньшей мере на 99% или на 100% идентична типичной последовательности crRNA дикого типа (например, crRNA из *S.pyogenes*) на отрезке из по меньшей мере 6 смежных нуклеотидов.

Термин "tracrRNA" в настоящем изобретении в общем относится к такой нуклеиновой кислоте, которая по последовательности по меньшей мере на 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 или 100% идентична и/или аналогична последовательности типичной tracrRNA дикого типа (например, tracrRNA из *S.pyogenes*). tracrRNA может означать такую нуклеиновую кислоту, которая по последовательности не более чем на 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 или 100% идентична и/или аналогична последовательности типичной tracrRNA дикого типа (например, tracrRNA из *S.pyogenes*). tracrRNA может означать модифицированную форму tracrRNA, которая может содержать изменения нуклеотидов типа делеции, вставки или замены, вариант, мутантную или химерную форму. tracrRNA может означать такую нуклеиновую кислоту, которая по последовательности по меньшей мере на 60% идентична последовательности типичной tracrRNA дикого типа (например, tracrRNA из *S.pyogenes*) на отрезке из по меньшей мере 6 смежных нуклеотидов. Например, последовательность tracrRNA может быть по меньшей мере на 60%, по меньшей мере на 65%, по меньшей мере на 70%, по меньшей мере на 75%, по меньшей мере на 80%, по меньшей мере на 85%, по меньшей мере на 90%, по меньшей мере на 95%, по меньшей мере на 98%, по меньшей мере на 99% или на 100% идентична типичной последовательности tracrRNA дикого типа (например, tracrRNA из *S.pyogenes*) на отрезке из по меньшей мере 6 смежных нуклеотидов.

В настоящем изобретении "направляющая нуклеиновая кислота" может означать такую нуклеиновую кислоту, которая может гибридизоваться с другой нуклеиновой кислотой. Направляющей нуклеиновой кислотой может быть РНК. Направляющей нуклеиновой кислотой может быть ДНК. Направляющая нуклеиновая кислота может быть запрограммирована так, чтобы она сайт-специфически связывалась с последовательностью нуклеиновой кислоты. Нуклеиновая кислота мишени или целевая нуклеиновая кислота может содержать нуклеотиды. Направляющая нуклеиновая кислота может содержать нуклеотиды. Часть нуклеиновой кислоты-мишени может быть комплементарна части гидовой нуклеиновой кислоты. Та нить двухцепочечного целевого полинуклеотида, которая комплементарна и гибридизуется с направляющей нуклеиновой кислотой, может называться комплементарной нитью. Та нить двухцепочечного целевого полинуклеотида, которая комплементарна комплементарной нити и поэтому не может быть комплементарна направляющей нуклеиновой кислоте, может называться некомплемтарной нитью. Направляющая нуклеиновая кислота может содержать одну полинуклеотидную цепь и может называться "одинарной направляющей нуклеиновой кислотой". Направляющая нуклеиновая кислота может содержать две полинуклеотидные цепи и может называться "двойной направляющей нуклеиновой кислотой". Если не указано иначе, термин "направляющая нуклеиновая кислота" может быть включительным, относящимся и к одинарным гидовым нуклеиновым кислотам, и к двойным гидовым нуклеиновым кислотам.

Направляющая нуклеиновая кислота может содержать сегмент, который может упоминаться как "сегмент, нацеливающий на нуклеиновую кислоту" или "последовательность наводки на нуклеиновую кислоту". Нацеливающий на нуклеиновую кислоту сегмент может содержать подсегмент, который может называться "сегментом связывания белка" или "последовательностью связывания белка" или "сегментом связывания белка Cas".

Термин "сайт распознавания расщепления" применительно к пептидам в настоящем изобретении относится к такому сайту пептида, по которому может расщепляться химическая связь типа пептидной связи или дисульфидной связи. Расщепление может осуществляться различными способами. Расщепление пептидных связей может осуществляться, к примеру, ферментом типа протеазы или сплайсинг-белком (например, интеином). Расщепление дисульфидной связи может осуществляться, к примеру, ферментом типа оксидоредуктазы.

Термин "нацеливающая последовательность" в настоящем изобретении относится к такой последовательности нуклеотидов и соответствующей аминокислотной последовательности, которая кодирует нацеливающий полипептид, который опосредует локализацию (или удерживание) белка в субклеточной локализации, например, плазматической мембране или мембране заданной органеллы, ядре, цитозоле, митохондриях, эндоплазматическом ретикулуме (ER), аппарате Гольджи, хлоропластах, апопластах, пе-

роксисомах или других органеллах. Например, нацеливающая последовательность может направлять белок (например, рецепторный полипептид или адаптерный полипептид) в ядро, используя сигнал ядерной локализации (NLS); из ядра клетки, к примеру, в цитоплазму, используя сигнал ядерного экспорта (NES); в митохондриях, используя сигнал митохондриальной наводки; в эндоплазматический ретикулум (ER), используя сигнал удержания в ER; в пероксисомы, используя сигнал пероксисомной наводки; в плазматическую мембрану, используя сигнал мембранной локализации; либо их комбинации.

В настоящем изобретении "слияние" может относиться к белкам и/или нуклеиновым кислотам, содержащим одну или несколько несобственных последовательностей (например, частей). Слияние может включать одну или несколько одинаковых несобственных последовательностей. Слияние может включать одну или несколько разных несобственных последовательностей. Слияние может представлять собой химеру. Слияние может включать аффинную метку нуклеиновой кислоты. Слияние может включать баркод. Слияние может включать аффинную метку пептида. Слияние может обеспечивать субклеточную локализацию направляемого на место полипептида (например, сигнал ядерной локализации (NLS) для наведения на ядро, сигнал митохондриальной локализации для наведения на митохондрии, сигнал хлоропластной локализации для наведения на хлоропласты, сигнал удержания в эндоплазматическом ретикулуме (ER) и др.). Слияние может обеспечить несобственную последовательность (например, аффинную метку), которая может использоваться для отслеживания или очистки. Слияние может представлять собой небольшую молекулу типа биотина или краситель типа флуоресцентных красителей Alexa, красителя Cyanine3, красителя Cyanine5.

Слияние может относиться к любому белку с функциональным действием. Например, слитый белок может включать активность метилтрансферазы, активность деметилазы, активность дисмутазы, активность алкилирования, активность депуринизации, активность окисления, активность образования пиримидиновых димеров, активность интегразы, активность транспозазы, активность рекомбиназы, активность полимеразы, активность лигазы, активность геликазы, активность фотолиазы или активность гликозилазы, активность ацетилтрансферазы, активность деацетилазы, активность киназы, активность фосфатазы, активность убиквитин-лигазы, активность деубиквитинирования, активность аденилирования, активность деаденилирования, активность сумоилирования, активность десумоилирования, активность рибозилирования, активность дерибозилирования, активность миристоилирования, активность ремоделирования, активность протеазы, активность оксидоредуктазы, активность трансферазы, активность гидролазы, активность лиазы, активность изомеразы, активность синтазы, активность синтетазы или активность демиростоилирования. Эффекторный белок может модифицировать геномный локус. Слитый белок может быть продуктом слияния в белке Cas. Слитый белок может быть несобственной последовательностью в белке Cas.

В настоящем изобретении "несобственный" может означать последовательность нуклеиновой кислоты или полипептида, которая не встречается у нативной нуклеиновой кислоты или белка. Несобственные могут быть аффинные метки. Несобственные могут быть партнеры по слиянию. Несобственные могут быть последовательности природной нуклеиновой кислоты или полипептида, которые содержат мутации, вставки и/или делеции. Несобственная последовательность может проявлять или кодировать активность (например, ферментативную активность, активность метилтрансферазы, активность ацетилтрансферазы, активность киназы, активность убиквитинирования и т.д.), которую также может проявлять последовательность нуклеиновой кислоты и/или полипептида, с которой слита несобственная последовательность. Несобственную последовательность нуклеиновой кислоты или полипептида можно соединить с последовательностью природной нуклеиновой кислоты или полипептида (или ее вариантом) при помощи генной инженерии, получая химерную последовательность нуклеиновой кислоты и/или полипептида, кодирующую химерную нуклеиновую кислоту и/или полипептид.

В настоящем изобретении "лечение" или "обработка" означает любое из следующего: облегчение одного или нескольких симптомов заболевания, например, рака; предупреждение появления таких симптомов еще до того, как они возникнут; замедление или полное предотвращение прогрессирования заболевания (о чем может свидетельствовать удлинение промежутка между повторными эпизодами, замедление или предотвращение ухудшения симптомов и т.п.); ускорение наступления периода ремиссии; замедление необратимых повреждений, возникающих на прогрессирующей/хронической стадии болезни (как на первичной, так и на вторичной стадии); замедление наступления такой прогрессирующей стадии; или любые их комбинации.

В настоящем изобретении "вводить", "вводится", "введение" и их производные относятся к методам, которые могут применяться для доставки средств или композиций к требуемому месту биологического действия. Эти методы включают, без ограничения, парентеральное введение (например, внутривенное, подкожное, внутривнутрибрюшинное, внутримышечное, внутрисосудистое, интратекальное, интраназальное, интравитреальное, вливание и местное введение), трансмукозальное, пероральное введение, введение в виде свечей и топическое введение. Введение осуществляется любым способом, включая парентеральные. Парентеральное введение включает, например, внутривенное, внутримышечное, внутривнутриартериальное, интравенеральное, подкожное, внутривнутрибрюшинное, интравентрикулярное и внутривнутричерепное. Другие способы доставки включают, без ограничения, липосомные формы, внутривенное вливание,

трансплантации и т.п. Специалистам в данной области должны быть известны и другие способы введения терапевтически эффективного количества композиции настоящего изобретения для профилактики или снятия одного или нескольких симптомов, связанных с заболеванием.

Здесь представлены системы, способы и композиции для регуляции экспрессии целевых полинуклеотидов в клетках. В одном аспекте настоящего изобретения предусмотрены системы для регуляции экспрессии целевых полинуклеотидов в клетках. Типичная система включает: (a) химерный рецепторный полипептид, который подвергается модификации при связывании антигена, причем модификация рецептора включает конформационное изменение или химическую модификацию; (b) химерный адаптерный полипептид, который связывает рецептор в ответ на модификацию рецептора; (c) полипептид, модулирующий ген (GMP), содержащий исполнительный элемент, соединенный с сайтом распознавания расщепления, причем при расщеплении сайта распознавания расщепления исполнительный элемент активируется и образует комплекс с целевым полинуклеотидом; и (d) расщепляющий элемент, который расщепляет сайт распознавания расщепления, когда он находится вблизи сайта распознавания расщепления. Химерный рецепторный полипептид, химерный адаптерный полипептид, полипептид, модулирующий ген (GMP) и расщепляющий элемент данной системы могут располагаться в различных конфигурациях. Здесь описаны типичные, неограничительные конфигурации. В некоторых воплощениях GMP входит в состав внутриклеточной области химерного рецепторного полипептида, а расщепляющий элемент входит в состав химерного адаптерного полипептида. В некоторых воплощениях GMP входит в состав химерного адаптерного полипептида, а расщепляющий элемент входит в состав внутриклеточной области химерного рецепторного полипептида. В некоторых воплощениях расщепляющий элемент образует комплекс со вторым адаптерным полипептидом, который связывает химерный рецепторный полипептид в ответ на модификацию рецептора, а GMP входит в состав химерного адаптерного полипептида.

В типичной конфигурации GMP входит в состав внутриклеточной области химерного рецепторного полипептида, а расщепляющий элемент входит в состав химерного адаптерного полипептида. Химерный рецепторный полипептид в типичной конфигурации может включать (a) взаимодействующий с антигеном домен и (b) полипептид, модулирующий ген (GMP), содержащий исполнительный элемент, соединенный с сайтом распознавания расщепления. На фиг. 1 представлен типичный химерный рецепторный полипептид. Рецептор включает взаимодействующий с антигеном домен 101 и полипептид, модулирующий ген (GMP) 102. GMP 102 может содержать исполнительный элемент 102a, соединенный с сайтом распознавания расщепления 102b.

В некоторых воплощениях (i) химерный рецепторный полипептид подвергается модификации в ответ на связывание антигена, (ii) сайт распознавания расщепления расщепляется расщепляющим элементом в ответ на модификацию химерного рецепторного полипептида, (iii) исполнительный элемент образует комплекс с целевым полинуклеотидом после отщепления от химерного рецепторного полипептида по сайту распознавания расщепления, и (iv) химерный рецепторный полипептид не содержит SEQ ID NO: 39:

```
ILDYSFTGGAGRDIPPPQIEEACELPECQVDAGNKVCNLQCNNHACGWDGGDC
SLNFNDPWKNCTQSLQCWKYFSDGHCDSCNSAGCLFDGFDQCLTEGQCNPYDQYC
KDHFSGDGHCDQGCNSAECEWDGLDCAEHVPERLAAGTLVLVLLPPDQLRNNSEHFLR
ELSHVLHTNVVFKRDAQGQMMIFPPYGHHEELRKHPIKRSTVGWATSSLLPGTSGGRQR
RELDPMDIRGSIVYLEIDNRQCVQSSQCFQSATDVA AFLGALASLGS LNIPYKIEAVKSE
PVEPPLPSQLHLMYVAAA FVLLFFVGC GVLLSRKRRR.
```

Химерный рецепторный полипептид рассматриваемой системы может включать эндогенный рецептор либо его производное, вариант или фрагмент. Химерный рецепторный полипептид может специфически связываться по меньшей мере с одним антигеном (например, по меньшей мере с одним лигандом), к примеру, через взаимодействующий с антигеном домен (который также именуется здесь "внеклеточным сенсорным доменом"). Химерный рецепторный полипептид, в ответ на связывание лиганда, может подвергаться модификации типа конформационного изменения или химической модификации. Такие модификации могут рекрутировать к рецептору партнеров по связыванию (например, партнеров типа белков), включая, без ограничения, сигнальные белки, участвующие в событиях сигнализации и различных клеточных процессах. Сигнальные белки, к примеру, могут участвовать в регулировании (например, активации и/или деактивации) клеточных реакций типа запрограммированных изменений в экспрессии генов путем регуляции трансляции, регуляции транскрипции и эпигенетических модификаций, включая регуляцию метилирования, ацетилирования, фосфорилирования, убиквитинилирования, сумоилирования, рибозилирования и цитруллинирования. Конформационные изменения химерного рецепторного полипептида могут экспонировать один или несколько участков рецептора, которые ранее не были экспонированными, а экспонированные участки могут рекрутировать и/или связывать сигнальные белки. Химические модификации рецептора, к примеру, фосфорилирование и/или дефосфорилирование (например, по остаткам тирозина, серина, треонина и/или других подходящих аминокислот), также могут рекрутировать сигнальные белки, участвующие в регуляции внутриклеточных процессов. Сигнальные

белки могут связываться с рецептором прямо или косвенно, к примеру, в составе более крупного комплекса.

В некоторых воплощениях химерный рецепторный полипептид является трансмембранным рецептором. Типичный трансмембранный рецептор представлен на фиг. 2. Трансмембранный рецептор может заключаться в клеточной мембране и содержать как минимум внеклеточную область 201, пронизывающую мембрану типа плазматической мембраны область 202 и внутриклеточную область 203. Взаимодействующий с антигеном домен может входить в состав внеклеточной области, а GMP может входить в состав внутриклеточной области. Мембранные рецепторы могут детектировать по меньшей мере один сигнал типа небольшой молекулы, иона или белка из окружающей среды (например, внеклеточной и/или внутриклеточной среды) и могут запускать клеточные реакции через по меньшей мере один сигнализирующий каскад с участием других белков и сигнальных молекул. Некоторые рецепторы могут транслоцироваться из одного района клетки в другой, к примеру, из плазматической мембраны или цитоплазмы в ядро и наоборот. Такая транслокация может быть обусловлена связыванием лиганда с рецептором. Примеры мембранных рецепторов включают, без ограничения, рецепторы Notch; сопряженные с G-белком рецепторы (GPCRs); интегриновые рецепторы; кадгериновые рецепторы; каталитические рецепторы, в том числе рецепторы, обладающие ферментативной активностью, и рецепторы, которые сами не обладают ферментативной активностью, а действуют путем стимуляции нековалентно связанных ферментов (например, киназ); рецепторы смерти типа представителей суперсемейства рецепторов фактора некроза опухолей (TNFR); и иммуно-рецепторы.

В некоторых воплощениях химерный рецепторный полипептид включает рецептор Notch либо его производное, вариант или фрагмент. Рецепторы Notch - это трансмембранные белки, которые опосредуют сигнализацию межклеточных контактов и играют ключевую роль в развитии и других аспектах межклеточной коммуникации, например коммуникации между контактирующими клетками (принимающей клеткой и передающей клеткой). Рецепторы Notch, экспрессированные в принимающей клетке, распознают свои лиганды (семейство белков Delta), экспрессированные на передающей клетке. Сцепление Notch и Delta на этих контактирующих клетках ведет к двухстадийному протеолизу рецептора Notch, что в конечном счете вызывает высвобождение внутриклеточной части рецептора из мембраны в цитоплазму.

В некоторых воплощениях химерный рецепторный полипептид включает как минимум внеклеточную область (например, лиганд-связывающий домен) рецептора Notch либо его производного, варианта или фрагмента. В некоторых воплощениях химерный рецепторный полипептид включает, по меньшей мере, пронизывающую мембрану область Notch либо его производного, варианта или фрагмента. В некоторых воплощениях химерный рецепторный полипептид включает, по меньшей мере, внутриклеточную область (например, цитоплазматический домен) Notch либо его производного, варианта или фрагмента. Химерный рецепторный полипептид, содержащий Notch либо его производное, вариант или фрагмент, может рекрутировать партнера по связыванию. В некоторых воплощениях связывание лиганда с химерным рецептором, содержащим Notch либо его производное, вариант или фрагмент, приводит к конформационному изменению, химической модификации либо их комбинации, которые рекрутируют партнера по связыванию к рецептору.

В некоторых воплощениях химерный рецепторный полипептид содержит Notch либо его производное, вариант или фрагмент из числа Notch1, Notch2, Notch3, Notch4 или их гомологов.

В некоторых воплощениях химерный рецепторный полипептид содержит сопряженный с G-белком рецептор (GPCR) либо его производное, вариант или фрагмент. GPCRs в общем характеризуются наличием 7 пронизывающих мембрану α -спиралей и могут располагаться в виде третичной структуры, похожей на бочку, причем 7 трансмембранных спиралей образуют полость в пределах плазматической мембраны, которая служит в качестве связывающего лиганд домена. Лиганды также могут связываться с GPCR и по другим местам, к примеру, со внеклеточными петлями и/или N-концевым хвостом. Связывание лиганда может активировать ассоциированный G-белок, который будет затем функционировать в различных сигнальных путях. Чтобы дезактивировать эту сигнализацию, GPCR может сначала подвергнуться химической модификации путем фосфорилирования. Затем фосфорилирование может рекрутировать коадаптерные белки (например, белки аррестины) для дополнительной сигнализации.

В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внеклеточную область (например, лиганд-связывающий домен) GPCR либо его производного, варианта или фрагмента. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, пронизывающую мембрану область GPCR либо его производного, варианта или фрагмента. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внутриклеточную область (например, цитоплазматический домен) GPCR либо его производного, варианта или фрагмента. Химерный рецепторный полипептид, содержащий GPCR либо его производное, вариант или фрагмент, может рекрутировать партнера по связыванию. В некоторых воплощениях связывание лиганда с химерным рецептором, содержащим GPCR либо его производное, вариант или фрагмент, приводит к конформационному изменению, химической модификации либо их комбинации, которые рекрутируют партнера по связыванию к рецептору.

В некоторых воплощениях химерный рецепторный полипептид содержит GPCR либо его произ-

водное, вариант или фрагмент, выбранный из орфанных рецепторов класса А; орфанов класса В; орфанов класса С; вкусовых рецепторов 1 типа; вкусовых рецепторов 2 типа; 5-гидрокситриптаминовых рецепторов; ацетилхолиновых рецепторов (мускариновых); аденозиновых рецепторов; GPCRs адгезионного класса; адренорецепторов, ангиотензиновых рецепторов; апелиновых рецепторов; рецепторов желчных кислот; бомбезиновых рецепторов; брадикининовых рецепторов; кальцитониновых рецепторов; кальций-чувствительных рецепторов; рецепторов каннабиноидов; хемериновых рецепторов; хемокиновых рецепторов; холецистокининовых рецепторов; GPCRs класса Frizzled (например, рецепторов Wnt); рецепторов пептидов комплемента; рецепторов кортикотропин-рилизинг-факторов; дофаминовых рецепторов; эндотелиновых рецепторов; сопряженных G-белком эстрогеновых рецепторов; рецепторов формиловых пептидов; рецепторов свободных жирных кислот; ГАМК_B-рецепторов; галаниновых рецепторов; грелиновых рецепторов; семейства глюкогаоновых рецепторов; рецепторов гликопротеидных гормонов; рецепторов гонадотропин-рилизинг-гормонов; GPR18, GPR55 и GPR119; гистаминовых рецепторов; рецепторов гидроксикарбоновых кислот; рецепторов кисспептина; лейкотриеновых рецепторов; рецепторов лизофосфолипидов (LPA); рецепторов лизофосфолипидов (S1P); рецепторов меланин-концентрирующего гормона; рецепторов меланокортина; мелатониновых рецепторов; метаболитных глутаматных рецепторов; мотилиновых рецепторов; рецепторов нейромедина U; рецепторов нейропептида FF/нейропептида AF; рецепторов нейропептида S; рецепторов нейропептида W/нейропептида B; рецепторов нейропептида Y; рецепторов нейротензина; опиоидных рецепторов; рецепторов орексина; рецепторов оксоглутарата; P2Y-рецепторов; рецепторов паратиреоидного гормона; рецепторов фактора активации тромбоцитов; рецепторов прокинетицина; рецепторов пролактин-рилизинг-пептида; простанойдных рецепторов; активируемых протеиназой рецепторов; рецепторов QRFP; рецепторов пептидов семейства релаксина; соматоста-тиновых рецепторов; рецепторов сукцината; тахикининовых рецепторов; рецепторов тиротропин-рилизинг-гормона; рецепторов "следовых" аминов; рецепторов уротензина; рецепторов вазопрессина и окситоцина; рецепторов VIP и PACAP.

В некоторых воплощениях химерный рецепторный полипептид содержит GPCR, выбранный из группы, в которую входят: 5-гидрокситриптаминовый (серотониновый) рецептор 1A (HTR1A), 5-гидрокситриптаминовый (серотониновый) рецептор 1B (HTR1B), 5-гидрокситриптаминовый (серотониновый) рецептор 1D (HTR1D), 5-гидрокситриптаминовый (серотониновый) рецептор 1E (HTR1E), 5-гидрокситриптаминовый (серотониновый) рецептор 1F (HTR1F), 5-гидрокситриптаминовый (серотониновый) рецептор 2A (HTR2A), 5-гидрокситриптаминовый (серотониновый) рецептор 2B (HTR2B), 5-гидрокситриптаминовый (серотониновый) рецептор 2C (HTR2C), 5-гидрокситриптаминовый (серотониновый) рецептор 4 (HTR4), 5-гидрокситриптаминовый (серотониновый) рецептор 5A (HTR5A), 5-гидрокситриптаминовый (серотониновый) рецептор 5B (HTR5BP), 5-гидрокситриптаминовый (серотониновый) рецептор 6 (HTR6), 5-гидрокситриптаминовый (серотониновый) рецептор 7, сопряженный с аденилатциклазой (HTR7); холинергический мускариновый рецептор 1 (CHRM1), холинергический мускариновый рецептор 2 (CHRM2), холинергический мускариновый рецептор 3 (CHRM3), холинергический мускариновый рецептор 4 (CHRM4), холинергический мускариновый рецептор 5 (CHRM5); аденозиновый рецептор A1 (ADORA1), аденозиновый рецептор A2a (ADORA2A), аденозиновый рецептор A2b (ADORA2B), аденозиновый рецептор A3 (ADORA3); адгезионный сопряженный с G-белком рецептор A1 (ADGRA1), адгезионный сопряженный с G-белком рецептор A2 (ADGRA2), адгезионный сопряженный с G-белком рецептор A3 (ADGRA3), адгезионный сопряженный с G-белком рецептор B1 (ADGRB1), адгезионный сопряженный с G-белком рецептор B2 (ADGRB2), адгезионный сопряженный с G-белком рецептор B3 (ADGRB3); EGF-LAG-кадгериновый семипролетный рецептор G-типа 1 (CELSR1), EGF-LAG-кадгериновый семипролетный рецептор G-типа 2 (CELSR2), EGF-LAG-кадгериновый семипролетный рецептор G-типа 3 (CELSR3); адгезионный сопряженный с G-белком рецептор D1 (ADGRD1), адгезионный сопряженный с G-белком рецептор D2 (ADGRD2), адгезионный сопряженный с G-белком рецептор E1 (ADGRE1), адгезионный сопряженный с G-белком рецептор E2 (ADGRE2), адгезионный сопряженный с G-белком рецептор E3 (ADGRE3), адгезионный сопряженный с G-белком рецептор E4 (ADGRE4P), адгезионный сопряженный G-белком рецептор E5 (ADGRE5), адгезионный сопряженный с G-белком рецептор F1 (ADGRF1), адгезионный сопряженный с G-белком рецептор F2 (ADGRF2), адгезионный сопряженный с G-белком рецептор F3 (ADGRF3), адгезионный сопряженный с G-белком рецептор F4 (ADGRF4), адгезионный сопряженный с G-белком рецептор F5 (ADGRF5), адгезионный сопряженный с G-белком рецептор G1 (ADGRG1), адгезионный сопряженный G-белком рецептор G2 (ADGRG2), адгезионный сопряженный с G-белком рецептор G3 (ADGRG3), адгезионный сопряженный с G-белком рецептор G4 (ADGRG4), адгезионный сопряженный с G-белком рецептор G5 (ADGRG5), адгезионный сопряженный с G-белком рецептор G6 (ADGRG6), адгезионный сопряженный с G-белком рецептор G7 (ADGRG7), адгезионный сопряженный с G-белком рецептор L1 (ADGRL1), адгезионный сопряженный G-белком рецептор L2 (ADGRL2), адгезионный сопряженный с G-белком рецептор L3 (ADGRL3), адгезионный сопряженный с G-белком рецептор L4 (ADGRL4), адгезионный сопряженный с G-белком рецептор V (ADGRV1); адренорецептор α 1A (ADRA1A), адренорецептор α 1B (ADRA1B), адренорецептор α 1D (ADRA1D), адренорецептор α 2A (ADRA2A), адренорецептор α 2B (ADRA2B), адрено-

рецептор $\alpha 2C$ (ADRA2C), адренорецептор $\beta 1$ (ADRB1), адренорецептор $\beta 2$ (ADRB2), адренорецептор $\beta 3$ (ADRB3); рецептор ангиотензина-II 1-го типа (AGTR1), рецептор ангиотензина-II 2-го типа (AGTR2); апелиновый рецептор (APLNR); сопряженный с G-белком рецептор желчных кислот 1 (GPBAR1), рецептор нейромедина В (NMBR), рецептор гастрин-рилизинг-пептида (GRPR), рецептор-3 пептидов типа бомбезина (BRS3), брадикининовый рецептор B1 (BDKRB1), брадикининовый рецептор B2 (BDKRB2), кальцитониновый рецептор (CALCR), рецептор типа кальцитониновых рецепторов (CALCRL), кальций-чувствительный рецептор (CASR), сопряженный с G-белком рецептор класса C (GPRC6A), рецептор каннабиноидов 1 (мозг) (CNR1), рецептор каннабиноидов 2 (CNR2), хемериновый рецептор 1 типа хемокиновых рецепторов (CMKLR1), рецептор 1 хемокинов с мотивом C-C (CCR1), рецептор 2 хемокинов с мотивом C-C (CCR2), рецептор 3 хемокинов с мотивом C-C (CCR3), рецептор 4 хемокинов с мотивом C-C (CCR4), рецептор 5 хемокинов с мотивом C-C (ген/псевдоген) (CCR5), рецептор 6 хемокинов с мотивом C-C (CCR6), рецептор 7 хемокинов с мотивом C-C (CCR7), рецептор 8 хемокинов с мотивом C-C (CCR8), рецептор 9 хемокинов с мотивом C-C (CCR9), рецептор 10 хемокинов с мотивом C-C (CCR10); рецептор 1 хемокинов с мотивом C-X-C (CXCR1), рецептор 2 хемокинов с мотивом C-X-C (CXCR2), рецептор 3 хемокинов с мотивом C-X-C (CXCR3), рецептор 4 хемокинов с мотивом C-X-C (CXCR4), рецептор 5 хемокинов с мотивом C-X-C (CXCR5), рецептор 6 хемокинов с мотивом C-X-C (CXCR6), рецептор 1 хемокинов с мотивом C-X3-C (CX3CR1), рецептор 1 хемокинов с мотивом C (XCR1), атипичный рецептор хемокинов 1 (группы крови Duffy) (ACKR1), атипичный рецептор хемокинов 2 (ACKR2), атипичный рецептор хемокинов 3 (ACKR3), атипичный рецептор хемокинов 4 (ACKR4), рецептор типа рецептора 2 хемокинов с мотивом C-C (CCRL2); рецептор холецистокинина А (CCKAR), рецептор холецистокинина В (CCKBR), сопряженный с G-белком рецептор 1 (GPR1), рецептор-3 пептидов типа бомбезина (BRS3), сопряженный с G-белком рецептор 3 (GPR3), сопряженный с G-белком рецептор 4 (GPR4), сопряженный с G-белком рецептор 6 (GPR6), сопряженный с G-белком рецептор 12 (GPR12), сопряженный с G-белком рецептор 15 (GPR15), сопряженный с G-белком рецептор 17 (GPR17), сопряженный с G-белком рецептор 18 (GPR18), сопряженный с G-белком рецептор 19 (GPR19), сопряженный с G-белком рецептор 20 (GPR20), сопряженный с G-белком рецептор 21 (GPR21), сопряженный с G-белком рецептор 22 (GPR22), сопряженный с G-белком рецептор 25 (GPR25), сопряженный с G-белком рецептор 26 (GPR26), сопряженный с G-белком рецептор 27 (GPR27), сопряженный с G-белком рецептор 31 (GPR31), сопряженный с G-белком рецептор 32 (GPR32), сопряженный с G-белком рецептор 33 (ген/псевдоген) (GPR33), сопряженный с G-белком рецептор 34 (GPR34), сопряженный с G-белком рецептор 35 (GPR35), сопряженный с G-белком рецептор 37 (типа рецептора эндотелина типа В) (GPR37), сопряженный с G-белком рецептор-1 типа GPR37 (GPR37L1), сопряженный с G-белком рецептор 39 (GPR39), сопряженный с G-белком рецептор 42 (ген/псевдоген) (GPR42), сопряженный с G-белком рецептор 45 (GPR45), сопряженный с G-белком рецептор 50 (GPR50), сопряженный с G-белком рецептор 52 (GPR52), сопряженный с G-белком рецептор (GPR55), сопряженный с G-белком рецептор 61 (GPR61), сопряженный с G-белком рецептор 62 (GPR62), сопряженный с G-белком рецептор 63 (GPR63), сопряженный с G-белком рецептор 65 (GPR65), сопряженный с G-белком рецептор 68 (GPR68), сопряженный с G-белком рецептор 75 (GPR75), сопряженный с G-белком рецептор 78 (GPR78), сопряженный с G-белком рецептор 79 (GPR79), сопряженный с G-белком рецептор 82 (GPR82), сопряженный с G-белком рецептор 83 (GPR83), сопряженный с G-белком рецептор 84 (GPR84), сопряженный с G-белком рецептор 85 (GPR85), сопряженный с G-белком рецептор 87 (GPR87), сопряженный с G-белком рецептор 88 (GPR88), сопряженный с G-белком рецептор 101 (GPR101), сопряженный с G-белком рецептор 119 (GPR119), сопряженный с G-белком рецептор 132 (GPR132), сопряженный с G-белком рецептор 135 (GPR135), сопряженный с G-белком рецептор 139 (GPR139), сопряженный с G-белком рецептор 141 (GPR141), сопряженный с G-белком рецептор 142 (GPR142), сопряженный с G-белком рецептор 146 (GPR146), сопряженный с G-белком рецептор 148 (GPR148), сопряженный с G-белком рецептор 149 (GPR149), сопряженный с G-белком рецептор 150 (GPR150), сопряженный с G-белком рецептор 151 (GPR151), сопряженный с G-белком рецептор 152 (GPR152), сопряженный с G-белком рецептор 153 (GPR153), сопряженный с G-белком рецептор 160 (GPR160), сопряженный с G-белком рецептор 161 (GPR161), сопряженный с G-белком рецептор 162 (GPR162), сопряженный с G-белком рецептор 171 (GPR171), сопряженный с G-белком рецептор 173 (GPR173), сопряженный с G-белком рецептор 174 (GPR174), сопряженный с G-белком рецептор 176 (GPR176), сопряженный с G-белком рецептор 182 (GPR182), сопряженный с G-белком рецептор 183 (GPR183), содержащий богатый лейцином повтор сопряженный с G-белком рецептор 4 (LGR4), содержащий богатый лейцином повтор сопряженный с G-белком рецептор 5 (LGR5), содержащий богатый лейцином повтор сопряженный с G-белком рецептор 6 (LGR6); протоонкоген MAS1 (MAS1), протоонкоген типа MAS1 (MAS1L), представитель D семейства родственных MAS GPR (MRGPRD), представитель E семейства родственных MAS GPR (MRGPRE), представитель F семейства родственных MAS GPR (MRGPRF), представитель G семейства родственных MAS GPR (MRGPRG), представитель X1 семейства родственных MAS GPR (MRGPRX1), представитель X2 семейства родственных MAS GPR (MRGPRX2), представитель X3 семейства родственных MAS GPR (MRGPRX3), представитель X4 семейства родственных MAS GPR (MRGPRX4); опсин-3 (OPN3), опсин-4 (OPN4), опсин-5 (OPN5); пуриnergический рецептор P2Y (P2RY8), пуриnergический рецептор P2Y (P2RY10), связанный

со следовыми аминами рецептор 2 (TAAR2), связанный со следовыми аминами рецептор 3 (ген/псевдоген) (TAAR3), связанный со следовыми аминами рецептор 4 (TAAR4P), связанный со следовыми аминами рецептор 5 (TAAR5), связанный со следовыми аминами рецептор 6 (TAAR6), связанный со следовыми аминами рецептор 8 (TAAR8), связанный со следовыми аминами рецептор 9 (ген/псевдоген) (TAAR9); сопряженный с G-белком рецептор 156 (GPR156), сопряженный с G-белком рецептор 158 (GPR158), сопряженный с G-белком рецептор 179 (GPR179), сопряженный с G-белком рецептор класса C (GPRC5A), сопряженный с G-белком рецептор класса C (GPRC5B), сопряженный с G-белком рецептор класса C (GPRC5C), сопряженный с G-белком рецептор класса C (GPRC5D), рецептор-1 класса Frizzled (FZD1), рецептор-2 класса Frizzled (FZD2), рецептор-3 класса Frizzled (FZD3), рецептор-4 класса Frizzled (FZD4), рецептор-5 класса Frizzled (FZD5), рецептор-6 класса Frizzled (FZD6), рецептор-7 класса Frizzled (FZD7), рецептор-8 класса Frizzled (FZD8), рецептор-9 класса Frizzled (FZD9), рецептор-10 класса Frizzled (FZD10), рецептор smoothed класса Frizzled (SMO); рецептор-1 компонента компонента 3а (C3AR1), рецептор-1 компонента компонента 5а (C5AR1), рецептор-2 компонента компонента 5а (C5AR2); рецептор-1 кортикотропин-рилизинг-гормона (CRHR1), рецептор-2 кортикотропин-рилизинг-гормона (CRHR2); дофаминовый рецептор D1 (DRD1), дофаминовый рецептор D2 (DRD2), дофаминовый рецептор D3 (DRD3), дофаминовый рецептор D4 (DRD4), дофаминовый рецептор D5 (DRD5); эндотелиновый рецептор типа A (EDNRA), эндотелиновый рецептор типа B (EDNRB), рецептор формиловых пептидов 1 (FPR1), рецептор формиловых пептидов 2 (FPR2), рецептор формиловых пептидов 3 (FPR3), рецептор свободных жирных кислот 1 (FFAR1), рецептор свободных жирных кислот 2 (FFAR2), рецептор свободных жирных кислот 3 (FFAR3), рецептор свободных жирных кислот 4 (FFAR4), сопряженный с G-белком рецептор 42 (ген/псевдоген) (GPR42), рецептор гамма-аминомасляной кислоты (ГАМК) B1 (GABBR1), рецептор гамма-аминомасляной кислоты (ГАМК) B2 (GABBR2), галаниновый рецептор 1 (GALR1), галаниновый рецептор 2 (GALR2), галаниновый рецептор 3 (GALR3), рецептор стимулятора секреции гормона роста (GHSR), рецептор рилизинг-фактора гормона роста (GHRHR), рецептор гастроингибиторного полипептида (GIPR), рецептор глюкагоноподобного пептида 1 (GLP1R), рецептор глюкагоноподобного пептида 2 (GLP2R), глюкагоновый рецептор (GCGR), рецептор секретина (SCTR), рецептор фолликулостимулирующего гормона (FSHR), рецептор лютеинизирующего гормона/ хориогонадотропина (LHCGR), рецептор тиреостимулирующего гормона (TSHR), рецептор гонадотропин-рилизинг-гормона (GNRHR), рецептор гонадотропин-рилизинг-гормона 2 (псевдоген) (GNRHR2), сопряженный с G-белком рецептор 18 (GPR18), сопряженный с G-белком рецептор 55 (GPR55), сопряженный с G-белком рецептор 119 (GPR119), сопряженный с G-белком эстрогеновый рецептор 1 (GPER1), гистаминовый рецептор H1 (HRH1), гистаминовый рецептор H2 (HRH2), гистаминовый рецептор H3 (HRH3), гистаминовый рецептор H4 (HRH4), рецептор гидроксикарбоновых кислот 1 (HCAR1), рецептор гидроксикарбоновых кислот 2 (HCAR2), рецептор гидроксикарбоновых кислот 3 (HCAR3), рецептор KISS1 (KISS1R), рецептор лейкотриена B4 (LTB4R), рецептор-2 лейкотриена B4 (LTB4R2), рецептор-1 цистеиниловых лейкотриенов (CYSLTR1), рецептор-2 цистеиниловых лейкотриенов (CYSLTR2), рецептор оксозикозаноидов (OXE) 1 (OXER1), рецептор-2 формиловых пептидов (FPR2), рецептор лизофосфатидной кислоты 1 (LPAR1), рецептор лизофосфатидной кислоты 2 (LPAR2), рецептор лизофосфатидной кислоты 3 (LPAR3), рецептор лизофосфатидной кислоты 4 (LPAR4), рецептор лизофосфатидной кислоты 5 (LPAR5), рецептор лизофосфатидной кислоты 6 (LPAR6), рецептор-1 сфингозин-1-фосфата 1 (S1PR1), рецептор-2 сфингозин-1-фосфата (S1PR2), рецептор-3 сфингозин-1-фосфата 3 (S1PR3), рецептор-4 сфингозин-1-фосфата (S1PR4), рецептор-5 сфингозин-1-фосфата (S1PR5), рецептор-1 меланин-концентрирующего гормона (MCHR1), рецептор-2 меланин-концентрирующего гормона (MCHR2), рецептор меланокортина-1 (рецептор α -меланоцитстимулирующего гормона) (MC1R), рецептор меланокортина-2 (адренкортикотропного гормона) (MC2R), рецептор меланокортина-3 (MC3R), рецептор меланокортина-4 (MC4R), рецептор меланокортина-5 (MC5R), мелатониновый рецептор 1A (MTNR1A), мелатониновый рецептор 1B (MTNR1B), метаботропный глутаматный рецептор 1 (GRM1), метаботропный глутаматный рецептор 2 (GRM2), метаботропный глутаматный рецептор 3 (GRM3), метаботропный глутаматный рецептор 4 (GRM4), метаботропный глутаматный рецептор 5 (GRM5), метаботропный глутаматный рецептор 6 (GRM6), метаботропный глутаматный рецептор 7 (GRM7), метаботропный глутаматный рецептор 8 (GRM8), рецептор мотилина (MLNR), рецептор-1 нейромедина U (NMUR1), рецептор-2 нейромедина U (NMUR2), рецептор-1 нейропептида FF (NPFFR1), рецептор-2 нейропептида FF (NPFFR2), рецептор-1 нейропептида S (NPSR1), рецептор-1 нейропептидов B/W (NPBWR1), рецептор-2 нейропептидов B/W (NPBWR2), рецептор Y1 нейропептида Y (NPY1R), рецептор Y2 нейропептида Y (NPY2R), рецептор Y4 нейропептида Y (NPY4R), рецептор Y5 нейропептида Y (NPY5R), рецептор Y6 нейропептида Y (псевдо-ген) (NPY6R), рецептор-1 нейротензина (с высоким сродством) (NTSR1), рецептор-2 нейротензина (NTSR2), опиоидный рецептор дельта-1 (OPRD1), опиоидный рецептор каппа-1 (OPRK1), опиоидный рецептор мю-1 (OPRM1), рецептор-1 типа опиатных рецепторов (OPRL1), рецептор-1 гипокретина (орексина) (HCRTR1), рецептор-2 гипокретина (орексина) (HCRTR2), сопряженный с G-белком рецептор 107 (GPR107), сопряженный с G-белком рецептор 137 (GPR137), 1-й обонятельный рецептор подсемейства E семейства 51 (OR51E1), связанный с адипоцитами трансмембранный белок 1 (TPRA1), сопряженный с G-белком рецептор 143 (GPR143), сопряженный с G-

белком рецептор 157 (GPR157), рецептор оксоглутарата (α -кетоглутарата) 1 (OXGR1), пуринергический рецептор P2Y (P2RY1), пуринергический рецептор P2Y (P2RY2), пиримидинергический рецептор P2Y (P2RY4), пиримидинергический рецептор P2Y (P2RY6), пуринергический рецептор P2Y (P2RY11), пуринергический рецептор P2Y (P2RY12), пуринергический рецептор P2Y (P2RY13), пуринергический рецептор P2Y (P2RY14), рецептор паратиреоидного гормона 1 (PTH1R), рецептор паратиреоидного гормона 2 (PTH2R), рецептор фактора активации тромбоцитов (PTAFR), рецептор-1 прокинетицина (PROKR1), рецептор-2 прокинетицина (PROKR2), рецептор пролактин-рилизинг-гормона (PRLHR), рецептор простагландина D2 (DP) (PTGDR), рецептор-2 простагландина D2 (PTGDR2), рецептор-1 простагландина E (PTGER1), рецептор-2 простагландина E (PTGER2), рецептор-3 простагландина E (PTGER3), рецептор-4 простагландина E (PTGER4), рецептор простагландина F (PTGFR), рецептор простагландина I2 (простаглицина) (IP) (PTGIR), рецептор тромбосана A2 (TBXA2R), рецептор фактора свёртывания II тромбина (F2R), рецептор-1 трипсина типа F2R (F2RL1), рецептор-2 типа рецептора фактора свёртывания II тромбина (F2RL2), рецептор-3 тромбина/трипсина типа F2R (F2RL3), рецептор пироглутамилированных RF-амидных пептидов (QRFP), рецептор-1 пептидов семейства релаксина/подобных инсулину (RXFP1), рецептор-2 пептидов семейства релаксина/подобных инсулину (RXFP2), рецептор-3 пептидов семейства релаксина/подобных инсулину (RXFP3), рецептор-4 пептидов семейства релаксина/подобных инсулину (RXFP4), соматостатиновый рецептор 1 (SSTR1), соматостатиновый рецептор 2 (SSTR2), соматостатиновый рецептор 3 (SSTR3), соматостатиновый рецептор 4 (SSTR4), соматостатиновый рецептор 5 (SSTR5), рецептор сукцината 1 (SUCNR1), тахикининовый рецептор 1 (TACR1), тахикининовый рецептор 2 (TACR2), тахикининовый рецептор 3 (TACR3), 1-й вкусовой рецептор 1 типа (TAS1R1), 2-й вкусовой рецептор 1 типа (TAS1R2), 3-й вкусовой рецептор 1 типа (TAS1R3), 1-й вкусовой рецептор 2 типа (TAS2R1), 3-й вкусовой рецептор 2 типа (TAS2R3), 4-й вкусовой рецептор 2 типа (TAS2R4), 5-й вкусовой рецептор 2 типа (TAS2R5), 7-й вкусовой рецептор 2 типа (TAS2R7), 8-й вкусовой рецептор 2 типа (TAS2R8), 9-й вкусовой рецептор 2 типа (TAS2R9), 10-й вкусовой рецептор 2 типа (TAS2R10), 13-й вкусовой рецептор 2 типа (TAS2R13), 14-й вкусовой рецептор 2 типа (TAS2R14), 16-й вкусовой рецептор 2 типа (TAS2R16), 19-й вкусовой рецептор 2 типа (TAS2R19), 20-й вкусовой рецептор 2 типа (TAS2R20), 30-й вкусовой рецептор 2 типа (TAS2R30), 31-й вкусовой рецептор 2 типа (TAS2R31), 38-й вкусовой рецептор 2 типа (TAS2R38), 39-й вкусовой рецептор 2 типа (TAS2R39), 40-й вкусовой рецептор 2 типа (TAS2R40), 41-й вкусовой рецептор 2 типа (TAS2R41), 42-й вкусовой рецептор 2 типа (TAS2R42), 43-й вкусовой рецептор 2 типа (TAS2R43), 45-й вкусовой рецептор 2 типа (TAS2R45), 46-й вкусовой рецептор 2 типа (TAS2R46), 50-й вкусовой рецептор 2 типа (TAS2R50), 60-й вкусовой рецептор 2 типа (TAS2R60), рецептор тиротропин-рилизинг-гормона (TRHR), связанный со следовыми аминами рецептор 1 (TAAR1), рецептор уротензина 2 (UTS2R), рецептор аргинин-вазопрессина 1A (AVPR1A), рецептор аргинин-вазопрессина 1B (AVP R1B), рецептор аргинин-вазопрессина 2 (AVPR2), окситоциновый рецептор (OXTR), рецептор активирующего аденилатциклазу полипептида 1 (типофизарного) I типа (ADCYAP1R1), рецептор-1 вазоактивного кишечного пептида (VIPR1), рецептор-2 вазоактивного кишечного пептида 2 (VIPR2), их производные, варианты и фрагменты.

Химерный рецепторный полипептид, содержащий GPCR либо его производное, вариант или фрагмент, может связывать антиген, включающий любой подходящий лиганд GPCR или его производное, вариант или фрагмент. Неограничительные примеры лигандов, которые могут связываться с GPCR, включают (-)-адреналин, (-)-норадреналин, (лизо)фосфолипидные медиаторы, [des-Arg10]каллидин, [des-Arg9]брадикинин, [des-Gln14]грелин, [Hyp3]брадикинин, [Leu]энкефалин, [Met]энкефалин, 12-гидроксигепта-декатриеновую кислоту, 12R-НЕТЕ, 12S-НЕТЕ, 12S-НРЕТЕ, 15S-НЕТЕ, 17 β -эстрадиол, 20-гидрокси-ЛТВ4, 2-арахидоноилглицерин, 2-олеоил-LPA, 3-гидроксиоктановую кислоту, 5-гидрокси-триптамин, 5-оксо-15-НЕТЕ, 5-оксо-ЕТЕ, 5-оксо-ЕТРе, 5-оксо-ОДЕ, 5S-НЕТЕ, 5S-НРЕТЕ, 7 α ,25-дигидроксихолестерин, ацетилхолин, АСТН, аденозиндифосфат, аденозин, аденомедуллин-2/интермедиин, аденомедуллин, амилин, анандамид, ангиотензин II, ангиотензин III, аннексин I, ранний эндогенный лиганд рецептора апелина, апелин-13, апелин-17, апелин-36, индуцируемый аспирином липоксин A4, индуцируемый аспирином резолвин D1, АТФ, β -дефенсин 4А, большой динорфин, пептид 8-22 мозгового вещества надпочечников быка, брадикинин, C3a, C5a, Ca²⁺, родственный гену кальцитонина пептид, кальцитонин, катепсин G, ССК-33, ССК-4, ССК-8, CCL1, CCL11, CCL13, CCL14, CCL15, CCL16, CCL17, CCL19, CCL2, CCL20, CCL21, CCL22, CCL23, CCL24, CCL25, CCL26, CCL27, CCL28, CCL3, CCL4, CCL5, CCL7, CCL8, хемерин, хенодесоксихолевую кислоту, холевую кислоту, кортикотропин-рилизинг-гормон, CST-17, CX3CL1, CXCL1, CXCL10, CXCL11, CXCL12 α , CXCL12 β , CXCL13, CXCL16, CXCL2, CXCL3, CXCL5, CXCL6, CXCL7, CXCL8, CXCL9, цистеиниловые лейкотриены (CysLTs), урациловые нуклеотиды, дезоксихолевую кислоту, дигидросфингозин-1-фосфат, диолеилфосфатидную кислоту, дофамин, динорфин А, динорфин А-(1-13), динорфин А-(1-8), динорфин В, эндоморфин-1, эндотелин-1, эндотелин-2, эндотелин-3, F2L, свободные жирные кислоты, FSH, ГАМК, галанин, галанинподобный пептид, гастроингибиторный полипептид, гастрин-17, гастрин-рилизинг-пептид, грелин, GHRH, глюкагон, глюкагоноподобный пептид 1-(7-36)-амид, глюкагоноподобный пептид 1-(7-37), глюкагоно-

подобный 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, длинноцепочечные карбоновые кислоты, LPA, LTB4, LTC4, LTD4, LTE4, LXA4, Lys-[Нур3]-брадикинин, лизофосфатидилинозитол, лизофосфатидилсерин, среднецепочечные жирные кислоты, 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-олеилэтаноламид, обестатин, октопамин, орексин-А, орексин-В, окси-стеролы, окситоцин, PACAP-27, PACAP-38, PAF, панкреатический полипептид, 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, SRJF-14, SRIF-28, субстанцию P, янтарную кислоту, тромбин, тромбоксан A2, TIRP39, 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-гидроксимасляную кислоту, β -эндофин, β -MSH, β -неоэндофин, β -фенилэтиламин и γ -MSH.

В некоторых воплощениях химерный рецепторный полипептид содержит интегриновый рецептор, субъединицу интегринового рецептора или его производное, вариант или фрагмент. Интегриновые рецепторы представляют собой трансмембранные рецепторы, которые могут функционировать в качестве мостиков для взаимодействий типа клетка-клетка и клетка-внеклеточный матрикс (ECM). Интегриновые рецепторы обычно образуются в виде гетеродимеров, состоящих из α -субъединицы и β -субъединицы, которые связаны нековалентно. Существует не менее 18 α -субъединиц и не менее 8 β -субъединиц. Каждая субъединица обычно содержит внеклеточную область (напр, лиганд-связывающий домен), пронизывающий мембрану участок и внутриклеточную область (напр, цитоплазматический домен). В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внеклеточную область (например, лиганд-связывающий домен) субъединицы интегрин (например, α -субъединицы или β -субъединицы) либо её производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, пронизывающий мембрану участок субъединицы интегрин (напр, α -субъединицы или β -субъединицы) либо его производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внутриклеточную область (напр, цитоплазматический домен) субъединицы интегрин (напр, α -субъединицы или β -субъединицы) либо её производное, вариант или фрагмент. Химерный рецепторный полипептид, содержащий субъединицу интегрин либо её производное, вариант или фрагмент, может рекрутировать партнера по связыванию. В некоторых воплощениях связывание лиганда с химерным рецептором, содержащим субъединицу интегрин либо её производное, вариант или фрагмент, вызывает конформационное изменение, химическую модификацию либо их комбинацию, которые рекрутируют партнера по связыванию к рецептору.

В некоторых воплощениях химерный рецепторный полипептид содержит α -субъединицу интегринового рецептора либо её производное, вариант или фрагмент, выбранную из группы, состоящей из $\alpha 1$, $\alpha 2$, $\alpha 3$, $\alpha 4$, $\alpha 5$, $\alpha 6$, $\alpha 7$, $\alpha 8$, $\alpha 9$, $\alpha 10$, $\alpha 11$, αV , αL , αM , αX , αD , αE и αIIb . В некоторых воплощениях химерный рецепторный полипептид содержит β -субъединицу интегринового рецептора либо её производное, вариант или фрагмент, выбранную из группы, состоящей из $\beta 1$, $\beta 2$, $\beta 3$, $\beta 4$, $\beta 5$, $\beta 6$, $\beta 7$ и $\beta 8$. Химерные рецепторные полипептиды, содержащие α -субъединицу, β -субъединицу либо её производное, вариант или фрагмент, могут образовывать гетеродимеры (например, α -субъединица димеризуется с β -субъединицей) с образованием интегринового рецептора либо его производного, варианта или фрагмента. Неограничительные примеры интегриновых рецепторов включают рецепторы $\alpha 1\beta 1$, $\alpha 2\beta 1$, $\alpha 5\beta 1$, $\alpha 6\beta 1$, $\alpha 7\beta 1$, $\alpha 8\beta 1$, $\alpha 9\beta 1$, $\alpha 10\beta 1$, $\alpha V\beta 1$, $\alpha L\beta 1$, $\alpha M\beta 1$, $\alpha X\beta 1$, $\alpha D\beta 1$, $\alpha IIb\beta 1$, $\alpha E\beta 1$, $\alpha 1\beta 2$, $\alpha 2\beta 2$, $\alpha 3\beta 2$, $\alpha 4\beta 2$, $\alpha 5\beta 2$, $\alpha 6\beta 2$, $\alpha 7\beta 2$, $\alpha 8\beta 2$, $\alpha 9\beta 2$, $\alpha 10\beta 2$, $\alpha V\beta 2$, $\alpha L\beta 2$, $\alpha M\beta 2$, $\alpha X\beta 2$, $\alpha D\beta 2$, $\alpha IIb\beta 2$, $\alpha E\beta 2$, $\alpha 1\beta 3$, $\alpha 2\beta 3$, $\alpha 3\beta 3$, $\alpha 4\beta 3$, $\alpha 5\beta 3$, $\alpha 6\beta 3$, $\alpha 7\beta 3$, $\alpha 8\beta 3$, $\alpha 9\beta 3$, $\alpha 10\beta 3$, $\alpha V\beta 3$, $\alpha L\beta 3$, $\alpha M\beta 3$, $\alpha X\beta 3$, $\alpha D\beta 3$, $\alpha IIb\beta 3$, $\alpha E\beta 3$, $\alpha 1\beta 4$, $\alpha 2\beta 4$, $\alpha 3\beta 4$, $\alpha 4\beta 4$, $\alpha 5\beta 4$, $\alpha 6\beta 4$, $\alpha 7\beta 4$, $\alpha 8\beta 4$, $\alpha 9\beta 4$, $\alpha 10\beta 4$, $\alpha V\beta 4$, $\alpha L\beta 4$, $\alpha M\beta 4$, $\alpha X\beta 4$, $\alpha D\beta 4$, $\alpha IIb\beta 4$, $\alpha E\beta 4$, $\alpha 1\beta 5$, $\alpha 2\beta 5$, $\alpha 3\beta 5$, $\alpha 4\beta 5$, $\alpha 5\beta 5$, $\alpha 6\beta 5$, $\alpha 7\beta 5$, $\alpha 8\beta 5$, $\alpha 9\beta 5$, $\alpha 10\beta 5$, $\alpha V\beta 5$, $\alpha L\beta 5$, $\alpha M\beta 5$, $\alpha X\beta 5$, $\alpha D\beta 5$, $\alpha IIb\beta 5$, $\alpha E\beta 5$, $\alpha 1\beta 6$, $\alpha 2\beta 6$, $\alpha 3\beta 6$, $\alpha 4\beta 6$, $\alpha 5\beta 6$, $\alpha 6\beta 6$, $\alpha 7\beta 6$, $\alpha 8\beta 6$, $\alpha 9\beta 6$, $\alpha 10\beta 6$, $\alpha V\beta 6$, $\alpha L\beta 6$, $\alpha M\beta 6$, $\alpha X\beta 6$, $\alpha D\beta 6$, $\alpha IIb\beta 6$, $\alpha E\beta 6$, $\alpha 1\beta 7$, $\alpha 2\beta 7$, $\alpha 3\beta 7$, $\alpha 4\beta 7$, $\alpha 5\beta 7$, $\alpha 6\beta 7$, $\alpha 7\beta 7$, $\alpha 8\beta 7$, $\alpha 9\beta 7$, $\alpha 10\beta 7$, $\alpha V\beta 7$, $\alpha L\beta 7$, $\alpha M\beta 7$, $\alpha X\beta 7$, $\alpha D\beta 7$, $\alpha IIb\beta 7$, $\alpha E\beta 7$, $\alpha 1\beta 8$, $\alpha 2\beta 8$, $\alpha 3\beta 8$, $\alpha 4\beta 8$, $\alpha 5\beta 8$, $\alpha 6\beta 8$, $\alpha 7\beta 8$, $\alpha 8\beta 8$, $\alpha 9\beta 8$, $\alpha 10\beta 8$, $\alpha V\beta 8$, $\alpha \beta \beta$, $\alpha M\beta 8$, $\alpha X\beta 8$, $\alpha D\beta 8$, $\alpha IIb\beta 8$ и $\alpha E\beta 8$. Химерный рецепторный полипептид, содержащий субъединицу интегрин либо её

производное, вариант или фрагмент, может димеризоваться с эндогенной субъединицей интегрин (например, субъединицей интегрин дикого типа).

Химерный рецепторный полипептид, содержащий субъединицу интегрин либо её производное, вариант или фрагмент, может связывать антиген, содержащий любой подходящий лиганд интегрин либо его производное, вариант или фрагмент. Неограничительные примеры лигандов, которые могут связываться с интегриновым рецептором, включают белок основания пентона аденовируса, β -глокан, костный сиалопроtein (BSP), *Borrelia burgdorferi*, *Candida albicans*, коллагены (CN, например, CNI-IV), цитотактин/тенасцин-С, декорсин, денатурированный коллаген, дезинтегрины, Е-кадгерин, рецептор эховирусов 1, эпигрин, фактор X, Fc ϵ RII (CD23), фибрин (Fb), фибриноген (Fg), фибронектин (Fn), гепарин, белок Tat ВИЧ, iC3b, молекулы межклеточной адгезии (например, ICAM-1,2,3,4,5), инвазин, молекулы клеточной адгезии L1 (L1-CAM), ламинин, липополисахарид (LPS), MAdCAM-1, матриксную металлопротеиназу-2 (ММРe), ингибирующий нейтрофилы фактор (NIF), остеопонтин (OP или OPN), плазминоген, протромбин, фертилин сперматозоидов, тромбоспондин (TSP), молекулы адгезии сосудистых клеток 1 (VCAM-1), витронектин (VN или VTN) и фактор фон Виллебранда (vWF).

В некоторых воплощениях химерный рецепторный полипептид содержит молекулу кадгерина либо его производное, вариант или фрагмент. Молекулы кадгерина, которые могут функционировать и как лиганды, и как рецепторы, относятся к определенным белкам, участвующим в осуществлении клеточной адгезии. Молекулы кадгерина обычно состоят из пяти тандемных повторов внеклеточных доменов, одного пронизывающего мембрану сегмента и цитоплазматической области. Например, Е-кадгерин или CDH1 состоит из 5 повторов внеклеточного домена, одного трансмембранного домена и одного внутриклеточного домена. При фосфорилировании Е-кадгерина на участке внутриклеточного домена адаптерные белки типа β -катенина и p120-катенина могут связываться с рецептором. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внеклеточную область кадгерина либо её производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, пронизывающий мембрану участок кадгерина либо его производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внутриклеточную область (например, цитоплазматический домен) кадгерина либо её производное, вариант или фрагмент. Химерный рецепторный полипептид, содержащий кадгерин либо его производное, вариант или фрагмент, может рекрутировать партнера по связыванию. В некоторых воплощениях связывание лиганда с химерным рецептором, содержащим кадгерин либо его производное, вариант или фрагмент, вызывает конформационное изменение, химическую модификацию либо их комбинацию, которые рекрутируют партнера по связыванию к рецептору.

Химерный рецепторный полипептид может содержать кадгерин либо его производное, вариант или фрагмент, выбранный из классического кадгерина, кадгерина десмосом, протокадгерина и нестандартного кадгерина. В некоторых воплощениях химерный рецепторный полипептид содержит классический кадгерин либо его производное, вариант или фрагмент, выбранный из CDH1 (Е-кадгерин, эпителиальный), CDH2 (N-кадгерин, невральный), CDH12 (кадгерин 12, 2-го типа, N-кадгерин 2) и CDH3 (P-кадгерин, плацентарный). В некоторых воплощениях химерный рецепторный полипептид содержит кадгерин десмосом либо его производное, вариант или фрагмент, выбранный из десмоглеина (DSG1, DSG2, DSG3, DSG4) и десмоколина (DSC1, DSC2, DSC3). В некоторых воплощениях химерный рецепторный полипептид содержит протокадгерин либо его производное, вариант или фрагмент, выбранный из PCDH1, PCDH10, PCDH11X, PCDH11Y, PCDH12, PCDH15, PCDH17, PCDH18, PCDH19, PCDH20, PCDH7, PCDH8, PCDH9, PCDHA1, PCDHA10, PCDHA11, PCDHA12, PCDHA13, PCDHA2, PCDHA3, PCDHA4, PCDHA5, PCDHA6, PCDHA7, PCDHA8, PCDHA9, PCDHAC1, PCDHAC2, PCDHB1, PCDHB10, PCDHB11, PCDHB12, PCDHB13, PCDHB14, PCDHB15, PCDHB16, PCDHB17, PCDHB18, PCDHB2, PCDHB3, PCDHB4, PCDHB5, PCDHB6, PCDHB7, PCDHB8, PCDHB9, PCDHGA1, PCDHGA10, PCDHGA11, PCDHGA12, PCDHGA2, PCDHGA3, PCDHGA4, PCDHGA5, PCDHGA6, PCDHGA7, PCDHGA8, PCDHGA9, PCDHGB1, PCDHGB2, PCDHGB3, PCDHGB4, PCDHGB5, PCDHGB6, PCDHGB7, PCDHGC3, PCDHGC4, PCDHGC5, FAT, FAT2 и FAT). В некоторых воплощениях химерный рецепторный полипептид содержит нестандартный кадгерин, выбранный из CDH4 (R-кадгерин, сетчатка), CDH5 (VE-кадгерин, сосудистый эндотелий), CDH6 (K-кадгерин, почки), CDH7 (кадгерин 7, 2-го типа), CDH8 (кадгерин 8, 2-го типа), CDH9 (кадгерин 9, 2-го типа, T1-кадгерин), CDH10 (кадгерин 10, 2-го типа, T2-кадгерин), CDH11 (OB-кадгерин, остеобласты), CDH13 (T-кадгерин, X-кадгерин, сердце), CDH15 (M-кадгерин, мышечные трубочки), CDH16 (KSP-кадгерин), CDH17 (LI-кадгерин, печень и кишечник), CDH18 (кадгерин 18, 2-го типа), CDH19 (кадгерин 19, 2-го типа), CDH20 (кадгерин 20, 2-го типа), CDH23 (кадгерин 23, нейросенсорный эпителий), CDH24, CDH26, CDH28, CELSR1, CELSR2, CELSR3, CLSTN1, CLSTN2, CLSTN3, DCHS1, DCHS2, LOC389118, PCLKC, RESDA1 и RET.

Химерный рецепторный полипептид, содержащий кадгерин либо его производное, вариант или фрагмент, может связывать антиген, включающий любой подходящий лиганд кадгерина либо его производное, вариант или фрагмент. Лиганд кадгерина может включать, к примеру, другой рецептор кадгерина (например, рецептор кадгерина в клетке).

В некоторых воплощениях химерный рецепторный полипептид содержит каталитический рецептор либо его производное, вариант или фрагмент. Примеры каталитических рецепторов включают, без ограничения, рецепторные тирозинкиназы (RTK) и рецепторные треониновые/сериновые киназы (RTSK). Каталитические рецепторы типа RTK и RTSK обладают определенной ферментативной активностью. Например, RTK могут фосфорилировать белки-субстраты по остаткам тирозина, которые затем могут действовать как сайты связывания для белков-адаптеров. RTK обычно содержат N-концевой внеклеточный лиганд-связывающий домен, одну трансмембранную α -спираль и цитозольный C-концевой домен с активностью протеин-тирозинкиназы. Некоторые RTK состоят из единственного полипептида, тогда как некоторые являются димерами, состоящими из двух пар полипептидных цепей, к примеру рецептора инсулина и некоторых родственных рецепторов. Связывание лигандов с внеклеточными доменами этих рецепторов может активировать цитозольные киназные домены, что приводит к фосфорилированию как самих рецепторов, так и внутриклеточных белков-мишеней, распространяющих сигнал, вызванный связыванием лиганда. У некоторых RTKs связывание лиганда индуцирует димеризацию рецептора. Некоторые лиганды (например, такие факторы роста, как PDGF и NGF) сами являются димерами, состоящими из двух идентичных полипептидных цепей. Эти факторы роста могут непосредственно индуцировать димеризацию при одновременном связывании с двумя различными молекулами рецептора. Другие факторы роста (например, типа EGF) являются мономерами, но имеют два разных сайта связывания рецептора, которые могут перекрестно сшивать рецепторы. Индуцированная лигандом димеризация может приводить к аутофосфорилированию рецептора, при этом димеризованные полипептидные цепи перекрестно фосфорилируют друг друга. Некоторые рецепторы могут мультимеризоваться.

В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внеклеточную область (например, лиганд-связывающий домен) каталитического рецептора типа RTK либо её производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, пронизывающий мембрану участок каталитического рецептора типа RTK либо его производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внутриклеточную область (например, цитозольный домен) каталитического рецептора типа RTK либо её производное, вариант или фрагмент. Химерный рецепторный полипептид, содержащий RTK либо её производное, вариант или фрагмент, может рекрутировать партнера по связыванию. В некоторых воплощениях связывание лиганда с химерным рецептором, содержащим RTK либо её производное, вариант или фрагмент, вызывает конформационное изменение, химическую модификацию либо их комбинацию, которые рекрутируют партнера по связыванию к рецептору.

В некоторых воплощениях химерный рецепторный полипептид содержит RTK класса I (например, из семейства рецепторов эпидермального фактора роста (EGF), включая EGFR, семейства ErbB, включая ErbB-2, ErbB-3 и ErbB-4), RTK класса II (например, из семейства инсулиновых рецепторов, включая INSR, IGF-1R и IRR), RTK класса III (например, из семейства рецепторов тромбоцитарного фактора роста (PDGF), включая PDGFR- α , PDGFR- β , CSF-1R, KIT/SCFR и FLK2/FLT3), RTK класса IV (например, из семейства рецепторов фактора роста фибробластов (FGF), включая FGFR-1, FGFR-2, FGFR-3 и FGFR-4), RTK класса V (например, из семейства рецепторов фактора роста сосудистого эндотелия (VEGF), включая VEGFR1, VEGFR2 и VEGFR3), RTK класса VI (например, из семейства рецепторов фактора роста гепатоцитов (HGF), включая рецептор фактора роста гепатоцитов (HGFR/MET) и RON), RTK класса VII (например, из семейства рецепторов типа рецепторной киназы тропомиозина (Trk), включая TRKA, TRKB и TRKC), RTK класса VIII (например, из семейства рецепторов эфрина (Eph), включая EphA1, EphA2, EphA3, EphA4, EphA5, EphA6, EphA7, EphA8, EphB1, EphB2, EphB3, EphB4, EphB5 и EphB6), RTK класса IX (например, из семейства рецепторов AXL типа AXL, MER и TRYO3), RTK класса X (например, из семейства рецепторов LTK типа LTK и ALK), RTK класса XI (например, из семейства рецепторов TIE типа TIE и TEK), RTK класса XII (например, ROR1 и ROR2 из семейства рецепторов ROR), RTX класса XIII (например, из семейства рецепторов с доменом дискоидина (DDR) типа DDR1 и DDR2), RTK класса XIV (например, из семейства рецепторов RET типа RET), RTK класса XV (например, из семейства рецепторов KLG, включая PTK7), RTK класса XVI (например, из семейства рецепторов RYK, включая Ryk), RTK класса XVII (например, из семейства рецепторов MuSK типа MuSK) либо её производное, вариант или фрагмент.

Химерный рецепторный полипептид, содержащий RTK либо её производное, вариант или фрагмент, может связывать антиген, включающий любой подходящий лиганд RTK либо его производное, вариант или фрагмент. Неограничительные примеры лигандов RTK включают факторы роста, цитокины и гормоны. Факторы роста включают, к примеру, представителей семейства эпидермального фактора роста (например, эпидермальный фактор роста или EGF, связывающий гепарин EGF-подобный фактор роста или HB-EGF, трансформирующий фактор роста- α или TGF- α , амфирегулин или AR, эпирегулин или EPR, эпиген, бетацеллюлин или BTC, нейрегулин-1 или NRG1, нейрегулин-2 или NRG2, нейрегулин-3 или NRG3 и нейрегулин-4 или NRG4), семейства факторов роста фибробластов (например, FGF1, FGF2, FGF3, FGF4, FGF5, FGF6, FGF7, FGF8, FGF9, FGF10, FGF11, FGF12, FGF13, FGF14, FGF15/19, FGF16, FGF17, FGF18, FGF20, FGF21 и FGF23), семейства факторов роста сосудистого эндотелия (на-

пример, VEGF-A, VEGF-B, VEGF-C, VEGF-D и PIGF) и семейства тромбоцитарных факторов роста (например, PDGFA, PDGFB, PDGFC и PDGFD). Гормоны включают, к примеру, представителей семейства инсулина/IGF/релаксина (например, инсулин, инсулиноподобные факторы роста, пептиды семейства релаксинов, включая релаксин-1, релаксин-2, релаксин-3, специфичный для клеток Лейдига инсулиноподобный пептид (ген INSL3), ранний инсулиноподобный плацентарный пептид (ELIP) (ген INSL4), инсулиноподобный пептид 5 (ген INSL5) и инсулиноподобный пептид 6).

В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внеклеточную область (например, лиганд-связывающий домен) каталитического рецептора типа RTSK либо её производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, пронизывающий мембрану участок каталитического рецептора типа RTSK либо его производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внутриклеточную область (например, цитозольный домен) каталитического рецептора типа RTSK либо её производное, вариант или фрагмент. Химерный рецепторный полипептид, содержащий RTSK либо её производное, вариант или фрагмент, может рекрутировать партнера по связыванию. В некоторых воплощениях связывание лиганда с химерным рецептором, содержащим RTSK либо её производное, вариант или фрагмент, вызывает конформационное изменение, химическую модификацию либо их комбинацию, которые рекрутируют партнера по связыванию к рецептору.

Химерный рецепторный полипептид, содержащий RTSK либо её производное, вариант или фрагмент, может фосфорилировать субстрат по остаткам серина и/или треонина и может предпочитать определенные остатки на основе консенсусной последовательности. Химерный рецепторный полипептид может содержать RTSK типа I, RTSK типа II либо её производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид, содержащий рецепторную сериновую/треониновую киназу I типа, неактивен, если он не образует комплекса с рецептором типа II. В некоторых воплощениях химерный рецепторный полипептид, содержащий рецепторную сериновую/треониновую киназу II типа, содержит конститутивно активный киназный домен, который может фосфорилировать и активировать рецептор I типа при образовании комплекса с рецептором I типа. Рецепторная сериновая/треониновая киназа II типа может фосфорилировать киназный домен партнера I типа, вызывая вытеснение белковых партнеров. Вытеснение белковых партнеров может позволить связывание и фосфорилирование других белков, например, некоторых представителей семейства SMAD. Химерный рецепторный полипептид может содержать рецептор I типа либо его производное, вариант или фрагмент, выбранный из группы, состоящей из: ALK1 (ACVRL1), ALK2 (ACVR1A), ALK3 (BMPR1A), ALK4 (ACVR1B), ALK5 (TGF β R1), ALK6 (BMPR1B) и ALK7 (ACVR1C). Химерный рецепторный полипептид может содержать рецептор II типа либо его производное, вариант или фрагмент, выбранный из группы, состоящей из: TGF β R2, BMPR2, ACVR2A, ACVR2B и AMHR2 (AMHR). В некоторых воплощениях химерный рецепторный полипептид содержит рецептор TGF- β либо его производное, вариант или фрагмент.

В некоторых воплощениях химерный рецепторный полипептид содержит рецептор, который стимулирует нековалентно связанные внутриклеточные киназы типа Src-киназы (например, c-Src, Yes, Fyn, Fgr, Lck, Hck, Blk, Lyn и Frk) или JAK-киназы (например, JAK1, JAK2, JAK3 и TYK2), но не обладает собственной ферментативной активностью, либо его производное, вариант или фрагмент. К таковым относится суперсемейство рецепторов цитокинов типа рецепторов цитокинов и полипептидных гормонов. Рецепторы цитокинов обычно содержат N-концевой внеклеточный лиганд-связывающий домен, трансмембранный α -спираль и C-концевой цитозольный домен. Цитозольные домены рецепторов цитокинов, как правило, лишены какой-либо известной каталитической активности. Вместо этого рецепторы цитокинов могут функционировать в ассоциации с нерецепторными киназами (например, тирозинкиназами или треониновыми/сериновыми киназами), которые могут активироваться при связывании лиганда с рецептором. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внеклеточную область (например, лиганд-связывающий домен) каталитического рецептора, нековалентно связывающегося с внутриклеточной киназой (например, рецептора цитокинов), либо её производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, пронизывающий мембрану участок каталитического рецептора, нековалентно связывающегося с внутриклеточной киназой (например, рецептора цитокинов), либо его производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внутриклеточную область (например, цитозольный домен) каталитического рецептора, нековалентно связывающегося с внутриклеточной киназой (например, рецептора цитокинов), либо её производное, вариант или фрагмент. Химерный рецепторный полипептид, содержащий каталитический рецептор, нековалентно связывающийся с внутриклеточной киназой, либо его производное, вариант или фрагмент, может рекрутировать партнера по связыванию. В некоторых воплощениях связывание лиганда с химерным рецептором, содержащим каталитический рецептор, нековалентно связывающийся с внутриклеточной киназой, либо его производное, вариант или фрагмент, вызывает конформационное изменение, химическую модификацию либо их комбинацию, которые рекрутируют партнера по связыванию к рецептору.

В некоторых воплощениях химерный рецепторный полипептид содержит рецептор цитокина, например, рецептор цитокина I типа или рецептор цитокина II типа, либо его производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит рецептор интерлейкина (например, IL-2R, IL-3R, IL-4R, IL-5R, IL-6R, IL-7R, IL-9R, IL-11R, IL-12R, IL-13R, IL-15R, IL-21R, IL-23R, IL-27R или IL-31R), рецептор колониестимулирующего фактора (например, рецептор эритропоэтина, CSF-1R, CSF-2R, GM-CSFR или G-CSFR), рецептор гормона/нейропептида (например, рецептор гормона роста, рецептор пролактина или рецептор лептина) либо его производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит рецептор цитокина II типа либо его производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит рецептор интерферона (например, IFNAR1, IFNAR2 или IFNGR), рецептор интерлейкина (например, IL-10R, IL-20R, IL-22R или IL-28R), рецептор тканевого фактора (также называемого тканевым тромбопластином) либо его производное, вариант или фрагмент.

Химерный рецепторный полипептид, содержащий цитокиновый рецептор, может связывать антиген, содержащий любой подходящий лиганд рецептора цитокинов, либо его производное, вариант или фрагмент. Неограничительные примеры лигандов цитокиновых рецепторов включают интерлейкины (например, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-9, IL-10, IL-11, IL-12, IL-13, IL-15, IL-20, IL-21, IL-22, IL-23, IL-27, IL-28 и IL-31), интерфероны (например, IFN- α , IFN- β , IFN- γ), колониестимулирующие факторы (например, эритропоэтин, колониестимулирующий фактор макрофагов, колониестимулирующий фактор гранулоцитов/макрофагов или GM-CSF, колониестимулирующий фактор гранулоцитов или G-CSF) и гормоны (например, пролактин и лептин).

В некоторых воплощениях химерный рецепторный полипептид содержит рецептор смерти, рецептор, содержащий домен смерти, либо его производное, вариант или фрагмент. Рецепторы смерти часто участвуют в регуляции апоптоза и воспаления. Рецепторы смерти включают представителей семейства рецепторов TNF типа TNFR1, Fas-рецептор, DR4 (также известный как TRAIL-рецептор 1 или TRAILR1) и DR5 (также известный как TRAIL-рецептор 2 или TRAILR2). В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внеклеточную область (например, лиганд-связывающий домен) рецептора смерти либо её производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, пронизывающий мембрану участок рецептора смерти либо его производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внутриклеточную область (например, цитозольный домен) рецептора смерти либо её производное, вариант или фрагмент. Химерный рецепторный полипептид, содержащий рецептор смерти либо его производное, вариант или фрагмент, может подвергаться олигомеризации рецептора в ответ на связывание лиганда, что, в свою очередь, может привести к рекрутированию специализированных адаптерных белков и активации сигнальных каскадов типа каспазных каскадов. В некоторых воплощениях связывание лиганда с химерным рецептором, содержащим рецептор смерти либо его производное, вариант или фрагмент, вызывает конформационное изменение, химическую модификацию либо их комбинацию, которые рекрутируют партнера по связыванию к рецептору.

Химерный рецепторный полипептид, содержащий рецептор смерти, может связывать антиген, включающий любой подходящий лиганд рецептора смерти либо его производное, вариант или фрагмент. Неограничительные примеры лигандов, связывающихся с рецепторами смерти, включают TNF α , Fas-лиганд и родственный TNF лиганд, индуцирующий апоптоз (TRAIL).

В некоторых воплощениях химерный рецепторный полипептид содержит иммунорецептор либо его производное, вариант или фрагмент. Иммунорецепторы включают представителей суперсемейства иммуноглобулинов (IgSF), имеющих общие структурные характеристики с иммуноглобулинами, например домен, известный как иммуноглобулиновый домен или складка. Представители IgSF включают, без ограничения, рецепторы антигенов клеточной поверхности, корецепторы и костимулирующие молекулы иммунной системы и молекулы, участвующие в презентации антигенов лимфоцитам. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внеклеточную область (например, лиганд-связывающий домен) иммунорецептора либо её производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, пронизывающий мембрану участок иммунорецептора либо его производное, вариант или фрагмент. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, внутриклеточную область (например, цитозольный домен) иммунорецептора либо её производное, вариант или фрагмент. Химерный рецепторный полипептид, содержащий иммунорецептор либо его производное, вариант или фрагмент, может рекрутировать партнера по связыванию. В некоторых воплощениях связывание лиганда с химерным рецептором, содержащим иммунорецептор либо его производное, вариант или фрагмент, вызывает конформационное изменение, химическую модификацию либо их комбинацию, которые рекрутируют партнера по связыванию к рецептору.

В некоторых воплощениях химерный рецепторный полипептид содержит рецептор антигенов клеточной поверхности типа Т-клеточного рецептора (TCR), В-клеточного рецептора (BCR) либо его произ-

водное, вариант или фрагмент. Т-клеточные рецепторы обычно содержат две цепи, это либо цепи TCR- α и TCR- β , либо цепи TCR- δ и TCR- γ . Химерный рецепторный полипептид, содержащий TCR либо его производное, вариант или фрагмент, может связываться с главным комплексом гистосовместимости (МНС). В-клеточные рецепторы обычно содержат мембраносвязанный иммуноглобулин и элемент трансдукции сигнала. Химерный рецептор, содержащий BCR либо его производное, вариант или фрагмент, может связывать соответствующий антиген BCR. В некоторых воплощениях химерный рецепторный полипептид содержит по меньшей мере активационный мотив иммунорецептора на основе тирозина (ITAM), который находится в цитоплазматическом домене некоторых иммунорецепторов. В некоторых воплощениях химерный рецепторный полипептид содержит, по меньшей мере, ингибирующий мотив иммунорецептора на основе тирозина (ITIM), который находится в цитоплазматическом домене некоторых иммунорецепторов. Химерные рецепторные полипептиды, содержащие домены ITAM и/или ITIM, могут подвергаться фосфорилированию после связывания лиганда с взаимодействующим с антигеном доменом. Фосфорилированные участки могут служить местами стыковки для других белков, участвующих в сигнализации иммунных клеток.

Взаимодействующий с антигеном домен химерного рецепторного полипептида может связывать мембраносвязанный антиген, к примеру, антиген на внеклеточной поверхности клетки (например, клетки мишени). В некоторых воплощениях взаимодействующий с антигеном домен связывает антиген, который не связан с мембраной, к примеру, внеклеточный антиген, который секретируется из клетки (например, клетки мишени), или антиген, локализованный в цитоплазме клетки. Антигены (например, мембраносвязанные и не связанные с мембраной) могут быть связаны с заболеванием типа вирусной, бактериальной и/или паразитарной инфекции; воспалительным и/или аутоиммунным заболеванием; либо неоплазией типа рака и/или опухоли. Раковые антигены, к примеру, - это белки, которые вырабатываются опухолевыми клетками и могут вызывать иммунный ответ, в особенности опосредованный Т-клетками иммунный ответ. Выбор антигенсвязывающих участков химерного рецепторного полипептида может зависеть от того, на какой конкретный тип раковых антигенов они должны быть нацелены. В некоторых воплощениях опухолевый антиген содержит один или несколько антигенных раковых эпитопов, связанных со злокачественной опухолью. Злокачественные опухоли могут экспрессировать ряд белков, которые могут служить в качестве антигенов-мишеней для иммунной атаки. Взаимодействующие с антигеном домены могут связываться с сигналами клеточной поверхности, внеклеточного матрикса (ЕСМ), паракринными сигналами, юкстакринными сигналами, эндокринными сигналами, аутокринными сигналами, сигналами, которые могут запускать или контролировать генетические программы в клетках либо их комбинациями. В некоторых воплощениях взаимодействия между клеточными сигналами, с которыми связываются рекомбинантные полипептиды химерных рецепторов, затрагивают взаимодействия между клетками, между клетками и растворимыми химическими веществами и между клетками и матриксом или микроокружением.

Полипептид, модулирующий ген (GMP) химерного рецепторного полипептида, может содержать исполнительный элемент, соединенный с сайтом распознавания расщепления. Исполнительный элемент может включать нуклеазу (например, ДНК-нуклеазу и/или РНК-нуклеазу), модифицированную нуклеазу (например, ДНК-нуклеазу и/или РНК-нуклеазу), лишённую нуклеазной активности или обладающую меньшей нуклеазной активностью по сравнению с нуклеазой дикого типа, её производное, вариант или фрагмент. Исполнительный элемент может регулировать экспрессию или активность гена и/или редактировать последовательность нуклеиновой кислоты (например, гена и/или продукта гена). В некоторых воплощениях исполнительный элемент содержит ДНК-нуклеазу типа генно-инженерной (например, программируемой или наводимой) ДНК-нуклеазы для запуска геномного редактирования целевой последовательности ДНК. В некоторых воплощениях исполнительный элемент содержит РНК-нуклеазу типа генно-инженерной (например, программируемой или наводимой) РНК-нуклеазы для запуска геномного редактирования целевой последовательности РНК. В некоторых воплощениях исполнительный элемент обладает пониженной или минимальной нуклеазной активностью. Исполнительный элемент с пониженной или минимальной нуклеазной активностью может регулировать экспрессию и/или активность гена путем физического ограждения целевого полинуклеотида или рекрутирования дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит безнуклеазный ДНК-связывающий белок, происходящий из ДНК-нуклеазы, который может индуцировать активацию или репрессию транскрипции целевой последовательности ДНК. В некоторых воплощениях исполнительный элемент содержит безнуклеазный Связывающий белок, связывающий РНК, происходящий из РНК-нуклеазы, который может индуцировать активацию или репрессию транскрипции целевой последовательности РНК. В некоторых воплощениях исполнительный элемент управляется направляющей нуклеиновой кислотой. В некоторых воплощениях исполнительный элемент управляется направляющей ДНК. В некоторых воплощениях исполнительный элемент управляется направляющей РНК. Исполнительный элемент может регулировать экспрессию или активность гена и/или редактировать последовательность нуклеиновой кислоты, будь то экзогенной или эндогенной.

В исполнительных элементах может использоваться любая подходящая нуклеаза. К подходящим нуклеазам относятся, без ограничения, CRISPR-ассоциированные (Cas) белки или Cas-нуклеазы, в том числе CRISPR-ассоциированные (Cas) полипептиды I типа, CRISPR-ассоциированные (Cas) полипептиды II типа, CRISPR-ассоциированные (Cas) полипептиды III типа, CRISPR-ассоциированные (Cas) полипептиды IV типа, CRISPR-ассоциированные (Cas) полипептиды V типа и CRISPR-ассоциированные (Cas) полипептиды VI типа; нуклеазы с "цинковым пальцем" (ZFN); эффекторные нуклеазы типа активаторов транскрипции (TALEN); мегануклеазы; РНК-связывающие белки (RBP); CRISPR-ассоциированные РНК-связывающие белки; рекомбиназы; флиппазы; транспозазы; белки Argonaute (Ago) (например, прокариотические белки Argonaute (pAgo), археальные Argonaute (aAgo) и эукариотические Argonaute (eAgo)); их производные, варианты или фрагменты.

Регуляция генов может относиться к любым генам, представляющим интерес. Предусматривается, что это охватывает и генетические гомологи описанных здесь генов. Например, ген может проявлять определенную степень идентичности и/или гомологии с описанными здесь генами. Так, предусматривается, что можно модифицировать гены, которые проявляют гомологичность на 50, 55, 60, 65, 70, 75, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 или 100% (на уровне нуклеиновой кислоты или белка). Также предусматривается, что можно модифицировать гены, которые проявляют идентичность на 50, 55, 60, 65, 70, 75, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 или 100% (на уровне нуклеиновой кислоты или белка).

В некоторых воплощениях исполнительный элемент содержит CRISPR-ассоциированный (Cas) белок или Cas-нуклеазу, которая функционирует в неприродной системе CRISPR (Clustered Regular Interspaced Short Palindromic Repeats)/Cas (CRISPR-associated). У бактерий эта система может обеспечивать адаптивный иммунитет к чужеродной ДНК (Barrangou R. et al. "CRISPR provides acquired resistance against viruses in prokaryotes" *Science* (2007) 315: 1709-1712; Makarova K.S. et al. "Evolution and classification of the CRISPR-Cas systems" *Nat Rev Microbiol.* (2011) 9:467-477; Garneau J.E. et al. "The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA" *Nature* (2010) 468:67-71; Sapranaukas R. et al. "The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*" *Nucleic Acids Res* (2011) 39: 9275-9282).

У самых различных организмов, включая разнообразных млекопитающих, животных, растений и дрожжей, система CRISPR/Cas (например, модифицированная и/или немодифицированная) может использоваться в качестве инструмента инженерии генома. Система CRISPR/Cas может содержать гидовую нуклеиновую кислоту типа направляющей РНК (gRNA) в комплексе с белком Cas для прицельной регуляции экспрессии и/или активности гена или редактирования нуклеиновой кислоты. Направляемый РНК белок Cas (например, нуклеаза Cas типа нуклеазы Cas9) может специфически связывать целевой полинуклеотид (например, ДНК) зависимым от последовательности образом. Cas-белки, если они обладают нуклеазной активностью, могут расщеплять ДНК (Gasiunas G. et al. "Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria" *Proc Natl Acad Sci USA* (2012) 109: E2579-E2 86; Jinek M. et al. "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity" *Science* (2012) 337:816-821; Steinberg S.H. et al. "DNA interrogation by the CRISPR RNA-guided endonuclease Cas9" *Nature* (2014) 507:62; Deltcheva E. et al. "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III" *Nature* (2011) 471:602-607) и широко применяются для программируемого редактирования генома у различных организмов и в модельных системах (Cong L. et al. "Multiplex genome engineering using CRISPR Cas systems" *Science* (2013) 339:819-823; Jiang W. et al. "RNA-guided editing of bacterial genomes using CRISPR-Cas systems" *Nat. Biotechnol.* (2013) 31: 233-239; Sander J.D. & Joung J.K. "CRISPR-Cas systems for editing, regulating and targeting genomes" *Nature Biotechnol.* (2014) 32:347-355).

В некоторых случаях белок Cas подвергается мутации и/или модификации, образуя дефектный по нуклеазе белок или белок с пониженной нуклеазной активностью по сравнению с белком Cas дикого типа. Дефектный по нуклеазе белок может сохранять способность к связыванию ДНК, но может быть лишен или же обладать пониженной активностью расщепления нуклеиновой кислоты. Исполнительный элемент, содержащий Cas-нуклеазу (например, сохраняющую активность нуклеазы дикого типа, с пониженной активностью нуклеазы и/или лишенную активности нуклеазы), может функционировать в системе CRISPR/Cas для регуляции уровня и/или активности целевого гена или белка (например, снижения, повышения или устранения), белок Cas может связываться с целевым полинуклеотидом и предотвращать транскрипцию посредством физического ограждения или редактировать последовательность нуклеиновой кислоты для получения нефункциональных генных продуктов.

В некоторых воплощениях исполнительный элемент включает белок Cas, который образует комплекс с направляющей нуклеиновой кислотой типа направляющей РНК). В некоторых воплощениях исполнительный элемент включает белок Cas, который образует комплекс с одной направляющей нуклеиновой кислотой типа одинарной направляющей РНК (sgRNA). В некоторых воплощениях исполнительный элемент включает белок, связывающий РНК (RBP), необязательно в комплексе с направляющей нуклеиновой кислотой типа направляющей РНК (например, sgRNA), которая способна образовывать комплекс с белком Cas. На фиг. 3А схематически представлена система, включающая химерный рецеп-

торный полипептид, в которой исполнительный элемент содержит связывающий белок, связывающий РНК 300а, необязательно в комплексе с направляющей нуклеиновой кислотой (например, sgRNA). После высвобождения из РНК-связывающего белка (RBP), к примеру, при диссоциации направляющей нуклеиновой кислоты из RBP или расщеплении сайта распознавания расщепления 300с, Направляющая нуклеиновая кислота может образовать комплекс с белком Cas 300b, который способен регулировать экспрессию и/или активность гена или редактировать последовательность нуклеиновой кислоты. В некоторых воплощениях исполнительный элемент содержит безнуклеазный белок, связывающий ДНК, полученный из ДНК-нуклеазы, который может вызывать активацию или репрессию транскрипции целевой последовательности ДНК. В некоторых воплощениях исполнительный элемент содержит безнуклеазный связывающий белок, связывающий РНК, полученный из РНК-нуклеазы, который может вызывать активацию или репрессию транскрипции целевой последовательности РНК. Например, исполнительный элемент может содержать белок Cas, который лишен активности расщепления.

Может использоваться любая подходящая система CRISPR/Cas. Для обозначения системы CRISPR/Cas могут применяться различные системы обозначения. Типичные системы обозначения приведены в Makarova K.S. et al. "An updated evolutionary classification of CRISPR-Cas systems" *Nat Rev Microbiol.* (2015)13:722-736; и Shmakov S. et al. "Discovery and functional characterization of diverse class 2 CRISPR-Cas systems" *Mol Cell* (2015) 60:1-13. Система CRISPR/Cas может быть системой I типа, II типа, III типа, IV типа, V типа, VI типа или любой другой подходящей системой CRISPR/Cas. В настоящем изобретении система CRISPR/Cas может быть системой класса 1, класса 2 или любой другой подходящей системой CRISPR/Cas. Определение класса 1 или класса 2 основывается на генах, кодирующих эффекторный модуль. Системы класса 1 обычно содержат мультисубъединичный crRNA-эффекторный комплекс, тогда как системы класса 2 обычно содержат эффекторный комплекс из одного белка типа Cas9, Cpf1, C2c1, C2c2, C2c3 или crRNA. В системах CRISPR/Cas класса 1 для осуществления регуляции может использоваться комплекс из нескольких Cas-белков. Система CRISPR/Cas класса 1 может включать, к примеру, CRISPR/Cas I типа (например, I, IA, IB, IC, ID, IE, IF, IU), III типа (например, III, IIIA, IIIB, IIIC, IIID) и IV типа (например, IV, IVA, IVB). В системах CRISPR/Cas класса 2 для осуществления регуляции может использоваться один большой белок Cas. Система CRISPR/Cas класса 2 может включать, к примеру, CRISPR/Cas II типа (например, II, IIA, IIB) и V типа. Системы CRISPR могут быть взаимодополняющими и/или же функциональные модули могут располагаться в транс-положении, облегчая наведение на локус CRISPR. На фиг. 15 представлена иллюстрация, адаптированная из фиг. 2 в Makarova K.S. et al. "An updated evolutionary classification of CRISPR-Cas systems" *Nat Rev Microbiol.* (2015) 13:722-736, на которой представлена архитектура геномных локусов для различных подтипов системы CRISPR-Cas.

Исполнительный элемент, содержащий белок Cas, может быть представлен белком Cas класса 1 или класса 2. белок Cas может представлять собой белок Cas I типа, II типа, III типа, IV типа, V типа или VI типа, белок Cas может содержать один или несколько доменов. Неограничительные примеры доменов включают домен распознавания и/или связывания направляющей нуклеиновой кислоты, нуклеазные домены (например, ДНКазные или РНКазные домены, RuvC, HNH), ДНК-связывающий домен, РНК-связывающий домен, геликазные домены, домены взаимодействия белок-белок и домены димеризации. Домен распознавания и/или связывания направляющей нуклеиновой кислоты может взаимодействовать с направляющей нуклеиновой кислотой. Нуклеазный домен может содержать каталитическую активность для расщепления нуклеиновой кислоты. Нуклеазный домен может быть лишен каталитической активности для предотвращения расщепления нуклеиновой кислоты, белок Cas может представлять собой химерный белок Cas, слитый с другими белками или полипептидами. Белок Cas может быть химерой из различных Cas-белков, к примеру, содержащих домены из разных Cas-белков.

Неограничительные примеры Cas-белков включают c2c1, C2c2, c2c3, Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas5e (CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a, Cas8a1, Cas8a2, Cas8b, Cas8c, Cas9 (Csn1 или Csx12), Cas10, Cas10d, Cas10, Cas10d, CasF, CasG, CasH, Cpf1, Csy1, Csy2, Csy3, Cse1 (CasA), Cse2 (CasB), Cse3 (CasE), Cse4 (CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx4, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4 и Cul966, их гомологи или модифицированные варианты.

Белок Cas может происходить из любого подходящего организма. Неограничительные примеры включают

Streptococcus pyogenes, *Streptococcus thermophilus*, *Streptococcus* sp., *Staphylococcus aureus*, *Nocardiopsis dassonvillei*, *Streptomyces pristinae spiralis*, *Streptomyces viridochromogenes*, *Streptosporangium roseum*, *Alicyclobacillus acidocaldarius*, *Bacillus pseudomycolides*, *Bacillus selenitireducens*, *Exiguobacterium sibiricum*, *Lactobacillus delbrueckii*, *Lactobacillus salivarius*, *Microscilla marina*, *Burkholderiales bacterium*, *Polaromonas naphthalenivorans*, *Polaromonas* sp., *Crocospaera watsonii*, *Cyanothece* sp., *Microcystis aeruginosa*, *Pseudomonas aeruginosa*, *Synechococcus* sp., *Acetohalobium arabaticum*, *Ammonifex degenzii*, *Caldicelulosiruptor beccsii*, *Candidatus desulfuridis*, *Clostridium botulinum*, *Clostridium difficile*, *Finegoldia magna*, *Natranaerobius thermophilus*, *Pelotomaculum thermopropionicum*, *Acidithiobacillus caldus*, *Acidithiobacillus ferrooxidans*, *Allochromatium vinosum*, *Marinobacter* sp., *Nitrosococcus halophilus*, *Nitrosococcus watsoni*, *Pseudoalteromonas haloplanktis*, *Ktedonobacter racemifer*, *Methanohalobium evestigatum*, *Anabaena variabilis*, *Nodularia spumigena*, *Nostoc* sp., *Arthrospira maxima*, *Arthrospira platensis*, *Arthrospira* sp., *Lyngbya* sp., *Microcoleus chthonoplastes*, *Oscillatoria* sp., *Petrotoga mobilis*, *Thermosiphon africanus*, *Acarochloris marina*, *Leptotrichia shahii* и *Francisella novicida*.

В некоторых аспектах организм представлен *Streptococcus pyogenes* (*S.pyogenes*). В некоторых аспектах организм представлен *Staphylococcus aureus* (*S.aureus*). В некоторых аспектах организм представлен *Streptococcus thermophilus* (*S.thermophilus*).

Белок Cas может происходить из различных видов бактерий, включая, без ограничения, атипические

Veillonella, *Fusobacterium nucleatum*, *Filifactor alocis*, *Solobacterium moorei*, *Coprococcus catus*, *Treponema denticola*, *Peptoniphilus duerdenii*, *Catenibacterium mitsuokai*, *Streptococcus mutans*, *Listeria innocua*, *Staphylococcus pseudintermedius*, *Acidaminococcus intestine*, *Olsenella uli*, *Oenococcus kitaharae*, *Bifidobacterium bifidum*, *Lactobacillus rhamnosus*, *Lactobacillus gasseri*, *Finegoldia magna*, *Mycoplasma mobile*, *Mycoplasma gallisepticum*, *Mycoplasma ovipneumoniae*, *Mycoplasma canis*, *Mycoplasma synoviae*, *Eubacterium rectale*, *Streptococcus thermophilus*, *Eubacterium dolichum*, *Lactobacillus coryniformis* subsp. *torquens*, *Ilyobacter polytropus*, *Ruminococcus albus*, *Akkermansia muciniphila*, *Acidothermus cellulolyticus*, *Bifidobacterium longum*, *Bifidobacterium dentium*, *Corynebacterium diphtheria*, *Elusimicrobium minutum*, *Nitratifactor salsuginis*, *Sphaerochaeta globus*, *Fibrobacter succinogenes* subsp. *succinogenes*, *Bacteroides fragilis*, *Capnocytophaga ochracea*, *Rhodopseudomonas palustris*, *Prevotella micans*, *Prevotella ruminicola*, *Flavobacterium columnare*, *Aminomonas paucivorans*, *Rhodospirillum rubrum*, *Candidatus*, *Puniceispirillum marinum*, *Verminephrobacter eiseniae*, *Ralstonia syzygii*, *Dinoroseobacter shibae*, *Azospirillum*, *Nitrobacter hamburgensis*, *Bradyrhizobium*, *Wolinella succinogenes*, *Campylobacter jejuni* subsp. *jejuni*, *Helicobacter mustelae*, *Bacillus cereus*, *Acidovorax ebreus*, *Clostridium perfringens*, *Parvibaculum lavamentivorans*, *Roseburia intestinalis*, *Neisseria meningitidis*, *Pasteurella multocida* subsp. *multocida*, *Sutterella wadsworthensis*, *Proteobacterium*, *Legionella pneumophila*, *Parasutterella excrementihominis*, *Wolinella succinogenes* и *Francisella novicida*.

В настоящем изобретении белок Cas может быть белком дикого типа или модифицированной формой Cas-белка. белок Cas может представлять собой активный вариант, неактивный вариант или фрагмент белка дикого типа или модифицированного Cas-белка. белок Cas может включать изменения аминокислот типа делеции, вставки или замены, варианты, мутации, слияния, химеры или их комбинации относительно Cas-белка дикого типа, белок Cas может быть полипептидом, который по последовательности по меньшей мере на 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 или 100% идентичен или аналогичен типичному Cas-белку дикого типа, белок Cas может быть полипептидом, который по последовательности не более чем на 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100% идентичен и/или аналогичен типичному Cas-белку дикого типа. Варианты или фрагменты по последовательности могут быть по меньшей мере на 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99 или 100% идентичны или аналогичны белку дикого типа или модифицированному Cas-белку или его части.

Варианты или фрагменты могут быть нацелены на locus нуклеиновой кислоты в комплексе с направляющей нуклеиновой кислотой даже при отсутствии активности расщепления нуклеиновой кислоты.

Белок Cas может содержать один или несколько нуклеазных доменов типа ДНКазных доменов. Например, белок Cas9 может содержать нуклеазный домен типа RuvC и/или нуклеазный домен типа HNH. Домены RuvC и HNH могут разрезать разные нити двухцепочечной ДНК, образуя двухцепочечный разрыв в ДНК. белок Cas может содержать и только один нуклеазный домен (например, Cpf1 содержит домен RuvC, но не содержит домена HNH).

Белок Cas может содержать аминокислотную последовательность, которая по последовательности по меньшей мере на 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 91 92, 93, 94, 95, 96, 97, 98, 99 или 100% идентична или аналогична нуклеазному домену (например, домену RuvC, домену HNH) белка Cas дикого типа.

Белок Cas может быть модифицирован для оптимизации регуляции экспрессии генов. Белок Cas может быть модифицирован для повышения или уменьшения сродства связывания с нуклеиновой кислотой, специфичности связывания нуклеиновой кислоты и/или ферментативной активности. Белок Cas также может быть модифицирован для изменения любой другой активности или свойства белка типа стабильности. Например, может быть модифицирован, удален или инактивирован один или несколько нуклеазных доменов Cas-белка или же белок Cas может быть усечен для удаления доменов, которые не существенны для функции белка, или же для оптимизации (например, повышения или снижения) активности Cas-белка для регуляции экспрессии генов.

Белок Cas может быть слитым белком. Например, белок Cas может быть слит с доменом расщепления, доменом эпигенетической модификации, доменом активации транскрипции или доменом репрессии транскрипции. Белок Cas также может быть слит с гетерологичным полипептидом, обеспечивающим повышенную или пониженную стабильность. Слитый домен или гетерологичный полипептид может располагаться на N-конце, C-конце или внутри Cas-белка.

Белок Cas может быть представлен в любом виде. Например, белок Cas может быть представлен в виде белка типа одного лишь белка Cas или в комплексе с направляющей нуклеиновой кислотой. Белок Cas может быть представлен в виде нуклеиновой кислоты, кодирующей белок Cas, типа РНК (например, матричной РНК (мРНК)) или ДНК. Нуклеиновая кислота, кодирующая белок Cas, может быть оптимизирована по кодонам для эффективной трансляции белка в определенных клетках или организме.

Нуклеиновые кислоты, кодирующие белки Cas, могут быть стабильно встроены в геном клетки. Нуклеиновые кислоты, кодирующие Cas-белки, могут быть функционально связаны с промотором, работающим в клетке. Нуклеиновые кислоты, кодирующие Cas-белки, могут быть функционально связаны с промотором в экспрессирующей конструкции. Экспрессионные конструкции могут включать в себя конструкции из нуклеиновых кислот, способные направлять экспрессию нужного гена или другой последовательности нуклеиновой кислоты (например, гена Cas), которые могут переносить такую представляющую интерес последовательность нуклеиновой кислоты в клетки мишени.

В некоторых воплощениях белок Cas является "мертвым" белком Cas. Мертвым белком Cas может быть белок, который лишен активности расщепления нуклеиновой кислоты.

Белок Cas может включать модифицированную форму белка Cas дикого типа. Модифицированная форма Cas-белка дикого типа может включать изменения аминокислот (например, делеции, вставки или замены), которые снижают активность расщепления нуклеиновой кислоты у Cas-белка. Например, модифицированная форма Cas-белка может иметь менее 90%, менее 80%, менее 70%, менее 60%, менее 50%, менее 40%, менее 30%, менее 20%, менее 10%, менее 5% или менее 1% активности расщепления нуклеиновых кислот у Cas-белка дикого типа (например, Cas9 из *S. pyogenes*). Модифицированная форма Cas-белка может практически не обладать активностью расщепления нуклеиновой кислот. Когда белок Cas представлен модифицированной формой, практически не обладающей активностью расщепления нуклеиновой кислоты, то он может называться ферментативно неактивным и/или "мертвым" (сокращенно "d"). Мертвый белок Cas (например, dCas, dCas9) может связываться с целевым полинуклеотидом, но не может его расщеплять. В некоторых аспектах мертвый белок Cas представлен мертвым белком Cas9.

Полипептид dCas9 может связываться с одинарной направляющей РНК (sgRNA), чтобы активировать или репрессировать транскрипцию целевой ДНК. sgRNAs можно вводить в клетки, экспрессирующие генно-инженерный полипептид химерного рецептора. В некоторых случаях такие клетки содержат одну или несколько различных sgRNAs, которые нацелены на одну и ту же нуклеиновую кислоту. В других случаях sgRNAs нацелены на разные нуклеиновые кислоты в клетке. Нуклеиновые кислоты, на которые нацелена направляющая РНК, могут быть любыми, которые экспрессируются в клетках типа иммунных клеток. Целевой нуклеиновой кислотой может быть ген, участвующий в регуляции иммунных клеток. В некоторых воплощениях нуклеиновая кислота связана с раком. Нуклеиновая кислота, связанная с раком, может представлять собой ген клеточного цикла, ген клеточного ответа, ген апоптоза или ген фагоцитоза. Рекомбинантная направляющая РНК может распознаваться белком CRISPR, безнуклеазным белком CRISPR, их вариантами, производными или фрагментами.

Ферментативно неактивный может означать такой полипептид, который может связываться с последовательностью нуклеиновой кислоты в полинуклеотиде специфичным для последовательности образом, но не может расщеплять целевой полинуклеотид. Ферментативно неактивный сайт-направленный полипептид может содержать ферментативно неактивный домен (например, нуклеазный домен). Ферментативно неактивный может означать отсутствие активности. Ферментативно неактивный может озна-

чать практически полное отсутствие активности. Ферментативно неактивный может означать почти полное отсутствие активности. Ферментативно неактивный может означать активность на уровне менее 1%, менее 2%, менее 3%, менее 4%, менее 5%, менее 6%, менее 7%, менее 8%, менее 9% или менее 10% от типичной активности дикого типа (например, от активности расщепления нуклеиновой кислоты у Cas9 дикого типа).

У белка Cas может быть удален или мутирован один или несколько нуклеазных доменов (например, RuvC, HNH) с тем, чтобы они больше не были функциональными или обладали меньшей нуклеазной активностью. Например, если у белка Cas, содержащего по меньшей мере два нуклеазных домена (например, Cas9), удалить или мутировать один из нуклеазных доменов, то такой белок Cas, известный как никаза, может создавать одноцепочечные разрывы на участке распознавания РНК CRISPR (crRNA) в двухцепочечной ДНК, но не двухцепочечные разрывы. Такая никаза может расщеплять комплементарную нить или некомплеметарную нить, но не может расщеплять обе. Если у белка Cas удалить или мутировать все нуклеазные домены (например, нуклеазные домены RuvC и HNH у белка Cas9, нуклеазный домен RuvC у белка Cpf1), то такой белок Cas может обладать пониженной способностью или совсем не обладать способностью к расщеплению обеих нитей двухцепочечной ДНК. Примером мутации, которая может превратить белок Cas9 в никазу, является мутация D10A (аспартат на аланин в положении 10 у Cas9) в домене RuvC у Cas9 из *S.pyogenes*. Мутация H939A (гистидин на аланин в положении аминокислоты 839) или H840A (гистидин на аланин в положении аминокислоты 840) в домене HNH у Cas9 из *S.pyogenes* может превратить Cas9 в никазу. Примером мутации, которая может превратить белок Cas9 в мертвый Cas9, является мутация D10A (аспартат на аланин в положении 10 у Cas9) в домене RuvC и мутация H939A (гистидин на аланин в положении аминокислоты 839) или H840A (гистидин на аланин в положении аминокислоты 840) в домене HNH у Cas9 из *S.pyogenes*.

"Мертвый" белок Cas может содержать одну или несколько мутаций по сравнению с версией белка дикого типа. В результате мутации активность расщепления нуклеиновой кислоты у одного или нескольких доменов расщепления нуклеиновой кислоты может составлять менее 90%, менее 80%, менее 70%, менее 60%, менее 50%, менее 40%, менее 30%, менее 20%, менее 10%, менее 5% или менее 1% от Cas-белка дикого типа. В результате мутации один или несколько доменов расщепления нуклеиновой кислоты могут сохранить способность к расщеплению комплементарной нити целевой нуклеиновой кислоты, но уменьшится способность к расщеплению некомплеметарной нити целевой нуклеиновой кислоты. В результате мутации один или несколько доменов расщепления нуклеиновой кислоты могут сохранить способность к расщеплению некомплеметарной нити целевой нуклеиновой кислоты, но уменьшится способность к расщеплению комплементарной нити целевой нуклеиновой кислоты. В результате мутации один или несколько доменов расщепления нуклеиновой кислоты могут утратить способность к расщеплению и комплементарной нити, и некомплеметарной нити целевой нуклеиновой кислоты. Подлежащие мутации остатки в нуклеазном домене могут соответствовать одному или нескольким каталитическим остаткам нуклеазы. Например, для инактивации одного или нескольких доменов расщепления нуклеиновой кислоты (например, нуклеазных доменов) можно подвергать мутации такие остатки у типичного полипептида Cas9 дикого типа *S.pyogenes*, как Asp10, His840, Asn854 и Asn856. Остатки, подлежащие мутации в нуклеазном домене Cas-белка, могут соответствовать остаткам Asp10, His840, Asn854 и Asn856 у полипептида Cas9 дикого типа *S.pyogenes*, например, при определении по последовательности и/или по структурному совмещению.

В качестве неограниченных примеров можно подвергать мутации остатки D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986 и/или A987 (или соответствующие мутации у любых белков Cas). К примеру, D10A, G12A, G17A, E762A, H840A, N854A, N863A, H982A, H983A, A984A и/или D986A. Могут подойти и мутации, отличные от замены на аланин.

Мутацию D10A можно комбинировать с одной или несколькими мутациями H840A, N854A или N856A, получая белок Cas9, практически не обладающий активностью расщепления ДНК (например, мертвый белок Cas9). Мутацию H840A можно комбинировать с одной или несколькими мутациями D10A, N854A или N856A, получая сайт-направленный полипептид, практически не обладающий активностью расщепления ДНК. Мутацию N854A можно комбинировать с одной или несколькими мутациями H840A, D10A или N856A, получая сайт-направленный полипептид, практически не обладающий активностью расщепления ДНК. Мутацию N856A можно комбинировать с одной или несколькими мутациями H840A, N854A или D10A, получая сайт-направленный полипептид, практически не обладающий активностью расщепления ДНК.

В некоторых воплощениях белок Cas представлен белком Cas класса 2. В некоторых воплощениях белок Cas представлен белком Cas II типа. В некоторых воплощениях белок Cas представлен белком Cas9, модифицированной версией белка Cas9 или получен из белка Cas9. К примеру, это белок Cas9, не обладающий активностью расщепления. В некоторых воплощениях белок Cas9 представлен белком Cas9 из *S.pyogenes* (например, с номером доступа Q99ZW2 в SwissProt). В некоторых воплощениях белок Cas9 представлен Cas9 из *S. aureus* (например, с номером доступа J7RUA5 в SwissProt). В некоторых воплощениях белок Cas9 представлен модифицированной версией белка Cas9 из *S.pyogenes* или *S.aureus*. В некоторых воплощениях белок Cas9 получен из белка Cas9 из *S.pyogenes* или *S.aureus*. К примеру, это

Cas9 из *S.pyogenes* или *S.aureus*, не обладающий активностью расщепления. Cas9 обычно может означать полипептид, который по последовательности по меньшей мере на 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100% идентичен и/или аналогичен типичному полипептиду Cas9 дикого типа (например, Cas9 из *S.pyogenes*). Cas9 может означать полипептид, который по последовательности не более чем на 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100% идентичен и/или аналогичен типичному полипептиду Cas9 дикого типа (например, из *S.pyogenes*). Cas9 может означать белок дикого типа или модифицированную форму белка Cas9, которая может включать изменения аминокислот типа делеции, вставки или замены, варианты, мутации, слияния, химеры или их комбинации.

В некоторых воплощениях исполнительный элемент содержит "нуклеазу с цинковым пальцем" или "ZFN". ZFN означает слияние между доменом расщепления типа домена расщепления FokI и по меньшей мере одним мотивом "цинковый палец" (например, по меньшей мере с 2, 3, 4 или 5 мотивами "цинковый палец"), который может связывать такие полинуклеотиды, как ДНК и РНК. Гетеродимеризация двух индивидуальных ZFN по определенным положениям в полинуклеотиде при определенной ориентации и расстановке может привести к расщеплению полинуклеотида. Например, связывание ZFN с ДНК может вызвать двухцепочечный разрыв ДНК. Для того, чтобы два домена расщепления могли димеризоваться и расщеплять ДНК, две индивидуальные ZFN должны связаться с противоположными нитями ДНК так, чтобы их С-концы были на определенном расстоянии друг от друга. В некоторых случаях может потребоваться, чтобы линкерные последовательности между доменом с цинковым пальцем и доменом расщепления отделяли 5'-концы каждого сайта связывания на 5-7 пар оснований. В некоторых случаях домен расщепления сливаются с С-концом каждого домена с цинковым пальцем. Типичные ZFN включают, без ограничения, описанные в Urnov et al., *Nature Reviews Genetics*, 2010, 11: 636-646; Gaj et al., *Nat Methods*, 2012, 9(8):805-7; U.S. Patent Nos. 6534261; 6607882; 6746838; 6794136; 6824978; 6866997; 6933113; 6979539; 7013219; 7030215; 7220719; 7241573; 7241574; 7585849; 7595376; 6903185; 6479626; и U.S. Application Publication Nos. 2003/0232410 и 2009/0203140.

В некоторых воплощениях содержащий ZFN исполнительный элемент может создавать двухцепочечные разрывы в целевых полинуклеотидах типа ДНК. Двухцепочечный разрыв ДНК может привести к репарации разрыва ДНК, что способствует проведению модификации генов (например, редактированию нуклеиновой кислоты). Репарация разрыва ДНК может происходить путем соединения негомологических концов (NHEJ) или путем репарации по гомологии (HDR). При HDR предоставляется донорская матрица репарации ДНК, которая содержит фланкирующие сайты с участками гомологии к ДНК мишени. В некоторых воплощениях ZFN представлена никазой с цинковым пальцем, которая вызывает сайт-специфичные одноцепочечные разрывы или "ники" ДНК, что приводит к HDR. Описание никаз с цинковым пальцем приводится, например, в Ramirez et al., *Nucl Acids Res*, 2012, 40(12): 5560-8; Kim et al., *Genome Res*, 2012, 22(7): 1327-33. В некоторых воплощениях ZFN связывает полинуклеотиды (например, ДНК и/или РНК), но не способна их расщеплять.

В некоторых воплощениях домен расщепления исполнительного элемента, содержащего ZFN, содержит модифицированную форму домена расщепления дикого типа. Модифицированная форма домена расщепления может включать изменения аминокислот (например, делеции, вставки или замены), которые снижают активность расщепления нуклеиновой кислоты у домена расщепления. К примеру, активность расщепления нуклеиновой кислоты у модифицированной формы домена расщепления может составлять менее 90%, менее 80%, менее 70%, менее 60%, менее 50%, менее 40%, менее 30%, менее 20%, менее 10%, менее 5% или менее 1% от активности расщепления нуклеиновой кислоты у домена расщепления дикого типа. Модифицированная форма домена расщепления может практически не обладать активностью расщепления нуклеиновой кислоты. В некоторых воплощениях домен расщепления ферментативно неактивен.

В некоторых воплощениях исполнительный элемент содержит "TALEN" или "TAL-эффекторную нуклеазу". TALEN относится к генно-инженерным эффекторным нуклеазам типа активаторов транскрипции, которые обычно содержат центральный домен ДНК-связывающих tandemных повторов и расщепляющий домен. TALENs получают путем слияния ДНК-связывающего домена TAL-эффектора с доменом расщепления ДНК. В некоторых случаях ДНК-связывающий tandemный повтор содержит 33-35 аминокислот и содержит два гипервариабельных аминокислотных остатка в положениях 12 и 13, которые могут распознавать по меньшей мере одну определенную пару оснований ДНК. Белок эффектора типа активатора транскрипции (TALE) может быть слит с нуклеазой типа эндонуклеазы дикого типа или мутантной эндонуклеазы FokI или с каталитическим доменом FokI. Получено несколько мутаций у FokI для её использования в TALENs, которые, к примеру, улучшают специфичность или активность расщепления. Такие TALENs могут быть настроены на связывание любой желательной последовательности ДНК. TALENs можно использовать для получения модификаций генов (например, редактирования последовательности нуклеиновых кислот) путем создания двухцепочечных разрывов в целевой последовательности ДНК, которая в свою очередь подвергается NHEJ или HDR. В некоторых случаях для осуществления HDR предоставляется одноцепочечная донорская матрица репарации ДНК. Подробное описание TALENs и их применение для редактирования генов приводится, например, в U.S. Patent Nos. 8440431, 8440432, 8450471, 8586363 и 8697853; Scharenberg et al., *Curr Gene Ther.*, 2013, 13(4): 291-303; Gaj et al.,

Nat Methods, 2012, 9 (8): 805-7; Beurdeley et al., Nat Commun, 2013, 4:1762; и Joung and Sander, Nat Rev Mol Cell Biol., 2013, 14 (1): 49-55.

В некоторых воплощениях получают TALEN с пониженной нуклеазной активностью. В некоторых воплощениях нуклеазный домен TALEN содержит модифицированную форму нуклеазного домена дикого типа. Модифицированная форма нуклеазного домена может включать изменения аминокислот (например, делеции, вставки или замены), которые снижают активность расщепления нуклеиновой кислоты у нуклеазного домена. К примеру, активность расщепления нуклеиновой кислоты у модифицированной формы нуклеазного домена может составлять менее 90%, менее 80%, менее 70%, менее 60%, менее 50%, менее 40%, менее 30%, менее 20%, менее 10%, менее 5% или менее 1% от активности расщепления нуклеиновой кислоты у нуклеазного домена дикого типа. Модифицированная форма нуклеазного домена может практически не обладать активностью расщепления нуклеиновой кислоты. В некоторых воплощениях нуклеазный домен ферментативно неактивен.

В некоторых воплощениях белок эффектора типа активатора транскрипции (TALE) слит с доменом, который может модулировать транскрипцию и не содержит нуклеазы. В некоторых воплощениях белок эффектора типа активатора транскрипции (TALE) предназначается для функционирования в качестве активатора транскрипции. В некоторых воплощениях белок эффектора типа активатора транскрипции (TALE) предназначается для функционирования в качестве репрессора транскрипции. Например, ДНК-связывающий домен белка эффектора типа активатора транскрипции (TALE) может быть слит (например, соединен) с одним или несколькими доменами активации транскрипции или с одним или несколькими доменами репрессии транскрипции. Неограничительные примеры доменов активации транскрипции включают домен активации VP16 Herpes simplex и тетрамерный повтор домена активации VP16, например, домен активации VP64. Неограничительный пример домена репрессии транскрипции включает домен Krüppel-associated box (KRAB).

В некоторых воплощениях исполнительный элемент содержит мегануклеазу. Мегануклеазы обычно относятся к редко-расщепляющим эндонуклеазам или хоминг-эндонуклеазам, которые могут быть высокоспецифичными. Мегануклеазы могут распознавать участки ДНК мишени длиной по меньшей мере от 12 пар оснований, например, длиной от 12 до 40 пар оснований, от 12 до 50 пар оснований или от 12 до 60 пар оснований. Мегануклеазы могут представлять собой модульные ДНК-связывающие нуклеазы типа слитых белков, содержащих по меньшей мере один каталитический домен эндонуклеазы и по меньшей мере один ДНК-связывающий домен или белок, определяющий последовательность целевой нуклеиновой кислоты. ДНК-связывающий домен может содержать по меньшей мере один мотив, который распознает одно- или двухцепочечную ДНК. Мегануклеаза может быть мономерной или димерной. В некоторых воплощениях мегануклеаза является природной (встречается в природе) или дикого типа, а в других случаях мегануклеаза бывает неприродной, искусственной, инженерной, синтезированной, полученной методом рационального проектирования или вручную. В некоторых воплощениях мегануклеазы по настоящему изобретению включают мегануклеазу I-CreI, мегануклеазу I-CeuI, мегануклеазу I-MsoI, мегануклеазу I-SceI, их варианты, производные и фрагменты. Подробные описания полезных мегануклеаз и их применение при редактировании генов приводится, например, в Silva et al., Curr Gene Ther, 2011, 11(1): 11-27; Zaslavoskiy et al., BMC Bioinformatics, 2014, 15:191; Takeuchi et al., Proc Natl Acad Sci USA, 2014, 111 (11): 4061-4066; и в U.S. Patent Nos. 7842489; 7897372; 8021867; 8163514; 8133697; 8021867; 8119361; 8119381; 812436; и 8129134.

В некоторых воплощениях нуклеазный домен мегануклеазы содержит модифицированную форму нуклеазного домена дикого типа. Модифицированная форма нуклеазного домена может включать изменения аминокислот (например, делеции, вставки или замены), которые снижают активность расщепления нуклеиновой кислоты у нуклеазного домена. К примеру, активность расщепления нуклеиновой кислоты у модифицированной формы нуклеазного домена может составлять менее 90%, менее 80%, менее 70%, менее 60%, менее 50%, менее 40%, менее 30%, менее 20%, менее 10%, менее 5% или менее 1% от активности расщепления нуклеиновой кислоты у нуклеазного домена дикого типа. Модифицированная форма нуклеазного домена может практически не обладать активностью расщепления нуклеиновой кислоты. В некоторых воплощениях нуклеазный домен ферментативно неактивен. В некоторых воплощениях мегануклеаза может связывать ДНК, но не может расщеплять ДНК.

В некоторых воплощениях исполнительный элемент слит с одним или несколькими доменами репрессоров транскрипции, активаторов транскрипции, эпигенетическими доменами, рекомбиназными доменами, транспозазными доменами, флиппазными доменами, никазными доменами или их комбинациями. Домен активатора может включать один или несколько tandemных доменов активации, расположенных на С-конце белка. В некоторых случаях исполнительный элемент включает в себя один или несколько tandemных репрессорных доменов, расположенных на С-конце белка. Неограничительные примеры активаторных доменов включают GAL4, домен активации VP16 вируса Herpes simplex, VP64 (тетрамер из активационных доменов VP16 Herpes simplex), субъединицу p65 NF-κB, трансаактиватор R вируса Эпштейна-Барра (Rta), которые описаны в Chavez et al., Nat Methods, 2015, 12(4): 326-328; и U.S. Patent App. Publ. № 2014/0068797. Неограничительные примеры репрессорных доменов включают домен KRAB (Krüppel-associated box) из Kox1, домен взаимодействия Mad-mSIN3 (SID), репрессорный домен

ERF (ERD), которые описаны в Chavez et al., Nat Methods, 2015, 12(4): 326-328; и U.S. Patent App. Publ. № 2014/0068797. Исполнительный элемент также может быть слит с гетерологичным полипептидом, обеспечивающим повышенную или пониженную стабильность. Слитый домен или гетерологичный полипептид может располагаться на N-конце, на C-конце или внутри исполнительного элемента.

Исполнительный элемент может содержать гетерологичный полипептид для удобства отслеживания или очистки типа флуоресцентного белка, метки для очистки или эпитопной метки. Примеры флуоресцентных белков включают зеленые флуоресцентные белки (например, GFP, GFP-2, tagGFP, turboGFP, eGFP, Emerald, Azami Green, Monomeric Azami Green, CopGFP, AceGFP, ZsGreen1), желтые флуоресцентные белки (например, YFP, eYFP, Citrine, Venus, YPet, PhiYFP, ZsYellow1), синие флуоресцентные белки (например, eBFP, eBFP2, Azurite, mKalama1, GFPuv, Sapphire, T-Sapphire), голубые флуоресцентные белки (например, eCFP, Cerulean, CyPet, AmCyan1, Midoriishi-Cyan), красные флуоресцентные белки (mKate, mKate2, mPlum, DsRed monomer, mCherry, mRFP1, DsRed-Express, DsRed2, DsRed-Monomer, HcRed-Tandem, HcRed1, AsRed2, eqFP611, mRaspberry, mStrawberry, Jred), оранжевые флуоресцентные белки (mOrange, mKO, Kusabira-Orange, Monomeric Kusabira-Orange, m-Tangerine, tdTomato) и другие подходящие флуоресцентные белки. Примеры тегов включают глутатион-S-трансферазу (GST), хитин-связывающий белок (CBP), белок, связывающий мальтозу, тиоредоксин (TRX), poly(NANP), тандемный метка для аффинной очистки (TAP), мус, AcV5, AU1, AU5, E, ECS, E2, FLAG, гемагглютинин (HA), nus, Softag 1, Softag 3, Strep, SBP, Glu-Glu, HSV, KT3, S, SI, T7, V5, VSV-G, гистидин (His), биотин-карбокисильный белок-переносчик (BCCP) и кальмодулин.

В некоторых конфигурациях химерного рецепторного полипептида сайт распознавания расщепления в GMP может быть фланкирован взаимодействующим с антигеном доменом и исполнительным элементом. Исполнительный элемент может высвобождаться из GMP при расщеплении сайта распознавания расщепляющим элементом. Расщепляющий элемент может распознавать и/или расщеплять сайт распознавания расщепления, к примеру, если он находится в близости от сайта распознавания расщепления. Расщепляющий элемент может содержать полипептидную последовательность. Расщепляющий элемент может входить в состав химерного адаптерного полипептида. Расщепляющий элемент может составлять N-конец, C-конец или внутреннюю часть химерного адаптерного полипептида. В некоторых воплощениях расщепляющий элемент образует комплекс с химерным адаптерным полипептидом. Расщепляющий элемент может образовать комплекс с N-концом, C-концом или внутренней частью химерного адаптерного полипептида. На фиг. 3B представлена типичная компоновка различных компонентов рассматриваемой системы. Сайт распознавания расщепления 302b в GMP фланкирован взаимодействующим с антигеном доменом 301 и исполнительным элементом 302a, а расщепляющий элемент 304 входит в состав химерного адаптерного полипептида 303.

На фиг. 4A-D схематически представлено высвобождение исполнительного элемента из GMP. На фиг. 4A представлено связывание антигена с трансмембранным полипептидом химерного рецептора. Трансмембранный полипептид химерного рецептора включает внеклеточную область, содержащую взаимодействующий с антигеном домен 401, и внутриклеточную область, содержащую GMP. GMP включает в себя исполнительный элемент 402a, соединенный с сайтом распознавания при расщеплении 402b. В ответ на связывание антигена рецептор подвергается модификации путем фосфорилирования 403 во внутриклеточной области рецептора (фиг. 4B). После модификации (например, фосфорилирования) рецептора к рецептору привлекается адаптерный белок, содержащий участок связывания рецептора, как показано на фиг. 4C. Рецептор содержит расщепляющий элемент 404; расщепляющий элемент может образовывать комплекс с адаптером или соединяться, к примеру, пептидной связью и/или пептидным линкером с участком рецепторного связывания. Когда он находится в близости от сайта распознавания расщепления, расщепляющий элемент может расщеплять сайт распознавания, высвобождая исполнительный элемент из GMP, как показано на фиг. 4D. После высвобождения исполнительный элемент может поступать в ядро для регуляции экспрессии и/или активности целевого гена или для редактирования последовательности нуклеиновой кислоты. На фиг. 4E-H представлена аналогичная система, в которой модификация рецептора включает конформационное изменение. В некоторых воплощениях адаптерный белок привязан к мембране (например, в виде мембраносвязанного белка).

В некоторых воплощениях расщепляющий элемент расщепляет сайт распознавания только тогда, когда он находится в близости от сайта распознавания расщепления. Сайт распознавания расщепления может содержать полипептидную последовательность, которая является распознавательной последовательностью для протеазы. Расщепляющий элемент может обладать протеазной активностью, которая распознает полипептидную последовательность. Расщепляющий элемент, обладающий протеазной активностью, может представлять собой протеазу или её производное, вариант или фрагмент. Протеаза означает такой фермент, который осуществляет протеолиз, при котором полипептиды расщепляются на более мелкие полипептиды или аминокислоты. Для применения в качестве расщепляющего элемента подходят различные протеазы. Некоторые протеазы бывают очень неразборчивыми и гидролизуют широкий спектр белковых субстратов. Некоторые протеазы бывают высокоспецифичными и расщепляют субстраты только с определенной последовательностью, например последовательностью распознавания расщепления или доменом отщепления пептида. В некоторых воплощениях сайт распознавания расщеп-

ления содержит несколько последовательностей распознавания расщепления, причем каждая последовательность распознавания расщепления может распознаваться одним и тем же или разными расщепляющими элементами, обладающими протеазной активностью (например, протеазами). Специфичные к последовательности протеазы, которые можно использовать в качестве расщепляющих элементов, включают, без ограничения, суперсемейство протеаз CA, например, семейства C1, C2, C6, C10, C12, C16, C19, C28, C31, C32, C33, C39, C47, C51, C54, C58, C64, C65, C66, C67, C70, C71, C76, C78, C83, C85, C86, C87, C93, C96, C98 и C101, включая папаин (*Carica papaya*), бромелаин (*Ananas comosus*), катепсин К (печеночника) и кальпаин (*Homo sapiens*); суперсемейство протеаз CD, например, семейства C11, C13, C14, C25, C50, C80 и C84, как-то каспаза-1 (*Rattus norvegicus*) и сепараза (*Saccharomyces cerevisiae*); суперсемейство протеаз CE, например, семейства C5, C48, C55, C57, C63 и C79, включая аденаин (аденовируса 2 типа человека); суперсемейство протеаз CF, например, семейство C15, включая пироглутамилпептидазу I (*Bacillus amyloliquefaciens*); суперсемейство протеаз CL, например, семейства C60 и C82, включая сортазу А (*Staphylococcus aureus*); суперсемейство протеаз CM, например, семейство C18, включая пептидазу-2 вируса гепатита С (вируса гепатита С); суперсемейство протеаз CN, например, семейство C9, включая пептидазу nsP2 типа вируса Sindbis (вируса Sindbis); суперсемейство протеаз CO, например, семейство C40, включая дипептидилпептидазу VI (*Lysinibacillus sphaericus*); суперсемейство протеаз CP, например, семейство C97, включая пептидазу DeSI-1 (*Mus musculus*); суперсемейство протеаз PA, например, семейства C3, C4, C24, C30, C37, C62, C74 и C99, включая протеазу TEV (вируса гравировки табака); суперсемейство протеаз PB, например, семейства C44, C45, C59, C69, C89 и C95, включая предшественника амидофосфорилтрансферазы (*Homo sapiens*); суперсемейство протеаз PC, семейства C26 и C56, включая γ -глутамилгидролазу (*Rattus norvegicus*); суперсемейство протеаз PD, например, семейство C46, включая белок Hedgehog (*Drosophila melanogaster*); суперсемейство протеаз PE, например, семейство P1, включая аминокпептидазу DmpA (*Ochrobactrum anthropi*); и другие протеазы, например, семейства C7, C8, C21, C23, C27, C36, C42, C53 и C75. Дополнительные протеазы включают сериновые протеазы, например, из суперсемейства SB, например, семейства S8 и S53, включая субтилизин (*Bacillus licheniformis*); из суперсемейства SC, например, семейства S9, S10, S15, S28, S33 и S37, включая пролил-олигопептидазу (*Sus scrofa*); из суперсемейства SE, например, семейства S11, S12 и S13, включая D-Ala-D-Ala-пептидазу C (*Escherichia coli*); из суперсемейства SF, например, семейства S24 и S26, включая сигнальную пептидазу I (*Escherichia coli*); из суперсемейства SJ, например, семейства S16, S50 и S69, включая пептидазу IonA (*Escherichia coli*); из суперсемейства SK, например, семейства S14, S41 и S49, включая протеазу Clp (*Escherichia coli*); из суперсемейства SO, например, семейства S74, включая саморасщепляющий белок эндосиалидазы CIMCD фага K1F (фага K1F энтеробактерий); из суперсемейства SP, например, семейства S59, включая нуклеопорин 145 (*Homo sapiens*); из суперсемейства SR, например, семейства S60, включая лактоферрин (*Homo sapiens*); из суперсемейства SS, семейства S66, включая LD-карбокиспептидазу - тетрапептидазу муреина (*Pseudomonas aeruginosa*); из суперсемейства ST, например, семейства S54, включая ромбоид-1 (*Drosophila melanogaster*); из суперсемейства PA, например, семейства S1, S3, S6, S7, S29, S30, S31, S32, S39, S46, S55, S64, S65 и S75, включая химотрипсин А (*Bos taurus*); из суперсемейства PB, например, семейства S45 и S63, включая предшественника ацилазы пенициллина G (*Escherichia coli*); из суперсемейства PC, семейства S51, включая дипептидазу E (*Escherichia coli*); из суперсемейства PE, например, семейства P1, включая аминокпептидазу DmpA (*Ochrobactrum anthropi*); неопределенные, например, семейства S48, S62, S68, S71, S72, S79 и S81; треониновые протеазы, например, из суперсемейства PB, например, семейства T1, T2, T3 и T6, включая архейные протеасомы, β -компонент (*Thermoplasma acidophilum*); и из суперсемейства PE, например, семейства T5, включая орнитин-ацетилтрансферазу (*Saccharomyces cerevisiae*); аспарагиновые протеазы, например, BACE1, BACE2; катепсин D; катепсин E; химозин; напсин-А; непентезин; пепсин; плазмепсин; пресенилин; ренин; и протеаза ВИЧ-1, а также металлопротеиназы, например, экзопептидазы, металлоэкзопептидазы; эндопептидазы и металлоэндопептидазы. Последовательность распознавания расщепления (например, полипептидная последовательность) может распознаваться любой из описанных здесь протеаз.

В некоторых воплощениях сайт распознавания расщепления содержит последовательность распознавания расщепления (например, полипептидную последовательность или домен отщепления пептида), которая распознается протеазой, выбранной из группы, в которую входят: ахромопептидаза, аминокпептидаза, анкрод, ангиотензинпревращающий фермент, бромелаин, кальпаин, кальпаин I, кальпаин II, карбокиспептидаза А, карбокиспептидаза В, карбокиспептидаза G, карбокиспептидаза Р, карбокиспептидаза W, карбокиспептидаза Y, каспаза 1, каспаза 2, каспаза 3, каспаза 4, каспаза 5, каспаза 6, каспаза 7, каспаза 8, каспаза 9, каспаза 10, каспаза 11, каспаза 12, каспаза 13, катепсин В, катепсин С, катепсин D, катепсин Е, катепсин G, катепсин H, катепсин L, химопапаин, химаза, химотрипсин, кластрипаин, коллагеназа, комплемент C1r, комплемент C1s, фактор комплемента D, фактор комплемента I, кукумизин, дипептидилпептидаза IV, эластаза (лейкоцитарная), эластаза (панкреатическая), эндопротеиназа Arg-C, эндопротеиназа Asp-N, эндопротеиназа Glu-C, эндопротеиназа Lys-C, энтерокиназа, фактор Ха, фицин, фузин, гранзим А, гранзим В, протеаза ВИЧ, I Газа, тканевой калликреин, лейцин-аминокпептидаза (общая), лейцин-аминокпептидаза (цитозольная), лейцин-аминокпептидаза (микросомальная), матриксная металло-

протеаза, метионин-аминопептидаза, нейтраз, папаин, пепсин, плазмин, пролидаза, проназа Е, специфический антиген простаты, протеаза алкалофильная из *Streptomyces griseus*, протеаза из *Aspergillus*, протеаза из *Aspergillus saitoi*, протеаза из *Aspergillus sojae*, протеаза из *B.licheniformis* (щелочная или алкалаза), протеаза из *Bacillus polymyxa*, протеаза из *Bacillus sp.*, протеаза из *Rhizopus sp.*, протеаза S, протеасомы, протеиназа из *Aspergillus oryzae*, протеиназа З, протеиназа А, протеиназа К, белок С, пироглутамат-аминопептидаза, ренин, ренин, стрептокиназа, субтилизин, термолизин, тромбин, активатор тканевого плазминогена, трипсин, триптаза и урокиназа.

В табл. 1 перечислены типичные протеазы и соответствующие им распознавательные последовательности, которые могут использоваться в системах по изобретению.

Таблица 1

Типичные протеазы и соответствующие им распознающие последовательности

Название протеазы	Синонимы	Расознавательная последовательность
Arg-C	аргинил-пептидаза, эндопротеиназа Arg-C, тканевой калликреин	R-x
Asp-N	эндопротеиназа Asp-N, пептидил-Asp-металлоэндопептидаза	x-D
Asp-N (N-концевой Glu)	эндопротеиназа Asp-N, пептидил-Asp-металлоэндопептидаза	x-[DE]
BNPS или NCS/мочевина	3-бром-3-метил-2-(2-нитрофенилтио)-3Н-индол, BNPS-скатол, N-хлорсукцинимид/мочевина	W-x
Каспаза-1	ICE, интерлейкин-1 β -превращающий фермент	[FLWY]-x-[AHT]-D- {DEKPQR} (SEQ ID NO: 61)
Каспаза-10	Flice2, Mch4	I-E-A-D-x (SEQ ID NO: 62)
Каспаза-2	Ich-1, Nedd2	D-V-A-D- {DEKPQR} (SEQ ID NO: 63) или D-E-H-D- {DEKPQR} (SEQ ID NO: 64)
Каспаза-3	апопаин, CPP32, Yama	D-M-Q-D- {DEKPQR} (SEQ ID NO: 65) или D-E- V-D- {DEKPQR} (SEQ ID NO: 66)

Каспаза-4	ICE(rel)II, Ich-2, TX	L-E-V-D-{\DEKPQR} (SEQ ID NO: 67) или [LW]-E-H-D-{\DEKPQR} (SEQ ID NO: 68)
Каспаза-5	ICE(rel)III, TY	[LW]-E-H-D-x (SEQ ID NO: 69)
Каспаза-6	Mch2	V-E-[HI]-D-{\DEKPQR} (SEQ ID NO: 70)
Каспаза-7	CMH-1, ICE-LAP3, Mch-3	D-E-V-D-{\DEKPQR}(SEQ ID NO: 71)
Каспаза-8	FLICE, MASH, Mch5	[IL]-E-T-D-{\DEKPQR}(SEQ ID NO: 72)
Каспаза-9	ICE-Lap6, Mch6	L-E-H-D-x (SEQ ID NO: 73)
Химотрипсин		[FY]-{P} или W-{MP}
Химотрипсин (низкая специфичность)		[FLY]-{P} или W-{MP} или M-{PY} или H-{DMPW}
Клострипан	клостридиопептидаза В	R-x
CNBr	Цианогенбромид	M-x
CNBr (метил-Cys)	Цианогенбромид	M-x или x-C
CNBr (с кислотами)	Цианогенбромид	[MW]-x
Энтерокиназа	Энтеропептидаза	[DE](4)-K-x (SEQ ID NO: 74)
фактор Ха	коагуляционный фактор Ха	[AFGILTVM]-[DE]-G-R-x (SEQ ID NO: 75)
Муравьиная кислота		D-x
Glu-C (AmAc буфер)	эндопротеиназа Glu-C, протеаза V8, глутамил-эндопептидаза	E-x
Glu-C (Phos буфер)	эндопротеиназа Glu-C, протетазы V8, глутамил-эндопептидаза	[DE]-x
Гранзим В	протеиназа-2 цитотоксических Т-лимфоцитов, гранзим-2, гранзим В, лимфоцитарная протеаза, SECT, Т-клеточная сериновая протеаза 1-3E	I-E-P-D-x (SEQ ID NO: 76)
Протеаза HRV3C	протеаза 3С риновируса человека, пикорнаин 3С, протеаза 3С	L-E-V-L-F-Q-G-P (SEQ ID NO: 77)
Гидроксиламин	гидроксиламмония хлорид	N-G
Йодбензойная кислота	2-йодбензойная кислота	W-x
Lys-C	эндопротеиназа Lys-C, лизил-эндопептидаза	K-x
Lys-N	эндопротеиназа Lys-N, пептидил-Lys-металлоэндопептидаза, нейтральная протеиназа <i>Armillaria mellea</i>	x-K
Lys-N (Cys-модифицированная)	эндопротеиназа Lys-N, пептидил-Lys-металлоэндопептидаза, нейтральная протеиназа <i>Armillaria mellea</i>	x-[CK]
Мягкий кислотный гидролиз		D-P
NBS (длительное воздействие)	N-бромсукцинимид	[HWY]-x
NBS (короткое воздействие)	N-бромсукцинимид	[WY]-x

NTCB	2-нитро-5-тиоцианатобензойная кислота, 2-нитро-5-тиоцианобензойная кислота	x-C
Панкреатическая эластаза	панкреопептидаза E, эластаза-1	[AGSV]-x
Пепсин А	Пепсин	{HKR}-{P}-{R}-[FLWY]- {P} (SEQ ID NO: 78) или {HKR}-{P}-[FLWY]-x-{P} (SEQ ID NO: 79)
Пепсин А (низкая специфичность)	Пепсин	{HKR}-{P}-{R}-[FL]-{P} (SEQ ID NO: 80) или {HKR}-{P}-[FL]-x-{P} (SEQ ID NO: 81)
Пролил-эндопептидаза	пролил-олигопептидаза, фермент расщепления после пролина	[HKR]-P-{P}
Протеиназа К	эндопептидаза К, пептидаза К	[AEFILTVMWY]-x
Протеаза TEV	протеаза вируса гравировки табака, эндопептидаза ядерных включений-а	E-x-x-Y-x-Q-[GS] (SEQ ID NO: 82)
Термолизин	термофильная бактериальная протеаза	{DE}-[AFILMV]-{P}
Тромбин	фактор IIa	x-x-G-R-G-x (SEQ ID NO: 83) или [AFGILTVW]- [AFGILTVW]-P-R-{DE}- {DE} (SEQ ID NO: 84)
Трипсин	трипсин-1	x-[KR]-{P} или W-K-P или M-R-P, но не: [CD]-K-D или C-K-[HY] или C-R-K или R-R-[HR]
Трипсин (Arg-блокированный)		K-{P}
Трипсин (Cys-модифицированный)		[RKC]-{P}
Трипсин (Lys-блокированный)		R-{P}

При выборе протеаз для применения в качестве расщепляющих элементов их можно выбирать по желательным характеристикам, таким как селективность по пептидной связи, активность при определенных значениях pH, молекулярная масса и т.п. В табл. 2 приведены свойства типичных протеаз.

Таблица 2

Типичные протеазы и характеристики протеаз

Протеаза	Номер по КФ	Класс	Селективность по пептидной связи	pH-оптимум	Молек. масса (кДа)	Номер доступа
Эндопротейназы						
Трипсин (бычий)	3.4.21.4	сериновая	P ₁ -P ₁ ¹ - (P ₁ = Lys, Arg)	8,0-9,0	23,5	P00760 ^S
Химотрипсин (бычий)	3.4.21.1	сериновая	P ₁ -P ₁ ¹ - (P ₁ = ароматическая, P ₁ ¹ = неспецифическая)	7,5-8,5	25	P00766 ^S
Эндопротеиназа Asp-N (Pseudomonas fragi)	3.4.24.33	металло	P ₁ -Asp- (и -P ₁ -цистеиновая кислота)	6,0-8,0	27	φ

Эндопротеиназа Arg-C (подчелюстная железа мыши)	φ	сериновая	-Arg-P ₁ -	8,0-8,5	30	н/п
Эндопротеиназа Glu-C (протеаза V8) (Staphylococcus aureus)	3.4.21.19	сериновая	-Glu-P ₁ ¹ - (и -Asp-P ₁ ¹ -) (2)	8,0	27	P04188 ^S
Эндопротеиназа Lys-C (Lysobacter enzymogenes)	3.4.21.50	сериновая	-Lys-P ₁ ¹ -	8,0	30 ^{NR} 33 ^R	S77957 ^P
Пепсин (свиной)	3.4.23.1	аспартатная	P ₁ -P ₁ ¹ - (P ₁ = предпочтительно гидрофобная)	2,0-4,0	34,5	P00791 ^S
Термолизин (Bacillus thermoproteolyticus)	3.4.24.27	металло	P ₁ -P ₁ ¹ - (P ₁ = Leu, Phe, Ile, Val, Met, Ala)	7,0-9,0	37,5	P00800 ^S
Эластаза (свиная)	3.4.21.36	сериновая	P ₁ -P ₁ ¹ - (P ₁ = незаряженная, неароматическая)	7,8-8,5	25,9	P00772 ^S
Папаин (Carica papaya)	3.4.22.2	цистеиновая	P ₁ -P ₁ ¹ - (P ₁ = предпочтительно Arg, Lys)	6,0-7,0	23	P00784 ^S
Протеиназа К (Tritirachium album)	3.4.21.64	сериновая	P ₁ -P ₁ ¹ - (P ₁ = предпочтительно ароматическая, гидрофобная)	7,5-12,0	18,5	P06873 ^S
Субтилизин (Bacillus subtilis)	3.4.21.62	сериновая	P ₁ -P ₁ ¹ - (P ₁ = предпочтительно нейтральная/кислая)	7,0-11,0	30 ^S 27,3 ^L	P04189 ^S
Клострипаин (эндопротеиназа Arg-C) (Clostridium histolyticum)	3.4.22.8	цистеиновая	-Arg-P ₁ - (P ₁ = Pro предпочтительно)	7,1-7,6	59	P09870 ^S
Экзопептидазы						
Карбоксипептидаза А (бычья)	3.4.17.1	металло	P ₁ -P ₁ ¹ - (P ₁ не может быть Arg, Lys, Pro)	7,0-8,0	34,5	P00730 ^S
Карбоксипептидаза В (свиная)	3.4.17.2	металло	P ₁ -P ₁ ¹ - (P ₁ = Lys, Arg)	7,0-9,0	34,6	P00732 ^S
Карбоксипептидаза Р (Penicillium janthinellum)	φ	сериновая	P ₁ -P ₁ ¹ - (неспецифически)	4,0-5,0	51	н/п
Карбоксипептидаза Y (дрожжевая)	3.4.16.5	сериновая	P ₁ -P ₁ ¹ - (неспецифически)	5,5-6,5	61	P00729 ^S
Катепсин С	3.4.14.1	цистеиновая	X-P ₁ -P ₁ ¹ - (удаляет N-концевой дипептид)	5,5	210	н/п
Ациламиноацилпептидаза (свиная)	3.4.19.1	сериновая	Ac-P ₁ -P ₁ ¹ - (P ₁ = предпочтительно Ser, Ala, Met)	7,5	80 ^H 360 ^P	P19205 ^{S+}
Пироглутамат-аминопептидаза (бычья)	3.4.19.3	цистеиновая	P ₁ -P ₁ ¹ - (P ₁ = 5-оксопролин или пироглутамат)	7,0-9,0	70-80 ^B	н/п

В некоторых воплощениях сайт распознавания расщепления содержит первую часть последовательности интеина, которая реагирует со второй частью последовательности интеина с высвобождением исполнительного элемента. Для облегчения высвобождения исполнительного элемента из химерного рецепторного полипептида может использоваться гетерологическая система с разделенным интеином. Исполнительный элемент может ковалентно соединяться с первой частью последовательности интеина. Исполнительный элемент может соединяться через свой N-конец или C-конец с первой частью последовательности интеина. Вторая часть последовательности интеина может входить в состав химерного адаптерного полипептида. Вторая часть последовательности интеина может служить расщепляющим элементом. Первой частью или второй частью последовательности интеина может быть N-концевой интеин, C-концевой интеин или любая другая подходящая часть интеина, которая может способствовать высвобождению исполнительного элемента. Последовательности интеина могут происходить из любого подхо-

дящего источника. Первая и вторая часть могут быть из одного источника (например, организма, белка) или из разных источников.

В одном показательном примере, представленном на фиг. 13А, химерный рецепторный полипептид содержит исполнительный элемент 1301, ковалентно связанный (например, на своем N-конце или С-конце) пептидной связью с первой частью последовательности интеина 1302, содержащей N-конец интеина. Слитый с N-концом интеина исполнительный элемент может входить в контакт со второй частью последовательности интеина 1303, содержащей С-конец интеина, как показано на фиг. 13В, к примеру, со второй частью последовательности интеина, связанной с адаптерным полипептидом. При таком контакте между первой и второй частью последовательности интеина может происходить сайт-специфичное расщепление (например, на участке между исполнительным элементом и N-концом интеина), как показано на фиг. 13С, при этом высвобождается исполнительный элемент, как показано на фиг. 13D. В альтернативной конфигурации, представленной на фиг. 13Е-Н, исполнительный элемент связан и/или закомплексован с адаптерным полипептидом, а не с рецепторным полипептидом. В другом показательном примере исполнительный элемент может быть ковалентно связан (например, на своем N-конце или С-конце) пептидной связью с первой частью интеина, содержащей С-конец интеина. Слитый с С-концом интеина исполнительный элемент может входить в контакт со второй частью последовательности интеина, содержащей N-конец интеина. При таком контакте между первой и второй частью интеина может происходить сайт-специфичное расщепление (например, на подходящем участке между исполнительным элементом и С-концом интеина), при этом высвобождается исполнительный элемент.

В некоторых воплощениях сайт распознавания расщепления содержит дисульфидную связь. Дисульфидная связь может соединять исполнительный элемент с химерным рецепторным полипептидом. Дисульфидная связь может образовываться между одним или несколькими цистеинами исполнительного элемента и рецептором. Цистеины могут быть встроены в состав исполнительного элемента или рецептора. Цистеины могут входить в состав нативной или последовательности дикого типа. Цистеины могут присутствовать в линкерном пептиде, присоединенном к исполнительному элементу или рецептору. Расщеплению дисульфидной связи может способствовать, к примеру, изменение окислительно-восстановительных условий у дисульфидной связи. Изменение окислительно-восстановительных условий может приводить к восстановлению дисульфидной связи до тиолов и высвобождению исполнительного элемента. Расщеплению дисульфидной связи может способствовать расщепляющий элемент, содержащий окислительно-восстановительное средство, которое может катализировать восстановление дисульфидной связи. Окислительно-восстановительным средством может быть фермент либо его производное, вариант или фрагмент. Фермент может быть оксидоредуктазой. Примеры оксидоредуктаз включают протеин-дисульфидредуктазы, тиоредоксины, глутаредоксины, тиол-дисульфидоксидоредуктазы (например, DsbA, VdbA-D, MdbA и SdbA) и глутатион-дисульфидредуктазы. Окислительно-восстановительное средство может быть из любого подходящего источника, включая прокариотов и эукариотов. Для оптимальной активности фермента могут потребоваться кофакторы (например, никотинамидные кофакторы, флавины либо их производные и аналоги).

В одном показательном примере, представленном на фиг. 14А, химерный рецепторный полипептид содержит исполнительный элемент 1401, соединенный дисульфидной связью. Дисульфидная связь может расщепляться расщепляющим элементом 1402, содержащим фермент типа оксидоредуктазы, к примеру, оксидоредуктазы, закомплексованной и/или связанной с адаптерным полипептидом, как показано на фиг. 14В. При расщеплении дисульфидной связи может высвобождаться исполнительный элемент, как показано на фиг. 14С. Исполнительный элемент после высвобождения может транслоцироваться в ядро клетки, где он может регулировать экспрессию и/или активность гена или редактировать последовательность нуклеиновой кислоты, как показано на фиг. 14D. На фиг. 14Е-Н представлена альтернативная конфигурация, в которой исполнительный элемент закомплексован и/или связан с адаптерным полипептидом, а расщепляющий элемент (например, оксидоредуктаза) связан с рецептором.

В некоторых воплощениях химерный рецепторный полипептид содержит по меньшей мере одну нацеливающую последовательность, которая направляет транспорт рецептора в определенный район клетки. Нацеливающая последовательность может использоваться для направления транспортировки полипептида, с которым связана нацеливающая последовательность, в определенный район клетки. К примеру, нацеливающая последовательность может направлять рецептор в ядро клетки, используя сигнал ядерной локализации (NLS), за пределы ядра (например, в цитоплазму), используя сигнал ядерного экспорта (NES), в митохондрии, эндоплазматический ретикулум (ER), аппарат Гольджи, хлоропласты, апопласты, пероксисомы, плазматические мембраны или мембраны различных органелл клетки. В некоторых воплощениях нацеливающая последовательность содержит сигнал ядерного экспорта (NES) и направляет полипептид за пределы ядра, к примеру, в цитоплазму клетки. Нацеливающая последовательность может направлять полипептид в цитоплазму с использованием различных сигналов ядерного экспорта. Обычно сигналы ядерного экспорта представляют собой короткие аминокислотные последовательности из гидрофобных остатков (например, по меньшей мере 2, 3, 4 или 5 гидрофобных остатков), которые направляют белок на экспорт из ядра клетки в цитоплазму через комплекс ядерных пор с помощью ядерного транспорта. Не все субстраты NES могут конститутивно экспортироваться из ядра. В не-

которых воплощениях нацеливающая последовательность содержит сигнал ядерной локализации (NLS, например, NLS SV40) и направляет полипептид в ядро клетки. Нацеливающая последовательность может направлять полипептид в ядро клетки, используя различные сигналы ядерной локализации (NLS). Последовательность NLS может состоять из одной части или из двух частей.

Неограничительные примеры NLS включают последовательности NLS, происходящие из: NLS большого Т-антигена вируса SV40 с аминокислотной последовательностью PKKKRKV (SEQ ID NO: 40); NLS из нуклеоплазмина (например, двухчастного NLS нуклеоплазмина с последовательностью KRPAATKKAGQAKKKK (SEQ ID NO: 41)); NLS с-мус с аминокислотной последовательностью PA-AKRVKLD (SEQ ID NO: 42) или RQRRNELKRSP (SEQ ID NO: 43); NLS hRNPA1 M9 с последовательностью NQSSNFGPMKGGNFGGRSSGPYGGGGYYFAKPRNQGGY (SEQ ID NO: 44); последовательности RMRIZFKNKGKDTAELRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 45) из домена IBB альфа-импортина; последовательности VSRKRPRP (SEQ ID NO: 46) и PPKKARED (SEQ ID NO: 47) из Т-белка миомы; последовательности PPKKKPL (SEQ ID NO: 48) из р53 человека; последовательности SALDCKKKKMAP (SEQ ID NO: 49) из с-abl IV мыши; последовательности DRLRR (SEQ ID NO: 50) и PKQKKRK (SEQ ID NO: 51) из NS1 вируса гриппа; последовательности RKLKKIKKL (SEQ ID NO: 52) из дельта-антигена вируса гепатита; последовательности REKKKFLKRR (SEQ ID NO: 53) из белка Mx1 мыши; последовательности KRKGDEVDGVDEVAKKKKK (SEQ ID NO: 54) из полимеразы поли(ADP-рибозы) человека; и последовательности RKCLQAGMNLEARKTKK (SEQ ID NO: 55) из глюкокортикоидного рецептора стероидных гормонов человека.

В некоторых воплощениях нацеливающая последовательность включает нацеливающий на мембрану пептид и направляет полипептид в плазматическую мембрану или мембрану клеточной органеллы. Нацеливающая на мембрану последовательность может обеспечивать транспорт химерного полипептида трансмембранного рецептора на поверхностную клеточную мембрану или на другую клеточную мембрану. Связанные с клеточными мембранами молекулы содержат определенные участки, способствующие мембранной ассоциации, и такие участки могут быть включены в наводящую на мембрану последовательность. Например, некоторые белки содержат последовательности на N-конце или C-конце, которые ацилированы, и эти ацильные группировки способствуют связи с мембраной. Такие последовательности могут распознаваться ацилтрансферазами и зачастую соответствуют определенному мотиву в последовательности. Некоторые ацилирующие мотивы могут быть модифицированы одной ацильной группировкой (зачастую за ней следуют несколько положительно заряженных остатков (к примеру, с-Src человека) для улучшения ассоциации с анионными "головками" липидов), а другие могут быть модифицированы несколькими ацильными группировками. Например, N-концевая последовательность белковой тирозинкиназы Src может содержать одну миристоиловую группировку. Двойные участки ацилирования располагаются в N-концевых районах некоторых протеинкиназ типа подгруппы представителей семейства Src (например, Yes, Fyn, Lck) и α -субъединицы G-белков. Такие двойные участки ацилирования часто располагаются в пределах первых 18 аминокислот таких белков и соответствуют мотиву последовательности Met-Gly-Cys-Xaa-Cys (SEQ ID NO: 86), причем Met отщепляется, Gly подвергается N-ацилированию, а один из остатков Cys подвергается S-ацилированию. Зачастую Gly подвергается миристоилрованию, а Cys - пальмитоилрованию. Также могут использоваться участки ацилирования, соответствующие мотиву последовательности Cys-Ala-Ala-Xaa (так называемой "рамки СААХ"), которые могут быть модифицированы группировкой C15- или C10-изопренила, из C-концевых участков γ -субъединиц G-белков и других белков. Эти и другие ацилирующие мотивы включают, к примеру, приведенные в Gauthier-Campbell et al., *Molecular Biology of the Cell* 15: 2205-2217 (2004); Glabati et al., *Biochem. J.* 303: 697-700 (1994); и Zlakine et al., *J. Cell Science* 110: 673-679 (1997), и могут быть включены в наводящие последовательности, чтобы индуцировать мембранную локализацию.

В некоторых воплощениях в нацеливающую последовательность включена нативная последовательность из белка, содержащего ацилирующий мотив. Например, в некоторых воплощениях в N-концевой участок химерного полипептида может быть включена N-концевая часть Lck, Fyn или Yes или альфа-субъединицы G-белка, как-то первые 25 N-концевых аминокислот такого белка или меньше (например, от 5 до 20 аминокислот, от 10 до 19 аминокислот или от 15 до 19 аминокислот из нативной последовательности, необязательно с мутациями). В некоторых воплощениях с C-концом химерного полипептида может быть связана C-концевая последовательность из 25 или менее аминокислот из γ -субъединицы G-белка, содержащей последовательность мотива рамки СААХ (например, от 5 до 20 аминокислот, от 10 до 18 аминокислот или от 15 до 18 аминокислот из нативной последовательности, необязательно с мутациями).

Может использоваться любая нацеливающая на мембрану последовательность. В некоторых воплощениях такие последовательности включают, без ограничения, наводящие последовательности миристоилрования, наводящие последовательности пальмитоилрования, последовательности пренилирования (например, фарнезилирования, геранил-геранилирования, рамки СААХ), мотивы межбелкового взаимодействия или трансмембранные последовательности (с использованием сигнальных пептидов) из рецепторов. Примеры включают приведенные, к примеру, в ten Klooster J.P. et al., *Biology of Cell* (2007)

99, 1-12; Vincent S. et al., *Nature Biotechnology* 21: 936-40, 1098 (2003).

Существуют и другие белковые домены, которые могут усиливать удержание белка в различных мембранах. Например, домен гомологичности плекстрина (PH) в ~120 аминокислот встречается у более чем 200 белков человека, которые обычно участвуют во внутриклеточной сигнализации. PH-домены могут связываться с различными липидами фосфатидилинозитола (PI) в мембранах (например, PI (3,4,5)-P3, PI (3,4)-P2, PI (4,5)-P2) и тем самым играть ключевую роль в рекрутинге белков в различные мембранные или клеточные компартменты. Часто уровень фосфорилирования PI-липидов регулируется, к примеру, киназой PI-3 или PTEN, поэтому взаимодействие мембран с PH-доменами может быть не таким стабильным, как при ацилировании липидов.

В некоторых воплощениях в наводящих последовательностях, направляющих полипептиды в клеточную мембрану, может использоваться сигнальная последовательность мембранного якорения. Имеются различные последовательности мембранного якорения. Например, можно использовать сигнальные последовательности мембранного якорения из различных мембраносвязанных белков. Последовательности могут включать в себя таковые из: 1) интегральных мембранных белков I класса типа β -цепи рецептора IL-2 и β -цепи рецептора инсулина; 2) интегральных мембранных белков II класса типа нейтральной эндопептидазы; 3) белков III типа, таких как цитохром P450 NF-25 человека; и 4) белков IV типа, таких как Р-гликопротеин человека.

В некоторых воплощениях химерный рецепторный полипептид связан с доменом фолдинга полипептидов, который может способствовать сворачиванию белка. В некоторых воплощениях исполнительный элемент связан с доменом проникновения в клетки. Например, домен проникновения в клетки может происходить из белка TAT HIV-1, мотива проникновения в клетки TLM из вируса гепатита В человека, MPG, Rep-1, VP22, проникающего в клетки пептида вируса Herpes simplex или последовательности полиар-генинового пептида. Проникающий в клетки домен может располагаться на N-конце, C-конце или же где-то внутри исполнительного элемента.

Нацеливающая последовательность может быть связана с любым подходящим участком химерного рецепторного полипептида, к примеру, на N-конце, C-конце или во внутренней области рецептора. В некоторых воплощениях с рецептором связаны по меньшей мере две наводящие последовательности. У типичного химерного рецепторного полипептида, представленного на фиг. 5, первая нацеливающая последовательность 501a может быть связана с внеклеточной областью рецептора, а вторая нацеливающая последовательность 501b может быть связана с внутриклеточной областью рецептора типа GMP. Когда рецептор связан с несколькими наводящими последовательностями, к примеру, с наводящими последовательностями, направляющими в разные компартменты клетки, то окончательная локализация рецептора может определяться относительной силой наводящих последовательностей. Например, рецептор, содержащий и нацеливающую последовательность, включающую NES, и нацеливающую последовательность, включающую NLS, может локализоваться в цитоплазме, если NES будет сильнее, чем NLS. С другой стороны, если NLS будет сильнее, чем NES, то рецептор может локализоваться в ядре, даже если у рецептора будет присутствовать как сигнал ядерной локализации, так и сигнал ядерного экспорта. Нацеливающая последовательность может содержать несколько копий, к примеру, каждого из NLS и NES, для точной настройки клеточной локализации.

В некоторых случаях нацеливающая последовательность связана с исполнительным элементом. После высвобождения исполнительного элемента из GMP (и рецептора) при расщеплении сайта распознавания расщепления нацеливающая последовательность может направить исполнительный элемент в такое место клетки, которое отличается от рецептора. Например, химерный трансмембранный рецептор может содержать первую нацеливающую последовательность, которая направляет рецептор на плазматическую мембрану, а исполнительный элемент может отдельно содержать вторую нацеливающую последовательность, которая направляет локализацию в ядро клетки. Первоначально исполнительный элемент (входящий в состав рецептора) может локализоваться на плазматической мембране из-за первой нацеливающей последовательности. После высвобождения исполнительного элемента из GMP при расщеплении сайта распознавания расщепления исполнительный элемент может локализоваться в ядре клетки под действием целевой нацеливающей последовательности. В некоторых воплощениях исполнительный элемент транслоцируется в ядро клетки после расщепления сайта распознавания расщепления.

Связывание химерного адаптерного полипептида с химерным рецепторным полипептидом, когда рецепторный полипептид подвергается модификации при связывании с антигеном, может привести расщепляющий элемент в близость с сайтом распознавания расщепления. При расщеплении сайта распознавания исполнительный элемент может высвободиться из GMP. После высвобождения исполнительный элемент может образовывать комплекс с целевым полинуклеотидом, к примеру, в цитоплазме клетки или в ядре клетки. Образование комплекса исполнительного элемента с целевым полинуклеотидом может регулировать экспрессию и/или активность по меньшей мере одного гена или редактировать последовательность нуклеиновой кислоты.

В другой типичной конфигурации GMP входит в состав химерного адаптерного полипептида, а расщепляющий элемент входит в состав химерного рецепторного полипептида. Химерный адаптерный

полипептид в типичной конфигурации может содержать (а) участок связывания рецептора, который связывается с рецептором, подвергшимся модификации при связывании с антигеном; и (b) полипептид, модулирующий ген (GMP), связанный с участком рецепторного связывания, причем GMP содержит исполнительный элемент, связанный с сайтом распознавания расщепления; причем (i) сайт распознавания расщепления расщепляется расщепляющим элементом в ответ на связывание рецептора, а (ii) исполнительный элемент образует комплекс с целевым полинуклеотидом в ответ на расщепление сайта распознавания расщепления. На фиг. 6А представлен типичный химерный адаптерный полипептид. Химерный адаптерный полипептид может содержать участок связывания рецептора 601, связанный с GMP 602. GMP может содержать исполнительный элемент 602a, связанный с сайтом распознавания расщепления 602b.

Участком рецепторного связывания у химерного адаптерного полипептида может быть любой партнер по связыванию (например, белок), который может связываться с рецептором, либо его производное, вариант или фрагмент. В некоторых воплощениях адаптер содержит такого партнера по связыванию с рецептором, который связан с мембраной, или его производное, вариант либо фрагмент. В некоторых воплощениях адаптер содержит такого партнера по связыванию, который не связан с мембраной (например, внутриклеточный или цитозольный), либо его производное, вариант или фрагмент. Адаптерный полипептид может содержать домен рецепторного связывания сигнального белка или другого белка, рекрутирующегося к рецептору. Химерный адаптерный полипептид может рекрутироваться к рецептору в ответ на модификацию рецептора, например, конформационное изменение, химическую модификацию либо их комбинацию. Рецептор может подвергаться модификации в ответ на связывание лиганда. Рецепторы либо их производные, варианты или фрагменты и партнеры по связыванию (например, белки) либо их производные, варианты или фрагменты можно выбирать так, чтобы оптимизировать требуемый уровень рекрутинга адаптерного полипептида к рецептору.

В некоторых воплощениях химерный адаптерный полипептид содержит молекулу (например, белок) либо её производное, вариант или фрагмент, который привлекается к рецептору Notch при связывании рецептора Notch с лигандом. Химерный адаптерный полипептид может содержать белок, его производное, вариант или фрагмент, выбранный из группы, в которую входят пресенилин-1 (PSEN1), никастрин, продукт гена anterior pharynx-defective 1 (APH-1) и усилитель пресенилина 2 (PEN-2).

В некоторых воплощениях химерный адаптерный полипептид содержит молекулу (например, белок) либо её производное, вариант или фрагмент, который привлекается к GPCR при связывании GPCR с лигандом (например, к связанному с лигандом GPCR, претерпевшему конформационную и/или биохимическую модификацию). Химерный адаптерный полипептид может содержать белок, его производное, вариант или фрагмент, выбранный из группы, в которую входят AKAP79 (AKAP5) и AKAP250 (AKAP12, гравин), аррестин (например, β -аррестин), ATBP50, кальмодулин, DRIP78 (DNAJC14), Homer, GASP1, GEC1 (GABARAPL1), INAD, JAK2, LARG (ARHGEF12), MAGI2, MAGI3, M10 MHC, MPP3, MRAP и MRAP2, MUPP1 (MPDZ), нейрохондрин, NHERF1 (EBP50, SLC9A3R1), NHERF2 (SLC9A3R2), NINAA, ODR4, p85, PDZ-RhoGEF (ARHGEF11), периплакин, PICK1, PSD95, RACK1 (GNB2L1), RAMP1, RAMP2, RAMP3, RanBP2, REEP, RTP, RTP4, Shank, SNX1, синтрофин, спинофилин, TCTEXT1 (DYNLT1) и USP4.

В некоторых воплощениях химерный адаптерный полипептид содержит молекулу (например, белок) либо её производное, вариант или фрагмент, который привлекается к интегриновому рецептору при связывании его с лигандом. Примеры адаптерных белков, которые рекрутируются к интегриновым рецепторам, включают, без ограничения, структурные адаптерные белки, скелетные адаптерные белки и адаптерные белки, обладающие каталитической активностью. В некоторых воплощениях химерный адаптерный полипептид содержит белок либо его производное, вариант или фрагмент, выбранный из талина, киндлина, филамина и тензина. В некоторых воплощениях химерный адаптерный полипептид содержит белок либо его производное, вариант или фрагмент, выбранный из паксиллина и киндлина. В некоторых воплощениях химерный адаптерный полипептид содержит белок либо его производное, вариант или фрагмент, выбранный из киназы фокальных контактов (FAK), Src и протеинфосфатазы 2A (PP2A). В некоторых воплощениях химерный адаптерный полипептид содержит белок либо его производное, вариант или фрагмент, выбранный из RAB21, PTPN2, AUP1, BIN1, COL8A1 и ITGB1.

В некоторых воплощениях химерный адаптерный полипептид содержит молекулу (например, белок) либо её производное, вариант или фрагмент, который привлекается к рецептору кадгерина. Молекула может рекрутироваться к рецептору в результате модификации рецептора (например, химической модификации, например, фосфорилирования, и/или конформационного изменения). В некоторых воплощениях химерный адаптерный полипептид содержит белок либо его производное, вариант или фрагмент, выбранный из α -катенина, β -катенина, γ -катенина, катенина δ 1 (катенина p120), AJAP1, CTNND1, DLGAP5, TBC1D2, LIMA1, CAV1, TRPV4, CTNNB1, PIP5K1C, RAB8B, RAPGEF2, DDR1, PSEN1, CDH1, CDC27, CTNNA1 и EGFR.

В некоторых воплощениях химерный адаптерный полипептид содержит молекулу (например, белок) либо её производное, вариант или фрагмент, который привлекается к химерному рецепторному по-

липептиду, содержащему RTSK либо её производное, вариант или фрагмент. В некоторых воплощениях химерный адаптерный полипептид содержит белок либо его производное, вариант или фрагмент, выбранный из представителей семейства SMAD, включая SMAD1, SMAD2, SMAD3, SMAD5, SMAD6, SMAD7 и SMAD9 (иногда называется SMAD8); якорных SMAD для активации рецептора (SARA); белков SMURF (например, SMURF1, SMURF2); и их производных, вариантов или фрагментов. В некоторых воплощениях химерный адаптерный полипептид содержит молекулу (например, белок), которая привлекается к химерному рецепторному полипептиду, содержащему цитокиновый рецептор либо его производное, вариант или фрагмент. В некоторых воплощениях адаптерный полипептид содержит gp130, CD131, CD132 либо его производное, вариант или фрагмент.

В некоторых воплощениях химерный адаптерный полипептид содержит молекулу, которая привлекается к фосфорилированной RTK или RTSK либо к рецептору, фосфорилированному нековалентно связанной внутриклеточной киназой. Фосфорилирование определенных аминокислотных остатков (например, остатков тирозина) у активированного рецептора (например, химерного рецепторного полипептида) может создавать сайты связывания для таких молекул, как белки, содержащие домен гомологии-2 Src (SH2) и связывающий фосфотирозин домен (PTB). В некоторых воплощениях адаптерный полипептид включает белок, содержащий домен SH2, как-то ABL1, ABL2, BCAR3, BLK, BLNK, BMX, BTK, CHN2, CISH, CRK, CRKL, CSK, DAPP1, EAT-2, FER, FES, FGR, FRK, FYN, GADS, GRAP, GRAP2, GRB10, GRB14, GRB2, GRB7, HCK, HSH2D, INPP5D, INPPL1, ITK, JAK2, LCK, LCP2, LYN, MATK, NCK1, NCK2, PIK3R1, PIK3R2, PIK3R3, PLCG1, PLCG2, PTK6, PTPN11, PTPN6, RASA1, SAP, SH2B1, SH2B2, SH2B3, SH2D1A, SH2D1B, SH2D2A, SH2D3A, SH2D3C, SH2D4A, SH2D4B, SH2D5, SH2D6, SH3BP2, SHB, SHC1, SHC2, SHC3, SHC4, SHD, SHE, SHP1, SHP2, SLA, SLA2, SOCS1, SOCS2, SOCS3, SOCS4, SOCS5, SOCS6, SOCS7, SRC, SRMS, STAT1, STAT2, STAT3, STAT4, STAT5A, STAT5B, STAT6, SUPT6H, SYK, TEC, TENC1, TNS, TNS1, TNS3, TNS4, TXK, VAV1, VAV2, VAV3, YES1, ZAP70 либо его производное, вариант или фрагмент. В некоторых воплощениях адаптерный полипептид включает белок, содержащий домен PTB, как-то APBA1, APBA2, APBA3, EPS8, EPS8L1, EPS8L2, EPS8L3, TENC1, TNS, TNS1, TNS3, TNS4, DOK1, DOK2, DOK3, DOK4, DOK5, DOK6, DOK7, FRS2, FRS3, IRS1, IRS2, IRS3, IRS4, SHC1, SHC2, SHC3, SHC4, TLN1, TLN2, X11a либо его производное, вариант или фрагмент.

В некоторых воплощениях химерный адаптерный полипептид содержит белок, который привлекается к рецептору TNF, либо его производное, вариант или фрагмент. Такие белки иногда упоминаются как связанные с рецепторами TNF факторы или TRAFs и включают TRAF1, TRAF2, TRAF3, TRAF4, TRAF5, TRAF6 и TRAF7. В некоторых воплощениях химерный адаптерный полипептид содержит взаимодействующую с рецептором сериновую/треониновую протеинкиназу 1 (RIP1 или RIPK1) или взаимодействующую с рецептором сериновую/треониновую протеинкиназу 3 (RIP3 или RIPK3) либо её производное, вариант или фрагмент. В некоторых воплощениях химерный адаптерный полипептид содержит адаптерный белок, который привлекается к TNFR, типа Fas-ассоциированного белка с доменом смерти (FADD) и ассоциированного с рецептором фактора некроза опухолей 1-го типа белка с доменом смерти (TRADD), который связывает TRAF2, либо его производное, вариант или фрагмент.

В некоторых воплощениях химерный адаптерный полипептид содержит молекулу (например, белок) либо её производное, вариант или фрагмент, который привлекается к фосфорилированному ITAM, к примеру, ITAM химерного полипептидного рецептора, включающего иммунорецептор типа TCR. Фосфорилирование определенных остатков тирозина у активированного рецептора может создавать сайты связывания для таких молекул, как белки, содержащие домен гомологии-2 Src (SH2) и связывающий фосфотирозин домен (PTB). В некоторых воплощениях химерный адаптерный полипептид содержит ABL1, ABL2, BCAR3, BLK, BLNK, BMX, BTK, CHN2, CISH, CRK, CRKL, CSK, DAPP1, EAT-2, FER, FES, FGR, FRK, FYN, GADS, GRAP, GRAP2, GRB10, GRB14, GRB2, GRB7, HCK, HSH2D, INPP5D, INPPL1, ITK, JAK2, LCK, LCP2, LYN, MATK, NCK1, NCK2, PIK3R1, PIK3R2, PIK3R3, PLCG1, PLCG2, PTK6, PTPN11, PTPN6, RASA1, SAP, SH2B1, SH2B2, SH2B3, SH2D1A, SH2D1B, SH2D2A, SH2D3A, SH2D3C, SH2D4A, SH2D4B, SH2D5, SH2D6, SH3BP2, SHB, SHC1, SHC2, SHC3, SHC4, SHD, SHE, SHP1, SHP2, SLA, SLA2, SOCS1, SOCS2, SOCS3, SOCS4, SOCS5, SOCS6, SOCS7, SRC, SRMS, STAT1, STAT2, STAT3, STAT4, STAT5A, STAT5B, STAT6, SUPT6H, SYK, TEC, TENC1, TNS, TNS1, TNS3, TNS4, TXK, VAV1, VAV2, VAV3, YES1, ZAP70 либо его производное, вариант или фрагмент. В некоторых воплощениях химерный адаптерный полипептид содержит APBA1, APBA2, APBA3, EPS8, EPS8L1, EPS8L2, EPS8L3, TENC1, TNS, TNS1, TNS3, TNS4, DOK1, DOK2, DOK3, DOK4, DOK5, DOK6, DOK7, FRS2, FRS3, IRS1, IRS2, IRS3, IRS4, SHC1, SHC2, SHC3, SHC4, TLN1, TLN2, X11a либо его производное, вариант или фрагмент.

В некоторых конфигурациях химерный адаптерный полипептид рассматриваемой системы может включать полипептид, модулирующий ген (GMP). Как уже описано здесь, GMP может содержать исполнительный элемент, соединенный с сайтом распознавания расщепления. Исполнительный элемент может содержать нуклеазу (например, ДНК-нуклеазу и/или РНК-нуклеазу), модифицированную нуклеазу (например, ДНК-нуклеазу и/или РНК-нуклеазу), которая дефектна по нуклеазе или обладает меньшей нуклеазной активностью по сравнению с нуклеазой дикого типа, её вариант, производное или фрагмент, как уже описано здесь. Исполнительный элемент может регулировать экспрессию и/или активность гена или

редактировать последовательность нуклеиновой кислоты (например, гена и/или продукта гена). В некоторых воплощениях исполнительный элемент содержит ДНК-нуклеазу типа генно-инженерной (например, программируемой или наводимой) ДНК-нуклеазы, чтобы запускать геномное редактирование целевой последовательности ДНК. В некоторых воплощениях исполнительный элемент содержит РНК-нуклеазу типа генно-инженерной (например, программируемой или наводимой) РНК-нуклеазы, чтобы запускать редактирование целевой последовательности РНК. В некоторых воплощениях исполнительный элемент обладает пониженной или минимальной нуклеазной активностью. Исполнительный элемент с пониженной или минимальной нуклеазной активностью может регулировать экспрессию и/или активность гена путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов, эффективно подавляющих или усиливающих экспрессию целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит безнуклеазный ДНК-связывающий белок, происходящий из ДНК-нуклеазы, который может вызывать активацию или репрессию транскрипции целевой последовательности ДНК. В некоторых воплощениях исполнительный элемент содержит безнуклеазный связывающий белок, связывающий РНК, происходящий из РНК-нуклеазы, который может вызывать активацию или репрессию транскрипции целевой последовательности РНК. В некоторых воплощениях исполнительный элемент управляется направляющей нуклеиновой кислотой. В некоторых воплощениях исполнительный элемент управляется направляющей ДНК. В некоторых воплощениях исполнительный элемент управляется направляющей РНК. Исполнительный элемент может регулировать экспрессию или активность гена и/или редактировать последовательность нуклеиновой кислоты, будь то экзогенной или эндогенной. Например, исполнительный элемент может содержать белок Cas, лишенный активности расщепления.

В исполнительных элементах может использоваться любая подходящая нуклеаза. К подходящим нуклеазам относятся, без ограничения, CRISPR-ассоциированные (Cas) белки или Cas-нуклеазы, в том числе CRISPR-ассоциированные (Cas) полипептиды I типа, CRISPR-ассоциированные (Cas) полипептиды II типа, CRISPR-ассоциированные (Cas) полипептиды III типа, CRISPR-ассоциированные (Cas) полипептиды IV типа, CRISPR-ассоциированные (Cas) полипептиды V типа и CRISPR-ассоциированные (Cas) полипептиды VI типа; нуклеазы с "цинковым пальцем" (ZFN); эффекторные нуклеазы типа активаторов транскрипции (TALEN); мегануклеазы; РНК-связывающие белки (RBP); CRISPR-ассоциированные РНК-связывающие белки; рекомбиназы; флипазы; транспозазы; белки Argonaute (Ago) (например, прокариотические белки Argonaute (pAgo), археальные Argonaute (aAgo) и эукариотические Argonaute (eAgo)); их производные, варианты или фрагменты. В некоторых воплощениях исполнительный элемент содержит белок Cas, который образует комплекс с направляющей нуклеиновой кислотой типа направляющей РНК. В некоторых воплощениях исполнительный элемент содержит белок, связывающий РНК (RBP), необязательно в комплексе с направляющей нуклеиновой кислотой типа направляющей РНК, который способен образовывать комплекс с белком Cas.

В некоторых воплощениях исполнительный элемент содержит белок, связывающий РНК (RBP), необязательно в комплексе с направляющей нуклеиновой кислотой типа направляющей РНК, который способен образовывать комплекс с белком Cas. На фиг. 6B представлен типичный химерный адаптерный полипептид, у которого исполнительный элемент содержит белок, связывающий РНК 600a, необязательно в комплексе с направляющей нуклеиновой кислотой. После высвобождения из РНК-связывающего белка (RBP), к примеру, при диссоциации направляющей нуклеиновой кислоты из RBP или расщеплении сайта распознавания расщепления, направляющая нуклеиновая кислота может образовать комплекс с белком Cas 600b, который способен регулировать экспрессию и/или активность целевого полинуклеотида (например, экспрессию гена) или редактировать последовательность нуклеиновой кислоты. В некоторых воплощениях исполнительный элемент содержит безнуклеазный ДНК-связывающий белок, происходящий из ДНК-нуклеазы, который может вызывать активацию или репрессию транскрипции целевой последовательности ДНК. В некоторых воплощениях исполнительный элемент содержит безнуклеазный белок, связывающий РНК, происходящий из РНК-нуклеазы, который может вызывать активацию или репрессию транскрипции целевой последовательности РНК. Например, исполнительный элемент может содержать белок Cas, который лишен активности расщепления.

В некоторых воплощениях сайт распознавания расщепления фланкирован участком рецепторного связывания и исполнительным элементом. Исполнительный элемент может высвободиться из GMP при расщеплении сайта распознавания расщепления расщепляющим элементом. Расщепляющий элемент может распознавать и/или расщеплять сайт распознавания расщепления, к примеру, когда он находится в близости к сайту распознавания расщепления. Расщепляющий элемент может содержать полипептидную последовательность. В некоторых конфигурациях расщепляющий элемент входит в состав химерного рецепторного полипептида. Расщепляющий элемент может составлять N-конец, C-конец или внутреннюю часть рецептора. В некоторых воплощениях расщепляющий элемент образует комплекс с химерным рецепторным полипептидом. Расщепляющий элемент может входить в комплекс с N-концом, C-концом или внутренней частью химерного рецепторного полипептида. В типичной конфигурации, представленной на фиг. 7, сайт распознавания расщепления 702b фланкирован участком рецепторного связывания 701 и исполнительным элементом 702a, а расщепляющий элемент 706 входит в состав химерного

рецепторного полипептида 705.

На фиг. 8A-D схематически представлено высвобождение исполнительного элемента из GMP. На фиг. 8A представлено связывание антигена с химерным полипептидом трансмембранного рецептора. Химерный полипептид трансмембранного рецептора включает внеклеточную область, содержащую взаимодействующий с антигеном домен 805, и внутриклеточную область, содержащую расщепляющий элемент 806. Расщепляющий элемент может входить в комплекс с рецептором или соединяться с рецептором, к примеру, пептидной связью и/или пептидным линкером. GMP входит в состав химерного адаптерного полипептида. GMP, соединенный с участком рецепторного связывания 801, включает в себя исполнительный элемент 802a, соединенный с сайтом распознавания расщепления 802b. В ответ на связывание антигена рецептор подвергается модификации путем фосфорилирования 803 во внутриклеточной области рецептора (фиг. 8B). После модификации (например, фосфорилирования) рецептора к рецептору привлекается химерный адаптерный полипептид, как показано на фиг. 8C. Рецептор содержит расщепляющий элемент 806. Когда он находится в непосредственной близости от сайта распознавания расщепления, расщепляющий элемент может расщеплять сайт распознавания, высвобождая исполнительный элемент из GMP, как показано на фиг. 8D. После высвобождения исполнительный элемент может поступать в ядро для регуляции экспрессии и/или активности целевого гена или редактирования последовательности нуклеиновой кислоты. На фиг. 8E-H представлена аналогичная система, в которой модификация рецептора включает конформационное изменение. В некоторых воплощениях химерный адаптерный белок привязан к мембране (например, в виде мембраносвязанного белка).

В другой конфигурации расщепляющий элемент находится в комплексе со вторым адаптерным полипептидом, который связывает химерный рецепторный полипептид, когда рецепторный полипептид подвергается модификации. Показательный пример представлен на фиг. 9. Сайт распознавания расщепления 902b фланкирован участком рецепторного связывания 901 и исполнительным элементом 902a, а расщепляющий элемент 906 входит в состав второго адаптерного полипептида 907.

На фиг. 10A-D схематически представлено высвобождение исполнительного элемента из GMP. На фиг. 10A представлено связывание антигена с химерным полипептидом трансмембранного рецептора. Химерный полипептид трансмембранного рецептора включает внеклеточную область, содержащую взаимодействующий с антигеном домен, и внутриклеточную область. GMP, содержащий исполнительный элемент, соединенный с сайтом распознавания расщепления, входит в состав химерного адаптерного полипептида. Сайт распознавания расщепления 1002b фланкирован участком рецепторного связывания 1001 и исполнительным элементом 1002a. В ответ на связывание антигена рецептор подвергается модификации путем фосфорилирования 1003 во внутриклеточной области (фиг. 10B). После модификации (например, фосфорилирования) рецептора к нему привлекается химерный адаптерный полипептид, как показано на фиг. 10B. Второй адаптерный полипептид 1007, содержащий расщепляющий элемент 1006, тоже привлекается к модифицированному рецептору (фиг. 10C). Расщепляющий элемент может входить в комплекс со вторым адаптерным полипептидом или соединяться с ним, к примеру, пептидной связью и/или пептидным линкером. Когда он находится в близости от сайта распознавания расщепления, расщепляющий элемент может расщеплять сайт распознавания, высвобождая исполнительный элемент из GMP, как показано на фиг. 10D. После высвобождения исполнительный элемент может поступать в ядро для регуляции экспрессии и/или активности целевого гена или редактирования последовательности нуклеиновой кислоты. На фиг. 10E-H представлена аналогичная система, в которой модификация рецептора включает конформационное изменение. В некоторых воплощениях химерный адаптерный полипептид привязан к мембране (например, в виде мембраносвязанного белка). В некоторых воплощениях второй адаптерный полипептид привязан к мембране (например, в виде мембраносвязанного белка).

На фиг. 16A-D схематически представлено высвобождение исполнительного элемента в системе, включающей первый мембраносвязанный адаптер и второй цитоплазматический адаптер. На фиг. 16A представлена ассоциация первого мембраносвязанного адаптера, содержащего домен мембранной привязки 1601a (например, СААХ), сайт распознавания протеазы 1601b (например, TEV) и исполнительный элемент 1601c, с химерным трансмембранным рецептором 1602. Химерный трансмембранный рецептор может функционировать в качестве каркаса и включает в себя по меньшей мере два адаптер-связывающих сайта (например, EGFR или рецепторной тирозинкиназы (RTK)). Один адаптер-связывающий сайт может связываться с мембраносвязанным адаптером, как показано на фиг. 16B. В некоторых случаях связывание мембраносвязанного адаптера зависит от связывания антигена с рецептором. В некоторых системах мембраносвязанный адаптер располагается в близости от рецептора, и ассоциация может не зависеть от связывания антигена с рецептором. Как показано на фиг. 16B и 16C, при взаимодействии антигена с рецептором может условно рекрутироваться второй адаптерный белок, содержащий цитоплазматический рецепторсвязывающий элемент 1603a и протеазу 1603b, к другому адаптер-связывающему сайту рецептора. Содержащий протеазу второй адаптерный белок при рекрутировании к трансмембранному рецептору может расщеплять сайт распознавания протеазы 1601b у мембраносвязанной молекулы, тем самым высвобождая исполнительный элемент 1601c, как показано на фиг. 16D.

В некоторых воплощениях расщепляющий элемент расщепляет по сайту распознавания только то-

гда, когда он находится вблизи сайта распознавания расщепления. В некоторых воплощениях сайт распознавания расщепления содержит полипептидную последовательность (например, домен расщепления пептида), которая является распознавательной последовательностью для протеазы. Расщепляющий элемент обладает активностью протеазы, которая распознает полипептидную последовательность. Расщепляющий элемент, обладающий протеазной активностью, может представлять собой протеазу, включая, без ограничения, протеазы, уже описанные здесь, либо их производные, варианты или фрагменты. В некоторых воплощениях сайт распознавания расщепления содержит несколько последовательностей распознавания расщепления, причем каждая последовательность распознавания расщепления может распознаваться одним и тем же или разными расщепляющими элементами, обладающими протеазной активностью (например, протеазами).

В некоторых воплощениях сайт распознавания расщепления содержит первую часть последовательности интеина, которая реагирует со второй частью последовательности интеина с высвобождением исполнительного элемента. Для облегчения высвобождения исполнительного элемента из химерного адаптерного полипептида может использоваться гетерологическая система с разделенным интеином. Исполнительный элемент может ковалентно соединяться с первой частью последовательности интеина. Исполнительный элемент может соединяться через свой N-конец или C-конец с первой частью последовательности интеина. Расщепляющий элемент может содержать вторую часть последовательности интеина. Первой частью или второй частью последовательности интеина может быть N-концевой интеин, C-концевой интеин или любая другая подходящая часть интеина, которая может способствовать высвобождению исполнительного элемента. Последовательности интеина могут происходить из любого подходящего источника. Первая и вторая часть могут быть из одного источника (например, организма, белка) или из разных источников. В одном показательном примере исполнительный элемент может быть ковалентно связан (например, на своем N-конце или C-конце) пептидной связью с первой частью последовательности интеина, содержащей N-конец интеина. Слитый с N-концом интеина исполнительный элемент может входить в контакт со второй частью последовательности интеина, содержащей C-конец интеина. При таком контакте между первой и второй частью последовательности интеина может происходить сайт-специфичное расщепление (например, на участке между исполнительным элементом и N-концом интеина), при этом высвобождается исполнительный элемент. В другом показательном примере исполнительный элемент может быть ковалентно связан (например, на своем N-конце или C-конце) пептидной связью с первой частью последовательности интеина, содержащей C-конец интеина. Слитый с C-концом интеина исполнительный элемент может входить в контакт со второй частью последовательности интеина, содержащей N-конец интеина. При таком контакте между первой и второй частью интеина может происходить сайт-специфичное расщепление (например, на подходящем участке между исполнительным элементом и C-концом интеина), при этом высвобождается исполнительный элемент.

В некоторых воплощениях сайт распознавания расщепления содержит дисульфидную связь. Дисульфидная связь может соединять исполнительный элемент с участком рецепторного связывания у химерного адаптерного полипептида. Дисульфидная связь может образовываться между одним или несколькими цистеинами исполнительного элемента и участком рецепторного связывания. Цистеины могут быть встроены в исполнительный элемент или участок связывания рецептора. Цистеины могут входить в состав нативной или последовательности дикого типа исполнительного элемента или участка рецепторного связывания. Цистеины могут присутствовать в линкерном пептиде, присоединенном к исполнительному элементу или участку рецепторного связывания. Расщеплению дисульфидной связи может способствовать, к примеру, изменение окислительно-восстановительных условий у дисульфидной связи. Изменение окислительно-восстановительных условий может приводить к восстановлению дисульфидной связи до тиолов и высвобождению исполнительного элемента. Расщеплению дисульфидной связи может способствовать расщепляющий элемент, содержащий окислительно-восстановительное средство, которое может вызывать восстановление дисульфидной связи. Окислительно-восстановительным средством может быть фермент либо его производное, вариант или фрагмент. Фермент может быть оксидоредуктазой. Примеры оксидоредуктаз включают протеин-дисульфидредуктазы, тиоредоксины, глутаредоксины, тиол-дисульфид-оксидоредуктазы (например, DsbA, BdbA-D, MdbA и SdbA) и глутатион-дисульфидредуктазы. Окислительно-восстановительное средство может быть из любого подходящего источника, включая прокариотов и эукариотов. Для оптимальной активности фермента могут потребоваться кофакторы (например, никотинамидные кофакторы, флавины либо их производные и аналоги).

В некоторых воплощениях химерный адаптерный полипептид содержит по меньшей мере одну нацеливающую последовательность, которая направляет транспорт адаптера в определенный район клетки. К примеру, нацеливающая последовательность может направлять рецептор в ядро клетки, используя сигнал ядерной локализации (NLS), за пределы ядра (например, в цитоплазму), используя сигнал ядерного экспорта (NES), в митохондрию, эндоплазматический ретикулум (ER), аппарат Гольджи, хлоропласты, апопласты, пероксисомы, плазматические мембраны или мембраны различных органелл клетки. В некоторых воплощениях нацеливающая последовательность содержит сигнал ядерного экспорта (NES) и направляет химерный адаптерный полипептид за пределы ядра. В некоторых воплощениях нацеливающая

последовательность содержит сигнал ядерной локализации (NLS) и направляет адаптер в ядро клетки. Нацеливающая последовательность может направлять полипептид в ядро клетки, используя различные сигналы ядерной локализации (NLS). В некоторых воплощениях нацеливающая последовательность включает нацеливающий на мембрану пептид и направляет адаптер на плазматическую мембрану или мембрану клеточной органеллы. Нацеливающая последовательность может направлять полипептид на мембрану, используя сигнальную последовательность мембранного якорения, как описано выше. Доступны различные последовательности мембранного якорения.

Нацеливающая последовательность может быть связана с любым подходящим участком химерного адаптерного полипептида, к примеру, на N-конце, C-конце полипептида или во внутренней области адаптера. В некоторых воплощениях с адаптером связаны по меньшей мере две наводящие последовательности. Например, как показано на фиг. 11, первая нацеливающая последовательность 1101a может быть связана с участком рецепторного связывания адаптера, а вторая нацеливающая последовательность 1101b может быть связана с GMP адаптера, к примеру, с исполнительным элементом. Когда адаптер связан с несколькими наводящими последовательностями, к примеру, с наводящими последовательностями, направляющими в различные дислокации клетки, то окончательная локализация адаптера может определяться относительной силой наводящих последовательностей. К примеру, адаптер, содержащий и нацеливающую последовательность, включающую NES, и нацеливающую последовательность, включающую NLS, может локализоваться в цитоплазме, если NES будет сильнее, чем NLS. С другой стороны, если NLS будет сильнее, чем NES, то адаптер может локализоваться в ядре, даже если у адаптера будет присутствовать как сигнал ядерной локализации, так и сигнал ядерного экспорта. Нацеливающая последовательность может содержать несколько копий, к примеру, каждого из NLS и NES, для точной настройки клеточной локализации.

В некоторых случаях нацеливающая последовательность связана с исполнительным элементом. После высвобождения исполнительного элемента из GMP (и адаптера) при расщеплении сайта распознавания расщепления нацеливающая последовательность может направлять исполнительный элемент в такое место клетки, которое отличается от адаптера. Например, химерный адаптерный полипептид может содержать первую нацеливающую последовательность, которая направляет адаптер в цитоплазму клетки, а исполнительный элемент может отдельно содержать вторую нацеливающую последовательность, которая направляет локализацию в ядро клетки. Первоначально исполнительный элемент (входящий в состав адаптера) может локализоваться в цитоплазме клетки из-за первой нацеливающей последовательности. После высвобождения исполнительного элемента из GMP при расщеплении сайта распознавания расщепления исполнительный элемент может локализоваться в ядре клетки под действием второй нацеливающей последовательности. В некоторых воплощениях исполнительный элемент транслоцируется в ядро клетки после расщепления последовательности распознавания расщепления.

В некоторых воплощениях нацеливающая последовательность включает нацеливающий на мембрану пептид и направляет полипептид в плазматическую мембрану или мембрану клеточной органеллы. Нацеливающая на мембрану последовательность может обеспечивать транспорт химерного полипептида трансмембранного рецептора на поверхность клеточную мембрану или на другую клеточную мембрану. Может использоваться любая подходящая нацеливающая на мембрану последовательность, описанная здесь.

В некоторых воплощениях химерный адаптерный полипептид связан с доменом фолдинга полипептидов, который может способствовать сворачиванию белка. В некоторых воплощениях исполнительный элемент может быть связан с доменом проникновения в клетки. Например, домен проникновения в клетки может происходить из белка TAT HIV-1, мотива проникновения в клетки TLM из вируса гепатита В человека, MPG, Pep-1, VP22, проникающего в клетки пептида вируса Herpes simplex или последовательности полиаргининового пептида. Проникающий в клетки домен может располагаться на N-конце, C-конце или же где-то внутри исполнительного элемента.

Исполнительный элемент рассматриваемой системы после высвобождения из химерного адаптерного полипептида или химерного рецепторного полипептида может связываться с целевым полинуклеотидом и регулировать экспрессию и/или активность целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит активатор транскрипции для эффективного повышения экспрессии целевого полинуклеотида. Исполнительный элемент может содержать репрессор транскрипции для эффективного снижения экспрессии целевого полинуклеотида. В некоторых воплощениях целевой полинуклеотид содержит геномную ДНК. В некоторых воплощениях целевой полинуклеотид содержит участок плазмиды, к примеру, плазмиды, несущей экзогенный ген. В некоторых воплощениях целевой полинуклеотид содержит РНК, например, мРНК. В некоторых воплощениях целевой полинуклеотид содержит эндогенный ген или генный продукт. Исполнительный элемент может включать в себя одну или несколько копий сигнала ядерной локализации, что позволяет ему транслоцироваться в ядро клетки при отщеплении от GMP.

В некоторых аспектах химерный рецепторный полипептид представляет собой химерный внутриклеточный рецептор. Типичный химерный внутриклеточный рецептор включает (а) взаимодействующий

с антигеном домен, который специфически связывает антиген, и (b) исполнительный элемент, связанный со взаимодействующим с антигеном доменом. В некоторых воплощениях (i) химерный внутриклеточный рецептор подвергается модификации в ответ на связывание антигена, (ii) химерный полипептид внутриклеточного рецептора транслоцируется в ядро клетки в ответ на модификацию, а (iii) исполнительный элемент образует комплекс с целевым полинуклеотидом в ядре.

В некоторых воплощениях химерный внутриклеточный рецептор представлен ядерным рецептором. Например, химерный внутриклеточный рецепторный полипептид может содержать ядерный рецептор либо его производное, вариант или фрагмент, выбранный из α -рецептора тиреоидного гормона ($TR\alpha$), β -рецептора тиреоидного гормона ($TR\beta$), α -рецептора ретиноевой кислоты ($RAR\alpha$), β -рецептора ретиноевой кислоты ($RAR\beta$), γ -рецептора ретиноевой кислоты ($RAR\gamma$), активируемого пролифераторами пероксисом рецептора- α ($PPAR\alpha$), активируемого пролифераторами пероксисом рецептора- β/δ ($PPAR\beta/\delta$), активируемого пролифераторами пероксисом рецептора- γ ($PPAR\gamma$), $Rev-ErbA\alpha$, $Rev-ErbA\beta$, родственного RAR орфанного рецептора α ($ROR\alpha$), родственного RAR орфанного рецептора β ($ROR\beta$), родственного RAR орфанного рецептора γ ($ROR\gamma$), печеночного X-рецептора α , печеночного X-рецептора β , фарнезоидного X-рецептора, фарнезоидного X-рецептора β , рецептора витамина D, прегнанового X-рецептора, конститутивного рецептора андростанов, ядерного фактора 4- α гепатоцитов ($HNF4\alpha$), ядерного фактора 4- γ гепатоцитов ($HNF4\gamma$), ретиноидного X-рецептора α ($RXR\alpha$), ретиноидного X-рецептора β ($RXR\beta$), ретиноидного X-рецептора γ ($RXR\gamma$), тестикулярного рецептора 2 ($TR2$), тестикулярного рецептора 4 ($TR4$), гомолога гена дрозофилы *tailless* (TLX), специфичного для фоторецепторных клеток ядерного рецептора (PNR), фактора транскрипции I-вышележащего промотора куриного овальбумина (COUP-TFI), фактора транскрипции II-вышележащего промотора куриного овальбумина (COUP-TFII), родственного V-erbA рецептора ($EAR-2$), α -рецептора эстрогенов ($ER\alpha$), β -рецептора эстрогенов ($ER\beta$), родственного эстрогеновым рецептора α ($ERR\alpha$), родственного эстрогеновым рецептора β ($ERR\beta$), родственного эстрогеновым рецептора γ ($ERR\gamma$), глюкокортикоидного рецептора (GR), минералокортикоидного рецептора (MR), рецептора прогестерона (PR), рецептора андрогенов (AR), фактора роста нервов IB ($NGFIB$), родственного ядерным рецептора-1 ($NURR1$), орфанного рецептора 1 из нейронов ($NOR1$), стероидогенного фактора 1 ($SF1$), гомолога-1 печеночного рецептора ($LRH-1$) и ядерного фактора зародышевых клеток ($GCNF$).

Химерный внутриклеточный рецептор, содержащий ядерный рецептор либо его производное, вариант или фрагмент, может связывать антиген, содержащий любой подходящий лиганд ядерного рецептора, либо его производное, вариант или фрагмент. Неограничительные примеры лигандов ядерных рецепторов включают тиреоидный гормон, витамин A и родственные соединения, жирные кислоты, простагландины, гем, холестерин, ATRA, оксистеролы, витамин D, ксенобиотики, андростаны, ретиноиды, эстрогены, кортизол, альдостерон, прогестерон, тестостерон и фосфатидилинозитолы. В некоторых воплощениях антиген представлен гормоном.

На фиг. 12A-C схематически представлена система, включающая типичный внутриклеточный рецептор, содержащий ядерный рецептор. Система включает в себя рецептор 1200, содержащий исполнительный элемент 1201. В отсутствие связывания лиганда с ядерным рецептором рецептор может секвестрироваться в определенном компартменте клетки, например, в цитоплазме, путем взаимодействия со связывающим белком 1202, как показано на фиг. 12A. После связывания лиганда 1203 с внутриклеточным рецептором, как показано на фиг. 12B, рецептор может диссоциировать от связывающего белка 1202 и транслоцироваться в ядро. Исполнительный элемент 1201, который образует комплекс и/или связан с рецептором, проникает в ядро с рецептором, где он работает, регулируя экспрессию и/или активность целевого полинуклеотида (например, экспрессию гена) либо редактируя последовательность нуклеиновой кислоты.

Исполнительный элемент может содержать нуклеазу (например, ДНК-нуклеазу и/или РНК-нуклеазу), модифицированную нуклеазу (например, ДНК-нуклеазу и/или РНК-нуклеазу), которая дефектна по нуклеазе или обладает меньшей нуклеазной активностью по сравнению с нуклеазой дикого типа, её вариант, производное или фрагмент, как уже описано здесь. Исполнительный элемент может регулировать экспрессию и/или активность гена или редактировать последовательность нуклеиновой кислоты (например, гена и/или продукта гена). В некоторых воплощениях исполнительный элемент содержит ДНК-нуклеазу типа генно-инженерной (например, программируемой или наводимой) ДНК-нуклеазы, чтобы запускать геномное редактирование целевой последовательности ДНК. В некоторых воплощениях исполнительный элемент содержит РНК-нуклеазу типа генно-инженерной (например, программируемой или наводимой) РНК-нуклеазы, чтобы запускать редактирование целевой последовательности РНК. В некоторых воплощениях исполнительный элемент обладает пониженной или минимальной нуклеазной активностью. Исполнительный элемент с пониженной или минимальной нуклеазной активностью может регулировать экспрессию и/или активность гена путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов, эффективно подавляющих или усиливающих экспрессию целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит безнуклеазный ДНК-связывающий белок, происходящий из ДНК-нуклеазы, который может вызы-

вать активацию или репрессию транскрипции целевой последовательности ДНК. В некоторых воплощениях исполнительный элемент содержит безнуклеазный белок, связывающий РНК, происходящий из РНК-нуклеазы, который может вызывать активацию или репрессию транскрипции целевой последовательности РНК. В некоторых воплощениях исполнительный элемент управляется направляющей нуклеиновой кислотой. В некоторых воплощениях исполнительный элемент управляется направляющей ДНК. В некоторых воплощениях исполнительный элемент управляется направляющей РНК. Исполнительный элемент может регулировать экспрессию или активность гена и/или редактировать последовательность нуклеиновой кислоты, будь то экзогенной или эндогенной.

В исполнительных элементах может использоваться любая подходящая нуклеаза. К подходящим нуклеазам относятся, без ограничения, CRISPR-ассоциированные (Cas) белки или Cas-нуклеазы, в том числе CRISPR-ассоциированные (Cas) полипептиды I типа, CRISPR-ассоциированные (Cas) полипептиды II типа, CRISPR-ассоциированные (Cas) полипептиды III типа, CRISPR-ассоциированные (Cas) полипептиды IV типа, CRISPR-ассоциированные (Cas) полипептиды V типа и CRISPR-ассоциированные (Cas) полипептиды VI типа; нуклеазы с "цинковым пальцем" (ZFN); эффекторные нуклеазы типа активаторов транскрипции (TALEN); мегануклеазы; РНК-связывающие белки (RBP); CRISPR-ассоциированные РНК-связывающие белки; рекомбиназы; флиппазы; транспозазы; белки Argonaute (например, прокариотические белки Argonaute (pAgo), археальные Argonaute (aAgo) и эукариотические Argonaute (eAgo)); их производные, варианты или фрагменты. В некоторых воплощениях исполнительный элемент содержит белок Cas, который образует комплекс с направляющей нуклеиновой кислотой типа направляющей РНК.

Исполнительный элемент внутриклеточного рецептора после транслокации в ядро клетки (например, с внутриклеточным рецептором) может связываться с целевым полинуклеотидом для регуляции экспрессии и/или активности целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит активатор транскрипции для эффективного повышения экспрессии целевого полинуклеотида. Исполнительный элемент может содержать репрессор транскрипции для эффективного уменьшения экспрессии целевого полинуклеотида.

В некоторых воплощениях целевой полинуклеотид содержит геномную ДНК. В некоторых воплощениях целевой полинуклеотид содержит участок плазмиды, к примеру, плазмиды, несущей экзогенный ген. В некоторых воплощениях целевой полинуклеотид содержит РНК, к примеру, мРНК. В некоторых воплощениях целевой полинуклеотид содержит эндогенный ген или генный продукт.

В некоторых случаях с внутриклеточным рецептором связана нацеливающая последовательность. К примеру, нацеливающая последовательность может направлять рецептор в ядро клетки, используя сигнал ядерной локализации (NLS), за пределы ядра клетки (например, в цитоплазму), используя сигнал ядерного экспорта (NES), в митохондрии, эндоплазматический ретикулум (ER), Гольджи, хлоропласта, апопласты или пероксисомы. В некоторых воплощениях нацеливающая последовательность содержит сигнал ядерного экспорта (NES) и направляет рецептор за пределы клеточного ядра. В некоторых воплощениях нацеливающая последовательность содержит сигнал ядерной локализации (NLS) и направляет рецептор в ядро клетки. Нацеливающая последовательность может направлять рецептор в ядро клетки, используя различные сигналы ядерной локализации (NLS). В некоторых воплощениях химерный внутриклеточный рецептор связан с доменом фолдинга полипептидов, который может способствовать сворачиванию белка.

Рассматриваемая система может быть введена в различные клетки. Клетки могут быть *in vitro*. Клетки могут быть *in vivo*. Клетки могут быть *ex vivo*. Клетки могут быть выделенными клетками. Клетки могут быть клетками внутри организма. Клетки могут представлять собой организм. Клетки могут быть клетками в клеточной культуре. Клетки могут быть клетками из коллекции клеток. Клетки могут быть клетками млекопитающих или происходить из клеток млекопитающих. Клетки могут быть клетками грызунов или происходить из клеток грызунов. Клетки могут быть клетками человека или происходить из клеток человека. Клетки могут быть прокариотическими клетками или происходить из прокариотических клеток. Клетки могут быть бактериальными клетками или происходить из бактериальных клеток. Клетки могут быть архейными клетками или происходить из архейных клеток. Клетки могут быть эукариотическими клетками или происходить из эукариотических клеток. Клетки могут быть плюрипотентными стволовыми клетками. Клетки могут быть клетками растений или происходить из клеток растений. Клетки могут быть клетками животных или происходить из клеток животных. Клетки могут быть клетками беспозвоночных или происходить из клеток беспозвоночных. Клетки могут быть клетками позвоночных или происходить из клеток позвоночных. Клетки могут быть клетками микробов или происходить из клеток микробов. Клетки могут быть клетками грибов или происходить из клеток грибов. Клетки могут происходить из определенного органа или ткани.

Клетка может быть стволовой клеткой или клеткой-предшественником. Клетка может включать стволовую клетку (например, взрослые стволовую клетку, эмбриональные стволовую клетку, клетку iPS) и клетку-предшественник (например, сердечные клетки-предшественники, нервные клетки-предшественники и т.п.). Клетка может включать стволовую клетку и клетку-предшественник млекопитающих, вклю-

чая стволовую клетку грызунов, клетку-предшественник грызунов, стволовую клетку человека, клетку-предшественник человека и т.д. Клональная клетка может включать потомство клетки. Клетка может содержать целевую нуклеиновую кислоту. Клетка может находиться в живом организме. Клетка может быть генетически модифицированной клеткой. Клетка может быть клеткой-хозяином.

Клетка может быть тотипотентной и стволовой клеткой, однако, хотя в некоторых воплощениях настоящего изобретения может применяться термин "клетка", но он может не относиться к тотипотентным стволовым клеткам. Клетка может быть клеткой растений, но, хотя в некоторых воплощениях настоящего изобретения может применяться термин "клетка", он может не относиться к растительным клеткам. Клетка может быть плюрипотентной клеткой. Например, клетка может быть плюрипотентной гемопоэтической клеткой, которая может дифференцироваться в другие клетки в линии гемопоэтических клеток, но не может дифференцироваться в какие-либо другие негемопоэтические клетки. Клетка может быть способна развиваться в целый организм. Клетка может и не быть способной развиваться в целый организм. Клетка может представлять собой целый организм.

Клетка может быть первичной клеткой. Например, культуры первичных клеток можно пассировать 0 раз, 1 раз, 2 раза, 4 раза, 5 раз, 10 раз, 15 и более раз. Клетка может быть одноклеточным организмом. Клетку можно выращивать в культуре.

Клетка может быть пораженной болезнью клеткой. Больные клетки могут иметь измененный метаболизм, экспрессию генов и/или морфологические признаки. Больная клетка может представлять собой раковую клетку, диабетическую клетку и апоптотическую клетку. Больная клетка может быть представлена клеткой от больного субъекта. Типичные заболевания могут включать гематологические заболевания, раковые заболевания, метаболические заболевания, глазные заболевания, заболевания органов, заболевания опорно-двигательного аппарата, сердечные заболевания и др.

Если клетки представлены первичными клетками, то их можно получать от людей любыми способами. Например, лейкоциты можно получать методом афереза, лейкоцитафереза, разделения в градиенте плотности и т.п. Клетки из таких тканей, как кожа, мышцы, костный мозг, селезенка, печень, поджелудочная железа, легкие, кишечник, желудок и т.п. можно получать при помощи биопсии. Для диспергирования или суспендирования полученных клеток можно использовать соответствующий раствор. Таким раствором может быть сбалансированный солевой раствор (например, обычный физраствор, физраствор с фосфатным буфером (PBS), сбалансированный солевой раствор Хэнка и др.), для удобства дополненный фетальной телячьей сывороткой или другими природными факторами в сочетании с приемлемым буфером при низкой концентрации. Буферы могут включать HEPES, фосфатные буферы, лактатные буферы и т.п. Клетки можно использовать немедленно или же их можно хранить (например, путем замораживания). Замороженные клетки можно оттаивать и их можно использовать повторно. Клетки можно замораживать в DMSO, сыворотке, буферной среде (например, в 10% DMSO, 50% сыворотке, 40% буферной среде) и/или в каком-либо другом распространенном растворе, который применяется для хранения клеток при температуре заморозки.

Неограничительные примеры клеток, в которых может использоваться рассматриваемая система, включают, без ограничения, такие лимфоидные клетки, как В-клетки, Т-клетки (цитотоксические Т-клетки, Т-клетки природных киллеров, регуляторные Т-клетки, Т-хелперные клетки), клетки природных киллеров, индуцированные цитокинами клетки (CIK) (например, см. US 2008/0241194); такие миелоидные клетки, как гранулоциты (базофильные гранулоциты, эозинофильные гранулоциты, нейтрофильные гранулоциты/гиперсегментированные нейтрофилы), моноциты/макрофаги, эритроциты (ретикулоциты), тучные клетки, тромбоциты/мегакарициты, дендритные клетки; клетки эндокринной системы, включая щитовидную железу (эпителиальные, парафолликулярные клетки щитовидной железы), паращитовидную железу (главные клетки, оксифильные клетки паращитовидной железы), надпочечники (хромаффинные клетки), клетки шишковидной железы (пинеалоциты); клетки нервной системы, включая глиальные клетки (астроциты, микроглия), магнoцеллюлярные нейросекреторные клетки, звездчатые клетки, клетки Бёттхера и клетки гипofиза (гонадотропы, кортикотропы, тиротропы, соматотропы, лактотропы); клетки дыхательной системы, включая пневмоциты (пневмоциты I типа, пневмоциты II типа), клетки Клара, бокаловидные клетки, пылевые клетки; клетки системы кровообращения, включая кардиомиоциты, пероциты; клетки пищеварительной системы, включая клетки желудка (главные клетки желудка, париетальные клетки), бокаловидные клетки, клетки Панета, G-клетки, D-клетки, ECL-клетки, I-клетки, K-клетки, S-клетки; энтероэндокринные клетки, включая энтерохромаффинные клетки, клетки APUD, печеночные (гепатоциты, клетки Купфера), хрящевые/костные/мышечные; костные клетки, включая остеобласты, остеоциты, остеокласты, зубные (цементобласты, амелобласты); клетки хрящей, включая хондробласты, хондроциты; клетки кожи, включая трихоциты, кератиноциты, меланоциты (невоциты); мышечные клетки, включая миоциты; клетки мочевой системы, включая подоциты, юкстагломерулярные клетки, интрагломерулярные мезангиальные клетки/экстрагломерулярные мезангиальные клетки, клетки щеточной каемки проксимальных почечных канальцев, клетки macula densa; клетки репродуктивной системы, включая сперматозоиды, клетки Сертоли, клетки Лейдига, яйцеклетки; и другие клетки, включая адипоциты, фибробласты, клетки сухожилий, эпидермальные кератиноциты (дифференцирующиеся эпидермальные клетки), эпидермальные базальные клетки (стволовые клетки), ногтевые

кератиноциты, базальные клетки ногтевого ложа (стволовые клетки), медуллярные клетки стержня волос, кортикальные клетки стержня волос, кутикулярные клетки стержня волос, кутикулярные клетки влагалища корня волос, клетки слоя Хаксли влагалища корня волос, клетки слоя Хенле влагалища корня волос, клетки наружного влагалища корня волос, клетки матрикса волос (стволовые клетки), клетки влажного многослойного барьерного эпителия, поверхностные эпителиальные клетки многослойного плоского эпителия роговицы, ротовой полости, пищевода, анального канала, дистальной уретры и влагалища, базальные клетки (стволовые клетки) эпителия роговицы, ротовой полости, пищевода, анального канала, дистальной уретры и влагалища, клетки мочевого эпителия (выстилающего мочевой пузырь и мочевые протоки), экзокринные клетки секреторного эпителия, мукоидные клетки слюнных желез (секреция полисахаридов), серозные клетки слюнных желез (секреция гликопротеиновых ферментов), клетки желез фон Эбнера на языке (омывают вкусовые сосочки), клетки молочной железы (секреция молока), клетки слезных желез (секреция слез), клетки церуминозных желез в ушах (секреция ушной серы), темные эккринные клетки потовых желез (секреция гликопротеинов), светлые эккринные клетки потовых желез (секреция небольших молекул), апокринные клетки потовых желез (секреция пахучих веществ, чувствительная к половым гормонам), клетки желез Молла в веках (специализированных потовых желез), клетки сальных желез (секреция богатого липидами кожного сала), клетки желез Боумана в носу (омывают обонятельный эпителий), клетки желез Бруннера в двенадцатиперстной кишке (ферменты и щелочная слизь), клетки семенных пузырьков (выделяют компоненты семенной жидкости, в том числе фруктозу для плавающей спермы), клетки предстательной железы (выделяют компоненты семенной жидкости), клетки бульбоуретральных желез (секреция слизи), клетки желез Бартолина (секреция вагинальной смазки), клетки желез Литтре (секреция слизи), клетки эндометрия матки (секреция углеводов), изолированные бокаловидные клетки дыхательного и пищеварительного тракта (секреция слизи), мукоидные клетки выстилки желудка (секреция слизи), зимогенные клетки желудочных желез (секреция пепсиногена), обкладочные клетки желудочных желез (секреция соляной кислоты), ацинарные клетки поджелудочной железы (секреция бикарбоната и пищеварительных ферментов), клетки Панета тонкой кишки (секреция лизоцима), легочные пневмоциты II типа (секреция сурфактанта), клетки Клара в легких, секретирующие гормоны клетки, клетки передней доли гипофиза, соматотропы, лактотропы, тиротропы, гонадотропы, кортикотропы, клетки промежуточной доли гипофиза, магноцеллюлярные нейросекреторные клетки, клетки кишечника и дыхательных путей, клетки щитовидной железы, эпителиальные и парафолликулярные клетки щитовидной железы, клетки парашитовидной железы, главные и оксифильные клетки парашитовидной железы, клетки надпочечников, хромоаффинные клетки, клетки Лейдига в яичках, клетки theca interna фолликул яичников, клетки желтого тела лопнувших фолликул яичников, лютеиновые клетки гранулозы, лютеиновые клетки теки, юкстагломерулярные клетки (секреция ренина), клетки macula densa почек, клетки метаболизма и отложения, клетки с барьерной функцией (легкие, кишечник, экзокринные железы и урогенитальный тракт), почечные клетки, пневмоциты I типа (выстилающие воздушные просветы легких), клетки протоков поджелудочной железы (центроацинарные клетки), клетки гладких мышц протоков (потовых желез, слюнных желез, молочных желез и т.д.), клетки протоков (семявыносящих, предстательной железы и т.д.), эпителиальные клетки, выстилающие закрытые внутренние полости тела, ресничные клетки с пропульсивной функцией, клетки, секретирующие внеклеточный матрикс, сократительные клетки, клетки скелетных мышц, стволовые клетки скелетных мышц, клетки сердечных мышц, клетки крови и иммунной системы, эритроциты, мегакарициты (предшественники тромбоцитов), моноциты, макрофаги соединительной ткани (различного типа), эпидермальные клетки Лангерганса, остеокласты (в костях), дендритные клетки (в лимфоидных тканях), клетки микроглии (в центральной нервной системе), нейтрофильные гранулоциты, эозинофильные гранулоциты, базофильные гранулоциты, тучные клетки, Т-хелперные клетки, Т-супрессорные клетки, цитотоксические Т-клетки, Т-клетки природных киллеров, В-клетки, природные клетки-киллеры, ретикулоциты, стволовые клетки и коммитированные предшественники клеток крови и иммунной системы (различного типа), плюрипотентные стволовые клетки, тотипотентные стволовые клетки, индуцированные плюрипотентные стволовые клетки, взрослые стволовые клетки, сенсорные клетки-преобразователи, клетки вегетативной нервной системы, клетки органов чувств и периферических нервов, нейроны и глиальные клетки центральной нервной системы, клетки хрусталика, пигментные клетки, меланоциты, клетки пигментного эпителия сетчатки, зародышевые клетки, оогонии/ооциты, сперматиды, сперматоциты, клетки-сперматогонии (стволовые клетки для сперматоцитов), сперматозоиды, поддерживающие клетки ("няньки"), клетки фолликул яичников, клетки Сертоли (в яичках), эпителиальные клетки тимуса, интерстициальные клетки и интерстициальные почечные клетки.

Рассматриваемая система, введенная в клетки, может применяться для регулирования экспрессии целевого полинуклеотида (например, экспрессии гена). В одном аспекте изобретения предусмотрены способы регулирования экспрессии целевого полинуклеотида в клетках. В некоторых воплощениях способ включает (a) экспонирование химерного рецепторного полипептида с антигеном, причем (i) химерный рецепторный полипептид подвергается модификации под воздействием антигена и (ii) модификация рецептора включает конформационное изменение или химическую модификацию; (b) связывание химерного адаптерного полипептида с рецептором в ответ на модификацию с образованием комплекса ме-

жду ген-модулирующим полипептидом (GMP) и расщепляющим элементом, причем GMP содержит исполнительный элемент, соединенный с сайтом распознавания расщепления; и (с) расщепление сайта распознавания расщепления расщепляющим элементом, причем при расщеплении сайта распознавания расщепления активируется исполнительный элемент и образует комплекс с целевым полинуклеотидом. В некоторых воплощениях GMP входит в состав внутриклеточной области химерного рецепторного полипептида, а расщепляющий элемент входит в состав химерного адаптерного полипептида. В некоторых воплощениях GMP входит в состав химерного адаптерного полипептида, а расщепляющий элемент входит в состав внутриклеточной части химерного рецепторного полипептида. В некоторых воплощениях расщепляющий элемент образует комплекс со вторым адаптерным полипептидом, который связывает рецептор в ответ на модификацию рецептора, а GMP входит в состав химерного адаптерного полипептида.

Химерный рецепторный полипептид может представлять собой любой химерный рецепторный полипептид, описанный здесь. В некоторых воплощениях химерный рецепторный полипептид является трансмембранным рецептором. Например, химерный полипептид трансмембранного рецептора означает сопряженный с G-белком рецептор (GPCR) типа рецептора Wnt (например, из семейства рецепторов Frizzled); интегриновый рецептор; кадгериновый рецептор; каталитический рецептор, включая рецепторы, обладающие ферментативной активностью, и такие рецепторы, которые, не обладая собственной ферментативной активностью, действуют путем стимуляции нековалентно связанных ферментов (например, киназ); рецептор смерти типа представителей суперсемейства рецепторов фактора некроза опухоли; иммуорецептор типа Т-клеточных рецепторов; либо его производное, вариант или фрагмент. В некоторых воплощениях рецептор не содержит SEQ ID NO: 39.

Экспонирование химерного рецепторного полипептида, экспрессированного в клетке, с антигеном может проводиться *in vitro* и/или *in vivo*. Экспонирование химерного рецепторного полипептида, экспрессированного в клетках, с антигеном может означать приведение рецептора в контакт с антигеном, который может быть мембраносвязанным антигеном или не связанным с мембраной. В некоторых случаях антиген связан с мембраной клетки. В некоторых случаях антиген не связан с мембраной клетки. Экспонирование клеток с антигеном может проводиться *in vitro* путем культивирования клеток, экспрессирующих данную систему, в присутствии антигена. К примеру, клетки, экспрессирующие данную систему, можно культивировать в виде адгезированных клеток или же в суспензии, а антиген добавлять в среду для культивирования клеток. В некоторых случаях антиген экспрессируется клетками мишени, а экспонирование может включать совместное культивирование клеток, экспрессирующих данную систему, и клеток мишени, экспрессирующих антиген. Клетки можно культивировать совместно в подходящих средах для культивирования клеток различного типа, например, с добавками, факторами роста, ионами и т.п. Экспонирование клеток, экспрессирующих данную систему, с клетками мишени (например, клетками мишени, экспрессирующими антиген) может осуществляться *in vivo*, в некоторых случаях, путем введения клеток субъекту, к примеру, человеку, что позволяет клеткам сблизиться с клетками мишени через систему кровообращения.

Экспонирование может проводиться в течение любого подходящего периода времени, к примеру, по меньшей мере 1 мин, по меньшей мере 5 мин, по меньшей мере 10 мин, по меньшей мере 30 мин, по меньшей мере 1 ч, по меньшей мере 2 ч, по меньшей мере 3 ч, по меньшей мере 4 ч, по меньшей мере 5 ч, по меньшей мере 6 ч, по меньшей мере 7 ч, по меньшей мере 8 ч, по меньшей мере 12 ч, по меньшей мере 16 ч, по меньшей мере 20 ч, по меньшей мере 24 ч, по меньшей мере 2 дня, по меньшей мере 3 дня, по меньшей мере 4 дня, по меньшей мере 5 дней, по меньшей мере 6 дней, по меньшей мере 1 неделю, по меньшей мере 2 недели, по меньшей мере 3 недели, по меньшей мере 1 месяц или дольше.

Химерный рецепторный полипептид может связывать любой подходящий антиген, как описано здесь. В некоторых воплощениях химерный рецепторный полипептид является трансмембранным рецептором. В некоторых воплощениях химерный рецепторный полипептид является внутриклеточным рецептором. В некоторых воплощениях химерный рецепторный полипептид является ядерным рецептором. Взаимодействующий с антигеном домен химерного рецепторного полипептида может связывать мембрано-связанный антиген, к примеру, антиген на внеклеточной поверхности клетки (например, клетки мишени). В некоторых воплощениях взаимодействующий с антигеном домен связывает антиген, который не связан с мембраной, к примеру, внеклеточный антиген, который секретируется из клетки (например, клетки мишени), или антиген, локализованный в цитоплазме клетки (например, клетки мишени). Антигены (например, мембраносвязанные и не связанные с мембраной) могут быть связаны с заболеванием типа вирусной, бактериальной и/или паразитарной инфекции; воспалительным и/или аутоиммунным заболеванием; либо неоплазией типа рака и/или опухоли.

После экспозиции с антигеном химерный рецептор может подвергаться модификации. Модификация рецептора может включать конформационное изменение, химическую модификацию либо их комбинацию. Химическая модификация может включать, к примеру, фосфорилирование или дефосфорилирование по меньшей мере одного аминокислотного остатка рецептора. Фосфорилирование и/или дефосфорилирование может происходить, к примеру, по остатку тирозина, серина, треонина или по любому другому подходящему аминокислотному остатку химерного рецепторного полипептида. При связывании химерного адаптерного полипептида с химерным рецепторным полипептидом в ответ на модификацию

рецептора может образоваться комплекс между GMP и расщепляющим элементом. Образование комплекса между GMP и расщепляющим элементом может привести к расщеплению сайта распознавания расщепления расщепляющим элементом. В некоторых воплощениях сайт распознавания расщепления содержит полипептидную последовательность (например, домен расщепления пептидов), которая распознается расщепляющим элементом, содержащим активность протеазы. Расщепляющий элемент может содержать активность протеазы, распознающей эту полипептидную последовательность. Расщепляющим элементом, содержащим протеазную активность, может быть протеаза, включая, без ограничения, любые протеазы, уже описанные здесь, либо их производные, варианты или фрагменты. В некоторых воплощениях сайт распознавания расщепления содержит несколько последовательностей распознавания расщепления, причем каждая последовательность распознавания расщепления может распознаваться одним и тем же или разными расщепляющими элементами, содержащими протеазную активность (например, протеазами). В некоторых воплощениях модификация рецептора включает модификацию по нескольким сайтам модификации, причем каждая модификация эффективна для связывания химерного адаптерного полипептида. В некоторых воплощениях (i) GMP входит в состав химерного адаптерного полипептида, (ii) химерный адаптерный полипептид отделяется от химерного рецепторного полипептида после расщепления сайта распознавания расщепления, и (iii) еще один химерный адаптерный полипептид, содержащий GMP, связывается с модифицированным рецептором.

В некоторых воплощениях сайт распознавания расщепления содержит первую часть последовательности интеина, которая реагирует со второй частью последовательности интеина с высвобождением исполнительного элемента. Для облегчения высвобождения исполнительного элемента пептида может использоваться гетерологическая система с разделенным интеином. Исполнительный элемент может быть ковалентно связан с первой частью последовательности интеина. Исполнительный элемент может быть связан с первой частью последовательности интеина через свой N-конец или C-конец. Расщепляющий элемент может содержать вторую часть последовательности интеина. Первой частью или второй частью последовательности интеина может быть N-концевой интеин, C-концевой интеин или любая другая подходящая часть интеина, которая может способствовать высвобождению исполнительного элемента. Последовательности интеина могут происходить из любого подходящего источника. Первая и вторая часть могут быть из одного источника (например, организма, белка) или из разных источников. В одном показательном примере исполнительный элемент может быть ковалентно связан (например, на своем N-конце или C-конце) пептидной связью с первой частью последовательности интеина, содержащей N-конец интеина. Слитый с N-концом интеина исполнительный элемент может входить в контакт со второй частью последовательности интеина, содержащей C-конец интеина.

При таком контакте между первой и второй частью последовательности интеина может происходить сайт-специфичное расщепление (например, на участке между исполнительным элементом и N-концом интеина), при этом высвобождается исполнительный элемент. В другом показательном примере исполнительный элемент может быть ковалентно связан (например, на своем N-конце или C-конце) пептидной связью с первой частью последовательности интеина, содержащей C-конец интеина. Слитый с C-концом интеина исполнительный элемент может входить в контакт со второй частью последовательности интеина, содержащей N-конец интеина. При таком контакте между первой и второй частью интеина может происходить сайт-специфичное расщепление (например, на подходящем участке между исполнительным элементом и C-концом интеина), при этом высвобождается исполнительный элемент.

В некоторых воплощениях сайт распознавания расщепления содержит дисульфидную связь. В некоторых воплощениях расщепляющий элемент содержит активность оксидоредуктазы. Дисульфидная связь может соединять исполнительный элемент с частью химерного адаптерного полипептида или химерного рецепторного полипептида. Дисульфидная связь может образовываться между одним или несколькими цистеинами исполнительного элемента. Цистеины могут быть встроены в исполнительный элемент. Цистеины могут входить в состав нативной или последовательности дикого типа исполнительного элемента. Цистеины могут присутствовать в линкерном пептиде, присоединенном к исполнительному элементу. Расщеплению дисульфидной связи может способствовать, к примеру, изменение окислительно-восстановительных условий у дисульфидной связи. Изменение окислительно-восстановительных условий может приводить к восстановлению дисульфидной связи до тиолов и высвобождению исполнительного элемента. Расщеплению дисульфидной связи может способствовать расщепляющий элемент, содержащий окислительно-восстановительное средство, которое может вызывать восстановление дисульфидной связи. Окислительно-восстановительным средством может быть фермент либо его производное, вариант или фрагмент. Фермент может быть оксидоредуктазой. Примеры оксидоредуктаз включают протеин-дисульфидредуктазы, тиоредоксины, глутаредоксины, тиол-дисульфид-оксидоредуктазы (например, DsbA, VdbA-D, MdbA и SdbA) и глутатион-дисульфидредуктазы. Окислительно-восстановительное средство может быть из любого подходящего источника, включая прокариотов и эукариотов. Для оптимальной активности фермента могут потребоваться кофакторы (например, никотинамидные кофакторы, флавины либо их производные и аналоги).

GMP, как уже описано здесь, может содержать исполнительный элемент, соединенный с сайтом распознавания расщепления. Исполнительный элемент может содержать нуклеазу (например, ДНК-

нуклеазу и/или РНК-нуклеазу), модифицированную нуклеазу (например, ДНК-нуклеазу и/или РНК-нуклеазу), которая дефектна по нуклеазе или обладает меньшей нуклеазной активностью по сравнению с нуклеазой дикого типа, её вариант, производное или фрагмент, как уже описано здесь. Исполнительный элемент может регулировать экспрессию и/или активность гена или редактировать последовательность нуклеиновой кислоты (например, гена и/или продукта гена). В некоторых воплощениях исполнительный элемент содержит ДНК-нуклеазу типа генно-инженерной (например, программируемой или наводимой) ДНК-нуклеазы, чтобы запускать геномное редактирование целевой последовательности ДНК. В некоторых воплощениях исполнительный элемент содержит РНК-нуклеазу типа генно-инженерной (например, программируемой или наводимой) РНК-нуклеазы, чтобы запускать редактирование целевой последовательности РНК. В некоторых воплощениях исполнительный элемент обладает пониженной или минимальной нуклеазной активностью. Исполнительный элемент с пониженной или минимальной нуклеазной активностью может регулировать экспрессию и/или активность гена путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит безнуклеазный ДНК-связывающий белок, происходящий из ДНК-нуклеазы, который может вызывать активацию или репрессию транскрипции целевой последовательности ДНК. В некоторых воплощениях исполнительный элемент содержит безнуклеазный белок, связывающий РНК, происходящий из РНК-нуклеазы, который может вызывать активацию или репрессию транскрипции целевой последовательности РНК. В некоторых воплощениях исполнительный элемент управляется направляющей нуклеиновой кислотой. В некоторых воплощениях исполнительный элемент управляется направляющей ДНК. В некоторых воплощениях исполнительный элемент управляется направляющей РНК. Исполнительный элемент может регулировать экспрессию или активность гена и/или редактировать последовательность нуклеиновой кислоты, будь то экзогенной или эндогенной. При расщеплении сайта распознавания расщепления исполнительный элемент активируется и образует комплекс с целевым полинуклеотидом.

В исполнительных элементах может использоваться любая подходящая нуклеаза. К подходящим нуклеазам относятся, без ограничения, CRISPR-ассоциированные (Cas) белки или Cas-нуклеазы, в том числе CRISPR-ассоциированные (Cas) полипептиды I типа, CRISPR-ассоциированные (Cas) полипептиды II типа, CRISPR-ассоциированные (Cas) полипептиды III типа, CRISPR-ассоциированные (Cas) полипептиды IV типа, CRISPR-ассоциированные (Cas) полипептиды V типа и CRISPR-ассоциированные (Cas) полипептиды VI типа; нуклеазы с "цинковым пальцем" (ZFN); эффекторные нуклеазы типа активаторов транскрипции (TALEN); мегануклеазы; РНК-связывающие белки (RBP); CRISPR-ассоциированные РНК-связывающие белки; рекомбиназы; флиппазы; транспозазы; белки Argonaute (например, прокариотические Argonaute (pAgo), археальные Argonaute (aAgo) и эукариотические Argonaute (eAgo)); их производные, варианты или фрагменты.

В некоторых воплощениях исполнительный элемент содержит белок Cas, который образует комплекс с направляющей нуклеиновой кислотой типа направляющей РНК. В некоторых воплощениях исполнительный элемент содержит белок, связывающий РНК (RBP), необязательно в комплексе с направляющей нуклеиновой кислотой типа направляющей РНК, который способен образовывать комплекс с белком Cas. В некоторых воплощениях исполнительный элемент содержит белок Cas, лишенный активности расщепления.

Исполнительный элемент рассматриваемой системы может связываться с целевым полинуклеотидом и регулировать экспрессию и/или активность целевого полинуклеотида путем физического ограждения целевого полинуклеотида или рекрутирования дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит активатор транскрипции для эффективного повышения экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит репрессор транскрипции для эффективного снижения экспрессии целевого полинуклеотида.

В некоторых воплощениях целевой полинуклеотид содержит геномную ДНК. В некоторых воплощениях целевой полинуклеотид содержит участок плазмиды, к примеру, плазмиды, несущей экзогенный ген. В некоторых воплощениях целевой полинуклеотид содержит РНК, например, мРНК. В некоторых воплощениях целевой полинуклеотид содержит эндогенный ген или генный продукт. Исполнительный элемент может включать в себя одну или несколько копий сигнала ядерной локализации, что позволяет ему транслоцироваться в ядро клетки при отщеплении от GMP.

Целевой полинуклеотид в различных воплощениях изложенных здесь аспектов может представлять собой ДНК или РНК (например, мРНК). Целевой полинуклеотид может быть одноцепочечным или двухцепочечным. Целевым полинуклеотидом может быть геномная ДНК. Целевым полинуклеотидом может быть любой полинуклеотид, эндогенный или экзогенный для клетки. Например, целевым полинуклеотидом может быть полинуклеотид, находящийся в ядре эукариотической клетки. Целевым полинуклеотидом может быть последовательность, кодирующая продукт гена (например, белок), или же некодирующая последовательность (например, регуляторный полинуклеотид). В некоторых воплощениях целевой полинуклеотид включает участок плазмиды, например, плазмиды, несущей экзогенный ген. В некоторых

воплощениях целевой полинуклеотид содержит РНК, к примеру, мРНК. В некоторых воплощениях целевой полинуклеотид содержит эндогенный ген или генный продукт.

Целевой полинуклеотид может содержать целый ряд связанных с заболеваниями генов и полинуклеотидов, а также гены и полинуклеотиды, связанные с сигнальными биохимическими путями. Примеры целевых полинуклеотидов включают последовательности, связанные с сигнальными биохимическими путями, например гены или полинуклеотиды, связанные с сигнальными биохимическими путями. Примеры целевых полинуклеотидов включают связанные с заболеваниями гены или полинуклеотиды. "Связанный с заболеванием" ген или полинуклеотид означает такой ген или полинуклеотид, который дает продукты транскрипции или трансляции на аномальном уровне или в аномальном виде в клетках из пораженной болезнью ткани по сравнению с тканями или клетками в контроле без болезни. В некоторых воплощениях это ген, который экспрессируется на аномально высоком уровне. В некоторых воплощениях это ген, который экспрессируется на аномально низком уровне. Изменение экспрессии может коррелировать с возникновением и/или прогрессированием заболевания. Связанный с заболеванием ген также означает ген, содержащий мутации или генетические вариации, который непосредственно ответственен или находится в неравновесном сцеплении с генами, которые ответственны за этиологию заболевания. Продукты транскрипции или трансляции могут быть известны или неизвестны и могут быть на нормальном или аномальном уровне.

Примеры связанных с заболеваниями генов и полинуклеотидов можно получить из McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, Md.) и National Center for Biotechnology Information, National Library of Medicine (Bethesda, Md.), которые доступны на World Wide Web. Типичные гены, связанные с определенными заболеваниями и расстройствами, представлены в табл. 3 и 4. Примеры генов и полинуклеотидов, связанных с сигнальными биохимическими путями, приведены в табл. 5.

Мутации в этих генах и путях могут привести к выработыванию неправильных белков или белков в ненадлежащем количестве, что влияет на функцию.

Таблица 3

Заболевание	Гены
Неоплазия	PTEN; ATM; ATR; EGFR; ERBB2; ERBB3; ERBB4; Notch1; Notch2; Notch3; Notch4; AKT; AKT2; AKT3; HIF; HIF1a; HIF3a; Met; HRG; Bcl2; PPAR- α ; PPAR- γ ; WT1 (опухоль Вильмса); гены семейства рецепторов FGF (5 представителей: 1, 2, 3, 4, 5); CDKN2a; APC; RB (ретинобластома); MEN1; VHL; BRCA1; BRCA2; AR (андрогеновый рецептор); TSG101; IGF; рецептор IGF; Igf1 (4 варианта); Igf2 (3 варианта); рецептор Igf 1; рецептор Igf 2; Вах; Bcl2; семейство каспаз (9 представителей: 1, 2, 3, 4, 6, 7, 8, 9, 12); Kras; Arc
Возрастная дегенерация желтого пятна	Abcr; Ccl2; Cc2; ср (шерулоплазмин); Timp3; катепсин D; Vldlr; Ccr2
Шизофрения	нейрегулин1 (Nrg1); Erb4 (рецептор нейрегулина); комплексин 1 (Cplx1); триптофангидроксилаза Tph1; триптофангидроксилаза Tph2; нейрексин 1; GSK3; GSK3a; GSK3b
Заболевания	5-HTT (Slc6a4); COMT; DRD (Drd1a); SLC6A3; DAOA; DTNBP1; Dao (Dao1)
Заболевания с повторами тринуклеотидов	HTT (болезнь Хантингтона); SBMA/SMAX1/AR (болезнь Кеннеди); FXN/X25 (атаксия Фридриха); ATX3 (болезнь Мачадо-Джозефа); ATXN1 и ATXN2 (спиноцеребеллярная атаксия); DMPK (многотоническая дистрофия); атрофин-1 и Atn1 (болезнь DRPLA); CBP (Creb-BP – глобальная нестабильность); VLDLR (болезнь Альцгеймера); Atxn7; Atxn10
Синдром ломкой X-хромосомы	FMR2; FXR1; FXR2; mGLUR5
Нарушения секреции	APH-1 (α и β); пресенилин (Psen1); никастрин (Ncstn); PEN-2
Другие	Nos1; Parp1; Nat1; Nat2
Прионные заболевания	Prp
ALS	SOD1; ALS2; STEX; FUS; TARDBP; VEGF (VEGF-a; VEGF-b; VEGF-c)
Наркотическая зависимость	Prkce (алкоголь); Drd2; Drd4; ABAT (алкоголь); GRIA2; Grm5; Grin1; Htr1b; Grin2a; Drd3; Pdyn; Gria1 (алкоголь)
Аутизм	MeCP2; BZRAP1; MDGA2; Sema5A; нейрексин 1; ломкая X-хромосома (FMR2 (AFF2); FXR1; FXR2; Mglur5)
Болезнь Альцгеймера	E1; CHIP; UCH; UBB; Tau; LRP; PICALM; кластерин; PS1; SORL1; CR1; Vldlr; Uba1; Uba3; CHIP28 (Aqp1, аквапорин 1); Uchl1; Uchl3; APP

Воспаление	IL-10; IL-1 (IL-1a; IL-1b); IL-13; IL-17 (IL-17a (CTLA8); IL-17b; IL-17c; IL-17d; IL-17f); IL-23; Cx3cr1; ptpn22; TNFa; NOD2/CARD15 для IBD; IL-6; IL-12 (IL-12a; IL-12b); CTLA4; Cx3cl1
Болезнь Паркинсона	х-синуклеин; DJ-1; LRRK2; паркин; PINK1

Таблица 4

Гематологические заболевания и коагулопатии	Анемия (CDAN1, CDA1, RPS19, DBA, PKLR, PK1, NT5C3, UMPH1, PSN1, RHAG, RH50A, NRAMP2, SPTB, ALAS2, ANH1, ASB, ABCB7, ABC7, ASAT); синдром голых лимфоцитов (TAPBP, TPSN, TAP2, ABCB3, PSF2, RING11, MHC2TA, C2TA, RFX5, RFXAP, RFX5), коагулопатии (TBXA2R, P2RX1, P2X1); фактор H и фактор-1 типа H (HF1, CFH, HUS); фактор V и фактор VIII (MCFD2); недостаточность фактора VII (F7); недостаточность фактора X (F10); недостаточность фактора XI (F11); недостаточность фактора XII (F12, HAF); недостаточность фактора XIII (F13A1, F13A); недостаточность фактора XIII (F13B); анемия Фанкони (FANCA, FACA, FA1, FA, FAA, FAAP95, FAAP90, FLJ34064, FANCB, FANCC, FACC, BRCA2, FANCD1, FANCD2, FANCD, FACD, FAD, FANCE, FACE, FANCF, XRCC9, FANCG, BRIP1, BACH1, FANCI, PHF9, FANCL, FANCM, KIAA1596); гемофагоцитарные лимфогистиоцитозы (PRF1, HPLH2, UNC13D, MUNC13-4, HPLH3, HLH3, FHL3); гемофилия A (F8, F8C, HEMA); гемофилия B (F9, HEMB); геморрагические заболевания (PI, ATT, F5); лейкоцитарные дефекты и нарушения (ITGB2, CD18, LCAMB, LAD, EIF2B1, EIF2BA, EIF2B2, EIF2B3, EIF2B5, LVWM, CACH, CLE, EIF2B4); серповидноклеточная анемия (HBB); талассемия (HBA2, HBB, HBD, LCRB, HBA1)
Нарушения клеточной регуляции и онкологические заболевания	В-клеточная неходжкинская лимфома (BCL7A, BCL7); лейкемия (TAL1, TCL5, SCL, TAL2, FLT3, NBS1, NBS, ZNFN1A1, IK1, LYF1, HOXD4, HOX4B, BCR, CML, PHL, ALL, ARNT, KRAS2, RASK2, GMPS, AF10, ARHGEF12, LARG, KIAA0382, CALM, CLTH, CEBPA, CEBP, CHIC2, BTL, FLT3, KIT, PBT, LPP, NPM1, NUP214, D9S46E, CAN, CAIN, RUNX1, CBFA2, AML1, WHSC1L1, NSD3, FLT3, AF1Q, NPM1, NUMA1, ZNF145, PLZF, PML, MYL, STAT5B, AF10, CALM, CLTH, ARL11, ARLTS1, P2RX7, P2X7, BCR, CML, PHL, ALL, GRAF, NF1, VRNF, WSS, NFNS, PTPN11, PTP2C, SHP2, NS1, BCL2, CCND1, PRAD1, BCL1, TCRA, GATA1, GF1, ERYF1, NFE1, ABL1, NQO1, DIA4, NMOR1, NUP214, D9S46E, CAN, CAIN)
Воспалительные и иммунологические заболевания	СПИД (KIR3DL1, NKAT3, NKВ1, AMB11, KIR3DS1, IFNG, CXCL12, SDF1); аутоиммунный лимфопролиферативный синдром (TNFRSF6, APT1, FAS, CD95, ALPS1A); комбинированный иммунодефицит (IL2RG, SCIDX1, SCIDX, IMD4); HIV-1 (CCL5, SCYA5, D17S136E, TCP228), восприимчивость к ВИЧ или ВИЧ-инфекция (IL10, CSIF, CMKBR2, CCR2, CMKBR5, CCCKR5 (CCR5)); иммунодефициты (CD3E, CD3G,

	AICDA, AID, HIGM2, TNFRSF5, CD40, UNG, DGU, HIGM4, TNFSF5, CD40LG, HIGM1, IGM, FOXP3, IPEX, AIID, XPID, PIDX, TNFRSF14B, TAC1); воспаление (IL-10, IL-1 (IL-1a, IL-1b), IL-13, IL-17 (IL-17a (CTLA8), IL-17b, IL-17c, IL-17d, IL-17f), IL-23, Cx3cr1, ptpn22, TNFa, NOD2/CARD15 для IBD, IL-6, IL-12 (IL-12a, IL-12b), CTLA4, Cx3cl1); тяжелые комбинированные иммунодефициты (SCIDs) (JAK3, JAKL, DCLRE1C, ARTEMIS, SCIDA, RAG1, RAG2, ADA, PTPRC, CD45, LCA, IL7R, CD3D, T3D, IL2RG, SCIDX1, SCIDX, IMD4)
Метаболические, печеночные, почечные и белковые заболевания	Амилоидная невропатия (TTR, PALB); амилоидоз (APOA1, APP, AAA, CVAP, AD1, GSN, FGA, LYZ, TTR, PALB); цирроз (KRT18, KRT8, CIRH1A, NAIC, TEX292, KIAA1988); кистозный фиброз (CFTR, ABCC7, CF, MRP7); болезни накопления гликогена (SLC2A2, GLUT2, G6PC, G6PT, G6PT1, GAA, LAMP2, LAMPB, AGL, GDE, GBE1, GYS2, PYGL, PFKM); аденома печени 142330 (TCF1, HNF1A, MODY3), печеночная недостаточность, ранняя форма, и неврологические расстройства (SCOD1, SCO1), недостаточность печеночной липазы (LIPC), гепатобластома, рак и карцинома печени (CTNNB1, PDGFRL, PDGRL, PRLTS, AXIN1, AXIN, CTNNB1, TP53, P53, LFS1, IGF2R, MPRI, MET, CASP8, MCH5; медуллярная кистозная болезнь почек (UMOD, HNFJ, FJHN, MCKD2, ADMCKD2); фенилкетонурия (PAH, PKU1, QDPR, DHPR, PTS); поликистозная болезнь почек и печени (FCYT, PKHD1, ARPKD, PKD1, PKD2, PKD4, PKDTS, PRKCSH, G19P1, PCLD, SEC63)
Мышечные и скелетные заболевания	Мышечная дистрофия Беккера (DMD, BMD, MYF6), мышечная дистрофия Дюшенна (DMD, BMD); мышечная дистрофия Эмери-Дрейфуса (LMNA, LMN1, EMD2, FPLD, CMD1A, HGPS, LGMD1B, LMNA, LMN1, EMD2, FPLD, CMD1A); фациоскапулогумеральная мышечная дистрофия (FSHMD1A, FSHD1A); мышечная дистрофия (FKRP, MDC1C, LGMD2I, LAMA2, LAMM, LARGE, KIAA0609, MDC1D, FCMD, TTID, MYOT, CAPN3, CANP3, DYSF, LGMD2B, SGCG, LGMD2C, DMDA1, SCG3, SGCA, ADL, DAG2, LGMD2D, DMDA2, SGCB, LGMD2E, SGCD, SGD, LGMD2F, CMD1L, TCAP, LGMD2G, CMD1N, TRIM32, HT2A, LGMD2H, FKRP, MDC1C, LGMD2I, TTN, CMD1G, TMD, LGMD2J, POMT1, CAV3, LGMD1C, SEPN1, SELN, RSMD1, PLEC1, PLTN, EBS1); остеопетроз (LRP5, BMND1, LRP7, LR3, OPPG, VBCH2, CLCN7, CLC7, OPTA2, OSTM1, GL, TCIRG1, TIRC7, OC116, OPTB1); мышечная атрофия (VAPB, VAPC, ALS8, SMN1, SMA1, SMA2, SMA3, SMA4, BSCL2, SPG17, GARS, SMAD1, CMT2D, HEXB, IGHMBP2, SMUBP2, CATF1, SMARD1)
Неврологические и нейронные заболевания	ALS (SOD1, ALS2, STEX, FUS, TARDBP, VEGF (VEGF-a, VEGF-b, VEGF-c); болезнь Альцгеймера (APP, AAA, CVAP, AD1, APOE, AD2, PSEN2, AD4, STM2, APBB2, FE65L1, NOS3, PLAUI, URK, ACE, DCP1, ACE1, MPO, PACIP1, PAXIP1L, PTIP, A2M, BLMH, BMH, PSEN1, AD3); аутизм (MeCP2, BZRAP1, MDGA2, Sema5A, нейрексин 1, GLO1, MECP2, RTT, PPMX, MRX16, MRX79, NLGN3, NLGN4, KIAA1260,

	<p>AUTSX2); синдром ломкой X-хромосомы (FMR2, FXR1, FXR2, mGLUR5); болезнь Хантингтона и сходные заболевания (HD, IT15, PRNP, PRIP, JPH3, JP3, HDL2, TBP, SCA17); болезнь Паркинсона (NR4A2, NURR1, NOT, TINUR, SNCAIP, TBP, SCA17, SNCA, NACP, PARK1, PARK4, DJ1, PARK7, LRRK2, PARK8, PINK1, PARK6, UCHL1, PARK5, SNCA, NACP, PARK1, PARK4, PRKN, PARK2, PDJ, DBH, NDUFV2); синдром Ретта (MECP2, RTT, PPMX, MRX16, MRX79, CDKL5, STK9, MECP2, RTT, PPMX, MRX16, MRX79, х-синуклеин, DJ-1); шизофрения (нейрегулин 1 (Nrg1), Erb4 (рецептор нейрегулина), комплексин 1 (Cplx1), триптофангидроксилаза Trp1, триптофангидроксилаза Trp2, нейрексин 1, GSK3, GSK3a, GSK3b, 5-HTT (Slc6a4), COMT, DRD (Drd1a), SLC6A3, DAOA, DTNBP1, Dao (Dao1)); нарушения секреции (APH-1 (α и β), пресенилин (Psen1), никастрин (Ncstn), PEN-2, Nos1, Parp1, Nat1, Nat2); заболевания с повторами тринуклеотидов (HTT (болезнь Хантингтона), SBMA/SMAX1/AR (болезнь Кеннеди), FXN/X25 (атаксия Фридриха), ATX3 (болезнь Мачадо-Джозефа), ATXN1 и ATXN2 (спиноцереbellарная атаксия), DMPK (миотоническая дистрофия), атрофин-1 и Atn1 (болезнь DRPLA), CBP (Creb-BP – глобальная нестабильность), VLDLR (болезнь Альцгеймера), Atxn7, Atxn10)</p>
Глазные заболевания и нарушения	<p>возрастная дегенерация желтого пятна (Abcr, Ccl2, Cc2, сr (церулоплазмин), Timpr3, катепсин D, Vldlr, Ccr2); катаракта (CRYAA, CRYA1, CRYBB2, CRYB2, PITX3, BFSP2, CP49, CP47, CRYAA, CRYA1, PAX6, AN2, MGDA, CRYBA1, CRYB1, CRYGC, CRYG3, CCL, LIM2, MP19, CRYGD, CRYG4, BFSP2, CP49, CP47, HSF4, CTM, HSF4, CTM, MIP, AQP0, CRYAB, CRYA2, STPP2, CRYBB1, CRYGD, CRYG4, CRYBB2, CRYB2, CRYGC, CRYG3, CCL, CRYAA, CRYA1, GJA8, CX50, CAE1, GJA3, CX46, CZP3, CAE3, CCM1, CAM, KRIT1); помутнение и дистрофия роговицы (APOA1, TGFB1, CSD2, CDGG1, CSD, BIGH3, CDG2, TACSTD2, TROP2, M1S1, VSX1, RINX, PPCD, PPD, KTCN, COL8A2, FECD, PPCD2, PIP5K3, CFD); врожденная плоская роговица (KERA, CNA2); глаукома (MYOC, TIGR, GLC1A, JOAG, GPOA, OPTN, GLC1E, FIP2, HYPL, NRP, CYP1B1, GLC3A, OPA1, NTG, NPG, CYP1B1, GLC3A); врожденный амавроз Лебера (CRB1, RP12, CRX, CORD2, CRD, RPGRIP1, LCA6, CORD9, RPE65, RP20, AIPL1, LCA4, GUCY2D, GUC2D, LCA1, CORD6, RDH12, LCA3); дистрофия желтого пятна (ELOVL4, ADMD, STGD2, STGD3, RDS, RP7, PRPH2, PRPH, AVMD, AOFMD, VMD2)</p>

Таблица 5

Клеточная функция	Гены
Сигнализация PI3K/AKT	PRKCE; ITGAM; ITGA5; IRAK1; PRKAA2; EIF2AK2; PTEN; EIF4E; PRKCZ; GRK6; MAPK1; TSC1; PLK1; AKT2; IKKBK; PIK3CA; CDK8; CDKN1B; NFKB2;

	BCL2; PIK3CB; PPP2R1A; MAPK8; BCL2L1; MAPK3; TSC2; ITGA1; KRAS; EIF4EBP1; RELA; PRKCD; NOS3; PRKAA1; MAPK9; CDK2; PPP2CA; PIM1; ITGB7; YWHAZ; ILK; TP53; RAF1; IKBKG; RELB; DYRK1A; CDKN1A; ITGB1; MAP2K2; JAK1; AKT1; JAK2; PIK3R1; CHUK; PDPK1; PPP2R5C; CTNNA1; MAP2K1; NFKB1; PAK3; ITGB3; CCND1; GSK3A; FRAP1; SFN; ITGA2; TTK; CSNK1A1; BRAF; GSK3B; AKT3; FOXO1; SGK; HSP90AA1; RPS6KB1
Сигнализация ERK/МАРК	PRKCE; ITGAM; ITGA5; HSPB1; IRAK1; PRKAA2; EIF2AK2; RAC1; RAP1A; TLN1; EIF4E; ELK1; GRK6; MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8; CREB1; PRKCI; PTK2; FOS; RPS6KA4; PIK3CB; PPP2R1A; PIK3C3; MAPK8; MAPK3; ITGA1; ETS1; KRAS; MYCN; EIF4EBP1; PPARG; PRKCD; PRKAA1; MAPK9; SRC; CDK2; PPP2CA; PIM1; PIK3C2A; ITGB7; YWHAZ; PPP1CC; KSR1; PXN; RAF1; FYN; DYRK1A; ITGB1; MAP2K2; PAK4; PIK3R1; STAT3; PPP2R5C; MAP2K1; PAK3; ITGB3; ESR1; ITGA2; MYC; TTK; CSNK1A1; CRKL; BRAF; ATF4; PRKCA; SRF; STAT1; SGK
Сигнализация рецепторов глюкокортикоидов	RAC1; TAF4B; EP300; SMAD2; TRAF6; PCAF; ELK1; MAPK1; SMAD3; AKT2; IKBKB; NCOR2; UBE2I; PIK3CA; CREB1; FOS; HSPA5; NFKB2; BCL2; MAP3K14; STAT5B; PIK3CB; PIK3C3; MAPK8; BCL2L1; MAPK3; TSC22D3; MAPK10; NRIP1; KRAS; MAPK13; RELA; STAT5A; MAPK9; NOS2A; PBX1; NR3C1; PIK3C2A; CDKN1C; TRAF2; SERPINE1; NCOA3; MAPK14; TNF; RAF1; IKBKG; MAP3K7; CREBBP; CDKN1A; MAP2K2; JAK1; IL8; NCOA2; AKT1; JAK2; PIK3R1; CHUK; STAT3; MAP2K1; NFKB1; TGFBR1; ESR1; SMAD4; CEBPB; JUN; AR; AKT3; CCL2; MMP1; STAT1; IL6; HSP90AA1
Сигнализация наведения аксонов	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; ADAM12; IGF1; RAC1; RAP1A; EIF4E; PRKCZ; NRP1; NTRK2; ARHGEF7; SMO; ROCK2; MAPK1; PGF; RAC2; PTPN11; GNAS; AKT2; PIK3CA; ERBB2; PRKCI; PTK2; CFL1; GNAQ; PIK3CB; CXCL12; PIK3C3; WNT11; PRKD1; GNB2L1; ABL1; MAPK3; ITGA1; KRAS; RHOA; PRKCD; PIK3C2A; ITGB7; GLI2; PXN; VASP; RAF1; FYN; ITGB1; MAP2K2; PAK4; ADAM17; AKT1; PIK3R1; GLI1; WNT5A; ADAM10; MAP2K1; PAK3; ITGB3; CDC42; VEGFA; ITGA2; EPHA8; CRKL; RND1; GSK3B; AKT3; PRKCA

Сигнализация эфриновых рецепторов	PRKCE; ITGAM; ROCK1; ITGA5; CXCR4; IRAK1; PRKAA2; EIF2AK2; RAC1; RAP1A; GRK6; ROCK2; MAPK1; PGF; RAC2; PTPN11; GNAS; PLK1; AKT2; DOK1; CDK8; CREB1; PTK2; CFL1; GNAQ; MAP3K14; CXCL12; MAPK8; GNB2L1; ABL1; MAPK3; ITGA1; KRAS; RHOA; PRKCD; PRKAA1; MAPK9; SRC; CDK2; PIM1; ITGB7; PXN; RAF1; FYN; DYRK1A; ITGB1; MAP2K2; PAK4; AKT1; JAK2; STAT3; ADAM10; MAP2K1; PAK3; ITGB3; CDC42; VEGFA; ITGA2; EPHA8; TTK; CSNK1A1; CRKL; BRAF; PTPN13; ATF4; AKT3; SGK
Сигнализация актинового цитоскелета	ACTN4; PRKCE; ITGAM; ROCK1; ITGA5; IRAK1; PRKAA2; EIF2AK2; RAC1; INS; ARHGEF7; GRK6; ROCK2; MAPK1; RAC2; PLK1; AKT2; PIK3CA; CDK8; PTK2; CFL1; PIK3CB; MYH9; DIAPH1; PIK3C3; MAPK8; F2R; MAPK3; SLC9A1; ITGA1; KRAS; RHOA; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; ITGB7; PPP1CC; PXN; VIL2; RAF1; GSN; DYRK1A; ITGB1; MAP2K2; PAK4; PIP5K1A; PIK3R1; MAP2K1; PAK3; ITGB3; CDC42; APC; ITGA2; TTK; CSNK1A1; CRKL; BRAF; VAV3; SGK
Сигнализация болезни Хантингтона	PRKCE; IGF1; EP300; RCOR1; PRKCZ; HDAC4; TGM2; MAPK1; CAPNS1; AKT2; EGFR; NCOR2; SP1; CAPN2; PIK3CA; HDAC5; CREB1; PRKC1; HSPA5; REST; GNAQ; PIK3CB; PIK3C3; MAPK8; IGF1R; PRKD1; GNB2L1; BCL2L1; CAPN1; MAPK3; CASP8; HDAC2; HDAC7A; PRKCD; HDAC11; MAPK9; HDAC9; PIK3C2A; HDAC3; TP53; CASP9; CREBBP; AKT1; PIK3R1; PDPK1; CASP1; APAF1; FRAP1; CASP2; JUN; BAX; ATF4; AKT3; PRKCA; CLTC; SGK; HDAC6; CASP3
Сигнализация апоптоза	PRKCE; ROCK1; BID; IRAK1; PRKAA2; EIF2AK2; BAK1; BIRC4; GRK6; MAPK1; CAPNS1; PLK1; AKT2; IKBKB; CAPN2; CDK8; FAS; NFKB2; BCL2; MAP3K14; MAPK8; BCL2L1; CAPN1; MAPK3; CASP8; KRAS; RELA; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; TP53; TNF; RAF1; IKBKG; RELB; CASP9; DYRK1A; MAP2K2; CHUK; APAF1; MAP2K1; NFKB1; PAK3; LMNA; CASP2; BIRC2; TTK; CSNK1A1; BRAF; BAX; PRKCA; SGK; CASP3; BIRC3; PARP1
Сигнализация В-клеточных рецепторов	RAC1; PTEN; LYN; ELK1; MAPK1; RAC2; PTPN11; AKT2; IKBKB; PIK3CA; CREB1; SYK; NFKB2; CAMK2A; MAP3K14; PIK3CB; PIK3C3; MAPK8; BCL2L1; ABL1; MAPK3; ETS1; KRAS; MAPK13; RELA; PTPN6; MAPK9; EGR1; PIK3C2A; BTK; MAPK14; RAF1; IKBKG; RELB; MAP3K7; MAP2K2; AKT1; PIK3R1; CHUK; MAP2K1; NFKB1; CDC42; GSK3A; FRAP1; BCL6; BCL10; JUN; GSK3B; ATF4; AKT3; VAV3; RPS6KB1
Сигнализация экстравазации лейкоцитов	ACTN4; CD44; PRKCE; ITGAM; ROCK1; CXCR4; CYBA; RAC1; RAP1A; PRKCZ; ROCK2; RAC2; PTPN11;

	MMP14; PIK3CA; PRKCI; PTK2; PIK3CB; CXCL12; PIK3C3; MAPK8; PRKD1; ABL1; MAPK10; CYBB; MAPK13; RHOA; PRKCD; MAPK9; SRC; PIK3C2A; BTK; MAPK14; NOX1; PXN; VIL2; VASP; ITGB1; MAP2K2; CTNND1; PIK3R1; CTNNB1; CLDN1; CDC42; F11R; ITK; CRKL; VAV3; CTTN; PRKCA; MMP1; MMP9
Интегриновая сигнализация	ACTN4; ITGAM; ROCK1; ITGA5; RAC1; PTEN; RAP1A; TLN1; ARHGEF7; MAPK1; RAC2; CAPNS1; AKT2; CAPN2; PIK3CA; PTK2; PIK3CB; PIK3C3; MAPK8; CAV1; CAPN1; ABL1; MAPK3; ITGA1; KRAS; RHOA; SRC; PIK3C2A; ITGB7; PPP1CC; ILK; PXN; VASP; RAF1; FYN; ITGB1; MAP2K2; PAK4; AKT1; PIK3R1; TNK2; MAP2K1; PAK3; ITGB3; CDC42; RND3; ITGA2; CRKL; BRAF; GSK3B; AKT3
Сигнализация острой фазы	IRAK1; SOD2; MYD88; TRAF6; ELK1; MAPK1; PTPN11; AKT2; IKKBK; PIK3CA; FOS; NFKB2; MAP3K14; PIK3CB; MAPK8; RIPK1; MAPK3; IL6ST; KRAS; MAPK13; IL6R; RELA; SOCS1; MAPK9; FTL; NR3C1; TRAF2; SERPINE1; MAPK14; TNF; RAF1; PDK1; IKKBK; RELB; MAP3K7; MAP2K2; AKT1; JAK2; PIK3R1; CHUK; STAT3; MAP2K1; NFKB1; FRAP1; CEBPB; JUN; AKT3; IL1R1; IL6
Сигнализация PTEN	ITGAM; ITGA5; RAC1; PTEN; PRKCZ; BCL2L11; MAPK1; RAC2; AKT2; EGFR; IKKBK; CBL; PIK3CA; CDKN1B; PTK2; NFKB2; BCL2; PIK3CB; BCL2L1; MAPK3; ITGA1; KRAS; ITGB7; ILK; PDGFRB; INSR; RAF1; IKKBK; CASP9; CDKN1A; ITGB1; MAP2K2; AKT1; PIK3R1; CHUK; PDGFRA; PDPK1; MAP2K1; NFKB1; ITGB3; CDC42; CCND1; GSK3A; ITGA2; GSK3B; AKT3; FOXO1; CASP3; RPS6KB1
Сигнализация p53	PTEN; EP300; BBC3; PCAF; FASN; BRCA1; GADD45A; BIRC5; AKT2; PIK3CA; CHEK1; TP53INP1; BCL2; PIK3CB; PIK3C3; MAPK8; THBS1; ATR; BCL2L1; E2F1; PMAIP1; CHEK2; TNFRSF10B; TP73; RB1; HDAC9; CDK2; PIK3C2A; MAPK14; TP53; LRDD; CDKN1A; HIPK2; AKT1; PIK3R1; RRM2B; APAF1; CTNNB1; SIRT1; CCND1; PRKDC; ATM; SFN; CDKN2A; JUN; SNAI2; GSK3B; BAX; AKT3
Сигнализация рецепторов арилуглевородов	HSPB1; EP300; FASN; TGM2; RXRA; MAPK1; NQO1; NCOR2; SP1; ARNT; CDKN1B; FOS; CHEK1; SMARCA4; NFKB2; MAPK8; ALDH1A1; ATR; E2F1; MAPK3; NRIP1; CHEK2; RELA; TP73; GSTP1; RB1; SRC; CDK2; AHR; NFE2L2; NCOA3; TP53; TNF; CDKN1A; NCOA2; APAF1; NFKB1; CCND1; ATM; ESR1; CDKN2A; MYC; JUN; ESR2; BAX; IL6; CYP1B1; HSP90AA1
Сигнализация метаболизма ксенобиотиков	PRKCE; EP300; PRKCZ; RXRA; MAPK1; NQO1; NCOR2; PIK3CA; ARNT; PRKCI; NFKB2; CAMK2A; PIK3CB; PPP2R1A; PIK3C3; MAPK8; PRKD1; ALDH1A1; MAPK3; NRIP1; KRAS; MAPK13; PRKCD;

	GSTP1; MAPK9; NOS2A; ABCB1; AHR; PPP2CA; FTL; NFE2L2; PIK3C2A; PPARGC1A; MAPK14; TNF; RAF1; CREBBP; MAP2K2; PIK3R1; PPP2R5C; MAP2K1; NFKB1; KEAP1; PRKCA; EIF2AK3; IL6; CYP1B1; HSP90AA1
Сигнализация SAPK/JNK	PRKCE; IRAK1; PRKAA2; EIF2AK2; RAC1; ELK1; GRK6; MAPK1; GADD45A; RAC2; PLK1; AKT2; PIK3CA; FADD; CDK8; PIK3CB; PIK3C3; MAPK8; RIPK1; GNB2L1; IRS1; MAPK3; MAPK10; DAXX; KRAS; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; TRAF2; TP53; LCK; MAP3K7; DYRK1A; MAP2K2; PIK3R1; MAP2K1; PAK3; CDC42; JUN; TTK; CSNK1A1; CRKL; BRAF; SGK
Сигнализация PPAγ/RXR	PRKAA2; EP300; INS; SMAD2; TRAF6; PPARA; FASN; RXRA; MAPK1; SMAD3; GNAS; IKBKB; NCOR2; ABCA1; GNAQ; NFKB2; MAP3K14; STAT5B; MAPK8; IRS1; MAPK3; KRAS; RELA; PRKAA1; PPARGC1A; NCOA3; MAPK14; INSR; RAF1; IKBKG; RELB; MAP3K7; CREBBP; MAP2K2; JAK2; CHUK; MAP2K1; NFKB1; TGFBR1; SMAD4; JUN; IL1R1; PRKCA; IL6; HSP90AA1; ADIPOQ
Сигнализация NF-KB	IRAK1; EIF2AK2; EP300; INS; MYD88; PRKCZ; TRAF6; TBK1; AKT2; EGFR; IKBKB; PIK3CA; BTRC; NFKB2; MAP3K14; PIK3CB; PIK3C3; MAPK8; RIPK1; HDAC2; KRAS; RELA; PIK3C2A; TRAF2; TLR4; PDGFRB; TNF; INSR; LCK; IKBKG; RELB; MAP3K7; CREBBP; AKT1; PIK3R1; CHUK; PDGFRA; NFKB1; TLR2; BCL10; GSK3B; AKT3; TNFAIP3; IL1R1
Сигнализация нейрорегулинов	ERBB4; PRKCE; ITGAM; ITGA5; PTEN; PRKCZ; ELK1; MAPK1; PTPN11; AKT2; EGFR; ERBB2; PRKCI; CDKN1B; STAT5B; PRKD1; MAPK3; ITGA1; KRAS; PRKCD; STAT5A; SRC; ITGB7; RAF1; ITGB1; MAP2K2; ADAM17; AKT1; PIK3R1; PDPK1; MAP2K1; ITGB3; EREG; FRAP1; PSEN1; ITGA2; MYC; NRG1; CRKL; AKT3; PRKCA; HSP90AA1; RPS6KB1
Сигнализация Wnt и β-катенина	CD44; EP300; LRP6; DVL3; CSNK1E; GJA1; SMO; AKT2; PIN1; CDH1; BTRC; GNAQ; MARK2; PPP2R1A; WNT11; SRC; DKK1; PPP2CA; SOX6; SFRP2; ILK; LEF1; SOX9; TP53; MAP3K7; CREBBP; TCF7L2; AKT1; PPP2R5C; WNT5A; LRP5; CTNNB1; TGFBR1; CCND1; GSK3A; DVL1; APC; CDKN2A; MYC; CSNK1A1; GSK3B; AKT3; SOX2
Инсулиновый рецептор	PTEN; INS; EIF4E; PTPN11; PRKCZ; MAPK1; TSC1; PTPN11; AKT2; CBL; PIK3CA; PRKCI; PIK3CB; PIK3C3; MAPK8; IRS1; MAPK3; TSC2; KRAS; EIF4EBP1; SLC2A4; PIK3C2A; PPP1CC; INSR; RAF1; FYN; MAP2K2; JAK1; AKT1; JAK2; PIK3R1; PDPK1; MAP2K1; GSK3A; FRAP1; CRKL; GSK3B; AKT3; FOXO1; SGK; RPS6KB1
Сигнализация IL-6	HSPB1; TRAF6; MAPKAPK2; ELK1; MAPK1; PTPN11; IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK3;

	MAPK10; IL6ST; KRAS; MAPK13; IL6R; RELA; SOCS1; MAPK9; ABCB1; TRAF2; MAPK14; TNF; RAF1; IKBKG; RELB; MAP3K7; MAP2K2; IL8; JAK2; CHUK; STAT3; MAP2K1; NFKB1; CEBPB; JUN; IL1R1; SRF; IL6
Печеночный холестаг	PRKCE; IRAK1; INS; MYD88; PRKCZ; TRAF6; PPARA; RXRA; IKBKB; PRKCI; NFKB2; MAP3K14; MAPK8; PRKD1; MAPK10; RELA; PRKCD; MAPK9; ABCB1; TRAF2; TLR4; TNF; INSR; IKBKG; RELB; MAP3K7; IL8; CHUK; NR1H2; TJP2; NFKB1; ESRI; SREBF1; FGFR4; JUN; IL1R1; PRKCA; IL6
Сигнализация IGF-1	IGF1; PRKCZ; ELK1; MAPK1; PTPN11; NEDD4; AKT2; PIK3CA; PRKCI; PTK2; FOS; PIK3CB; PIK3C3; MAPK8; IGF1R; IRS1; MAPK3; IGFBP7; KRAS; PIK3C2A; YWHAZ; PXN; RAF1; CASP9; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; IGFBP2; SFN; JUN; CYR61; AKT3; FOXO1; SRF; CTGF; RPS6KB1
Опосредованная NRF2 реакция на окислительный стресс	PRKCE; EP300; SOD2; PRKCZ; MAPK1; SQSTM1; NQO1; PIK3CA; PRKCI; FOS; PIK3CB; PIK3C3; MAPK8; PRKD1; MAPK3; KRAS; PRKCD; GSTP1; MAPK9; FTL; NFE2L2; PIK3C2A; MAPK14; RAF1; MAP3K7; CREBBP; MAP2K2; AKT1; PIK3R1; MAP2K1; PPIB; JUN; KEAP1; GSK3B; ATF4; PRKCA; EIF2AK3; HSP90AA1
Активация фиброза печени/ звездчатых клеток печени	EDN1; IGF1; KDR; FLT1; SMAD2; FGFR1; MET; PGF; SMAD3; EGFR; FAS; CSF1; NFKB2; BCL2; MYH9; IGF1R; IL6R; RELA; TLR4; PDGFRB; TNF; RELB; IL8; PDGFRA; NFKB1; TGFB1; SMAD4; VEGFA; BAX; IL1R1; CCL2; HGF; MMP1; STAT1; IL6; CTGF; MMP9
Сигнализация PPAR	EP300; INS; TRAF6; PPARA; RXRA; MAPK1; IKBKB; NCOR2; FOS; NFKB2; MAP3K14; STAT5B; MAPK3; NRIP1; KRAS; PPARG; RELA; STAT5A; TRAF2; PPARGC1A; PDGFRB; TNF; INSR; RAF1; IKBKG; RELB; MAP3K7; CREBBP; MAP2K2; CHUK; PDGFRA; MAP2K1; NFKB1; JUN; IL1R1; HSP90AA1
Сигнализация Fcε-R1	PRKCE; RAC1; PRKCZ; LYN; MAPK1; RAC2; PTPN11; AKT2; PIK3CA; SYK; PRKCI; PIK3CB; PIK3C3; MAPK8; PRKD1; MAPK3; MAPK10; KRAS; MAPK13; PRKCD; MAPK9; PIK3C2A; BTK; MAPK14; TNF; RAF1; FYN; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; AKT3; VAV3; PRKCA
Сигнализация рецепторов, сопряженных с G-белком	PRKCE; RAP1A; RGS16; MAPK1; GNAS; AKT2; IKBKB; PIK3CA; CREB1; GNAQ; NFKB2; CAMK2A; PIK3CB; PIK3C3; MAPK3; KRAS; RELA; SRC; PIK3C2A; RAF1; IKBKG; RELB; FYN; MAP2K2; AKT1; PIK3R1; CHUK; PDPK1; STAT3; MAP2K1; NFKB1; BRAF; ATF4; AKT3; PRKCA
Метаболизм инозитолфосфата	PRKCE; IRAK1; PRKAA2; EIF2AK2; PTEN; GRK6; MAPK1; PLK1; AKT2; PIK3CA; CDK8; PIK3CB; PIK3C3; MAPK8; MAPK3; PRKCD; PRKAA1; MAPK9; CDK2; PIM1; PIK3C2A; DYRK1A; MAP2K2; PIP5K1A;

	PIK3R1; MAP2K1; PAK3; ATM; TTK; CSNK1A1; BRAF; SGK
Сигнализация PDGF	EIF2AK2; ELK1; ABL2; MAPK1; PIK3CA; FOS; PIK3CB; PIK3C3; MAPK8; CAV1; ABL1; MAPK3; KRAS; SRC; PIK3C2A; PDGFRB; RAF1; MAP2K2; JAK1; JAK2; PIK3R1; PDGFRA; STAT3; SPHK1; MAP2K1; MYC; JUN; CRKL; PRKCA; SRF; STAT1; SPHK2
Сигнализация VEGF	ACTN4; ROCK1; KDR; FLT1; ROCK2; MAPK1; PGF; AKT2; PIK3CA; ARNT; PTK2; BCL2; PIK3CB; PIK3C3; BCL2L1; MAPK3; KRAS; HIF1A; NOS3; PIK3C2A; PXN; RAF1; MAP2K2; ELAVL1; AKT1; PIK3R1; MAP2K1; SFN; VEGFA; AKT3; FOXO1; PRKCA
Сигнализация клеток природных киллеров	PRKCE; RAC1; PRKCZ; MAPK1; RAC2; PTPN11; KIR2DL3; AKT2; PIK3CA; SYK; PRKCI; PIK3CB; PIK3C3; PRKD1; MAPK3; KRAS; PRKCD; PTPN6; PIK3C2A; LCK; RAF1; FYN; MAP2K2; PAK4; AKT1; PIK3R1; MAP2K1; PAK3; AKT3; VAV3; PRKCA
Регуляция клеточного цикла: контрольная точка G1/S	HDAC4; SMAD3; SUV39H1; HDAC5; CDKN1B; BTRC; ATR; ABL1; E2F1; HDAC2; HDAC7A; RB1; HDAC11; HDAC9; CDK2; E2F2; HDAC3; TP53; CDKN1A; CCND1; E2F4; ATM; RBL2; SMAD4; CDKN2A; MYC; NRG1; GSK3B; RBL1; HDAC6
Сигнализация Т-клеточных рецепторов	RAC1; ELK1; MAPK1; IKBKB; CBL; PIK3CA; FOS; NFKB2; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; RELA; PIK3C2A; BTK; LCK; RAF1; IKBKG; RELB; FYN; MAP2K2; PIK3R1; CHUK; MAP2K1; NFKB1; ITK; BCL10; JUN; VAV3
Сигнализация рецепторов смерти	CRADD; HSPB1; BID; BIRC4; TBK1; IKBKB; FADD; FAS; NFKB2; BCL2; MAP3K14; MAPK8; RIPK1; CASP8; DAXX; TNFRSF10B; RELA; TRAF2; TNF; IKBKG; RELB; CASP9; CHUK; APAF1; NFKB1; CASP2; BIRC2; CASP3; BIRC3
Сигнализация FGF	RAC1; FGFR1; MET; MAPKAPK2; MAPK1; PTPN11; AKT2; PIK3CA; CREB1; PIK3CB; PIK3C3; MAPK8; MAPK3; MAPK13; PTPN6; PIK3C2A; MAPK14; RAF1; AKT1; PIK3R1; STAT3; MAP2K1; FGFR4; CRKL; ATF4; AKT3; PRKCA; HGF
Сигнализация GM-CSF	LYN; ELK1; MAPK1; PTPN11; AKT2; PIK3CA; CAMK2A; STAT5B; PIK3CB; PIK3C3; GNB2L1; BCL2L1; MAPK3; ETS1; KRAS; RUNX1; PIM1; PIK3C2A; RAF1; MAP2K2; AKT1; JAK2; PIK3R1; STAT3; MAP2K1; CCND1; AKT3; STAT1
Сигнализация бокового амиотрофического склероза	BID; IGF1; RAC1; BIRC4; PGF; CAPNS1; CAPN2; PIK3CA; BCL2; PIK3CB; PIK3C3; BCL2L1; CAPN1; PIK3C2A; TP53; CASP9; PIK3R1; RAB5A; CASP1; APAF1; VEGFA; BIRC2; BAX; AKT3; CASP3; BIRC3
Сигнализация JAK/Stat	PTPN1; MAPK1; PTPN11; AKT2; PIK3CA; STAT5B; PIK3CB; PIK3C3; MAPK3; KRAS; SOCS1; STAT5A; PTPN6; PIK3C2A; RAF1; CDKN1A; MAP2K2; JAK1; AKT1; JAK2; PIK3R1; STAT3; MAP2K1; FRAP1; AKT3;

	STAT1
Метаболизм никотината и никотинамида	PRKCE; IRAK1; PRKAA2; EIF2AK2; GRK6; MAPK1; PLK1; AKT2; CDK8; MAPK8; MAPK3; PRKCD; PRKAA1; PBEF1; MAPK9; CDK2; PIM1; DYRK1A; MAP2K2; MAP2K1; PAK3; NT5E; TTK; CSNK1A1; BRAF; SGK
Сигнализация хемокинов	CXCR4; ROCK2; MAPK1; PTK2; FOS; CFL1; GNAQ; CAMK2A; CXCL12; MAPK8; MAPK3; KRAS; MAPK13; RHOA; CCR3; SRC; PPP1CC; MAPK14; NOX1; RAF1; MAP2K2; MAP2K1; JUN; CCL2; PRKCA
Сигнализация IL-2	ELK1; MAPK1; PTPN11; AKT2; PIK3CA; SYK; FOS; STAT5B; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; SOCS1; STAT5A; PIK3C2A; LCK; RAF1; MAP2K2; JAK1; AKT1; PIK3R1; MAP2K1; JUN; AKT3
Долговременная синаптическая депрессия	PRKCE; IGF1; PRKCZ; PRDX6; LYN; MAPK1; GNAS; PRKCI; GNAQ; PPP2R1A; IGF1R; PRKD1; MAPK3; KRAS; GRN; PRKCD; NOS3; NOS2A; PPP2CA; YWHAZ; RAF1; MAP2K2; PPP2R5C; MAP2K1; PRKCA
Сигнализация эстрогеновых рецепторов	TAF4B; EP300; CARM1; PCAF; MAPK1; NCOR2; SMARCA4; MAPK3; NR1P1; KRAS; SRC; NR3C1; HDAC3; PPARGC1A; RBM9; NCOA3; RAF1; CREBBP; MAP2K2; NCOA2; MAP2K1; PRKDC; ESR1; ESR2
Путь убиквитинирования белков	TRAF6; SMURF1; BIRC4; BRCA1; UCHL1; NEDD4; CBL; UBE2I; BTRC; HSPA5; USP7; USP10; FBXW7; USP9X; STUB1; USP22; B2M; BIRC2; PARK2; USP8; USP1; VHL; HSP90AA1; BIRC3
Сигнализация IL-10	TRAF6; CCR1; ELK1; IKBKB; SP1; FOS; NFKB2; MAP3K14; MAPK8; MAPK13; RELA; MAPK14; TNF; IKBKG; RELB; MAP3K7; JAK1; CHUK; STAT3; NFKB1; JUN; IL1R1; IL6
Активация VDR/RXR	PRKCE; EP300; PRKCZ; RXRA; GADD45A; HES1; NCOR2; SP1; PRKC1; CDKN1B; PRKD1; PRKCD; RUNX2; KLF4; YY1; NCOA3; CDKN1A; NCOA2; SPP1; LRP5; CEBPB; FOXO1; PRKCA
Сигнализация TGF- β	EP300; SMAD2; SMURF1; MAPK1; SMAD3; SMAD1; FOS; MAPK8; MAPK3; KRAS; MAPK9; RUNX2; SERPINE1; RAF1; MAP3K7; CREBBP; MAP2K2; MAP2K1; TGFBR1; SMAD4; JUN; SMAD5
Сигнализация Toll-подобных рецепторов	IRAK1; EIF2AK2; MYD88; TRAF6; PPARA; ELK1; IKBKB; FOS; NFKB2; MAP3K14; MAPK8; MAPK13; RELA; TLR4; MAPK14; IKBKG; RELB; MAP3K7; CHUK; NFKB1; TLR2; JUN
Сигнализация MAPK p38	HSPB1; IRAK1; TRAF6; MAPKAPK2; ELK1; FADD; FAS; CREB1; DDIT3; RPS6KA4; DAXX; MAPK13; TRAF2; MAPK14; TNF; MAP3K7; TGFBR1; MYC; ATF4; IL1R1; SRF; STAT1
Сигнализация нейротрофина/TRK	NTRK2; MAPK1; PTPN11; PIK3CA; CREB1; FOS; PIK3CB; PIK3C3; MAPK8; MAPK3; KRAS; PIK3C2A; RAF1; MAP2K2; AKT1; PIK3R1; PDPK1; MAP2K1; CDC42; JUN; ATF4

Активация FXR/RXR	INS; PPARA; FASN; RXRA; AKT2; SDC1; MAPK8; APOB; MAPK10; PPARG; MTTT; MAPK9; PPARGC1A; TNF; CREBBP; AKT1; SREBF1; FGFR4; AKT3; FOXO1
Долговременная синаптическая потенция	PRKCE; RAP1A; EP300; PRKCZ; MAPK1; CREB1; PRKCI; GNAQ; CAMK2A; PRKD1; MAPK3; KRAS; PRKCD; PPP1CC; RAF1; CREBBP; MAP2K2; MAP2K1; ATF4; PRKCA
Кальциевая сигнализация	RAP1A; EP300; HDAC4; MAPK1; HDAC5; CREB1; CAMK2A; MYH9; MAPK3; HDAC2; HDAC7A; HDAC11; HDAC9; HDAC3; CREBBP; CALR; CAMKK2; ATF4; HDAC6
Сигнализация EGF	ELK1; MAPK1; EGFR; PIK3CA; FOS; PIK3CB; PIK3C3; MAPK8; MAPK3; PIK3C2A; RAF1; JAK1; PIK3R1; STAT3; MAP2K1; JUN; PRKCA; SRF; STAT1
Сигнализация гипоксии в сердечно-сосудистой системе	EDN1; PTEN; EP300; NQO1; UBE2I; CREB1; ARNT; HIF1A; SLC2A4; NOS3; TP53; LDHA; AKT1; ATM; VEGFA; JUN; ATF4; VHL; HSP90AA1
Опосредованное LPS/IL-1 ингибирование функции RXR	IRAK1; MYD88; TRAF6; PPARA; RXRA; ABCA1; MAPK8; ALDH1A1; GSTP1; MAPK9; ABCB1; TRAF2; TLR4; TNF; MAP3K7; NR1H2; SREBF1; JUN; IL1R1
Активация LXR/RXR	FASN; RXRA; NCOR2; ABCA1; NFKB2; IRF3; RELA; NOS2A; TLR4; TNF; RELB; LDLR; NR1H2; NFKB1; SREBF1; IL1R1; CCL2; IL6; MMP9
Процессинг амилоида	PRKCE; CSNK1E; MAPK1; CAPNS1; AKT2; CAPN2; CAPN1; MAPK3; MAPK13; MAPT; MAPK14; AKT1; PSEN1; CSNK1A1; GSK3B; AKT3; APP
Сигнализация IL-4	AKT2; PIK3CA; PIK3CB; PIK3C3; IRS1; KRAS; SOCS1; PTPN6; NR3C1; PIK3C2A; JAK1; AKT1; JAK2; PIK3R1; FRAP1; AKT3; RPS6KB1
Регуляция клеточного цикла: контрольная точка повреждения ДНК в G2/M	EP300; PCAF; BRCA1; GADD45A; PLK1; BTRC; CHEK1; ATR; CHEK2; YWHAZ; TP53; CDKN1A; PRKDC; ATM; SFN; CDKN2A
Сигнализация оксида азота в сердечно-сосудистой системе	KDR; FLT1; PGF; AKT2; PIK3CA; PIK3CB; PIK3C3; CAV1; PRKCD; NOS3; PIK3C2A; AKT1; PIK3R1; VEGFA; AKT3; HSP90AA1
Метаболизм пуринов	NME2; SMARCA4; MYH9; RRM2; ADAR; EIF2AK4; PKM2; ENTPD1; RAD51; RRM2B; TJP2; RAD51C; NT5E; POLD1; NME1
Опосредованная cAMP сигнализация	RAP1A; MAPK1; GNAS; CREB1; CAMK2A; MAPK3; SRC; RAF1; MAP2K2; STAT3; MAP2K1; BRAF; ATF4
Митохондриальная дисфункция	SOD2; MAPK8; CASP8; MAPK10; MAPK9; CASP9; PARK7; PSEN1; PARK2; APP; CASP3
Сигнализация Notch	HES1; JAG1; NUMB; NOTCH4; ADAM17; NOTCH2; PSEN1; NOTCH3; NOTCH1; DLL4
Стресс-путь эндоплазматического ретикулума	HSPA5; MAPK8; XBPI; TRAF2; ATF6; CASP9; ATF4; EIF2AK3; CASP3
Метаболизм пиримидинов	NME2; AICDA; RRM2; EIF2AK4; ENTPD1; RRM2B; NT5E; POLD1; NME1
Сигнализация болезни	UCHL1; MAPK8; MAPK13; MAPK14; CASP9; PARK7;

Паркинсона	PARK2; CASP3
Сердечная и β-адренергическая сигнализация	GNAS; GNAQ; PPP2R1A; GNB2L1; PPP2CA; PPP1CC; PPP2R5C
Гликолиз/глюконеогенез	HK2; GCK; GPI; ALDH1A1; PKM2; LDHA; HK1
Сигнализация интерферона	IRF1; SOCS1; JAK1; JAK2; IFITM1; STAT1; IFIT3
Сигнализация Sonic Hedgehog	ARRB2; SMO; GLI2; DYRK1A; GLI1; GSK3B; DYRK1B
Метаболизм глицерофосфолипидов	PLD1; GRN; GPAM; YWHAZ; SPHK1; SPHK2
Деградация фосфолипидов	PRDX6; PLD1; GRN; YWHAZ; SPHK1; SPHK2
Метаболизм триптофана	SIAH2; PRMT5; NEDD4; ALDH1A1; CYP1B1; SIAH1
Деградация лизина	SUV39H1; EHMT2; NSD1; SETD7; PPP2R5C
Путь эксцизионной репарации нуклеотидов	ERCC5; ERCC4; XPA; XPC; ERCC1
Метаболизм крахмала и сахарозы	UCHL1; HK2; GCK; GPI; HK1
Метаболизм аминсахаров	NQO1; HK2; GCK; HK1
Метаболизм арахидоновой кислоты	PRDX6; GRN; YWHAZ; CYP1B1
Сигнализация циркадного ритма	CSNK1E; CREB1; ATF4; NR1D1
Система свертывания крови	BDKRB1; F2R; SERPINE1; F3
Сигнализация дофаминовых рецепторов	PPP2R1A; PPP2CA; PPP1CC; PPP2R5C
Метаболизм глутатиона	IDH2; GSTP1; ANPEP; IDH1
Метаболизм глицеролипидов	ALDH1A1; GPAM; SPHK1; SPHK2
Метаболизм линолеиновой кислоты	PRDX6; GRN; YWHAZ; CYP1B1
Метаболизм метионина	DNMT1; DNMT3B; AHCY; DNMT3A
Метаболизм пирувата	GLO1; ALDH1A1; PKM2; LDHA
Метаболизм аргинина и пролина	ALDH1A1; NOS3; NOS2A
Эйкозаноидная сигнализация	PRDX6; GRN; YWHAZ
Метаболизм фруктозы и маннозы	HK2; GCK; HK1
Метаболизм галактозы	HK2; GCK; HK1
Биосинтез стильбена, кумарина и лигнина	PRDX6; PRDX1; TYR
Путь презентации антигенов	CALR; B2M
Биосинтез стероидов	NQO1; DHCR7
Метаболизм бутаноата	ALDH1A1; NLGN1
Цикл лимонной кислоты	IDH2; IDH1
Метаболизм жирных кислот	ALDH1A1; CYP1B1
Метаболизм глицерофосфолипидов	PRDX6; CHKA
Метаболизм гистидина	PRMT5; ALDH1A1
Метаболизм инозитола	EROIL; APEX1
Метаболизм ксенобиотиков у цитохрома p450	GSTP1; CYP1B1
Метаболизм метана	PRDX6; PRDX1

Метаболизм фенилаланина	PRDX6; PRDX1
Метаболизм пропаноата	ALDH1A1; LDHA
Метаболизм селеноаминокислот	PRMT5; AHCY
Метаболизм сфинголипидов	SPHK1; SPHK2
Метаболизм аминокислот	PRMT5
Метаболизм андрогенов и эстрогенов	PRMT5
Метаболизм аскорбата и альдаратов	ALDH1A1
Биосинтез желчных кислот	ALDH1A1
Метаболизм цистеина	LDHA
Биосинтез жирных кислот	FASN
Сигнализация глутаматных рецепторов	GNB2L1
Опосредованная NRF2 реакция на окислительный стресс	PRDX1
Пентозофосфатный путь	GPI
Взаимопревращения пентоз и глюкоуроната	UCHL1
Метаболизм ретинола	ALDH1A1
Метаболизм рибофлавина	TYR
Метаболизм тирозина	PRMT5, TYR
Биосинтез убихинона	PRMT5
Деградация валина, лейцина и изолейцина	ALDH1A1
Метаболизм глицина, серина и треонина	CHKA
Деградация лизина	ALDH1A1
Боль/вкус	TRPM5; TRPA1
Боль	TRPM7; TRPC5; TRPC6; TRPC1; Cnr1; cnr2; Grk2; Trpa1; Pomc; Cgrp; Crf; Pka; Era; Nr2b; TRPM5; Prkaca; Prkacb; Prkar1a; Prkar2a
Митохондриальная функция	AIF; CytC; SMAC (Diablo); Aifm-1; Aifm-2
Возрастная неврология	BMP-4; хордин (Chrd); ноггин (Nog); WNT (Wnt2; Wnt2b; Wnt3a; Wnt4; Wnt5a; Wnt6; Wnt7b; Wnt8b; Wnt9a; Wnt9b; Wnt10a; Wnt10b; Wnt16); β -катенин; Dkk-1; белки типа Frizzled; Otx-2; Gbx2; FGF-8; рилин; Dab1; unc-86 (Pou4fl или Brn3a); Numb; Reln

Последовательность целевого полинуклеотида может содержать целевую нуклеиновую кислоту или последовательность протоспейсера (то есть последовательность, распознаваемую спейсерным участком направляющей нуклеиновой кислоты) длиной примерно в 20 нуклеотидов. Длина протоспейсера может составлять менее 20 нуклеотидов. Длина протоспейсера может составлять по меньшей мере 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 и более нуклеотидов. Длина последовательности протоспейсера может составлять не более 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 и более нуклеотидов. Последовательность протоспейсера может отстоять на 16, 17, 18, 19, 20, 21, 22 или 23 основания непосредственно 5' от первого нуклеотида РАМ. Последовательность протоспейсера может отстоять на 16, 17, 18, 19, 20, 21, 22 или 23 основания непосредственно 3' от последнего нуклеотида последовательности РАМ. Последовательность протоспейсера может отстоять на 20 оснований непосредственно 5' от первого нуклеотида последовательности РАМ. Последовательность протоспейсера может отстоять на 20 оснований непосредственно 3' от последнего нуклеотида последовательности РАМ. Последовательность целевой нуклеиновой кислоты может находиться с 5'- или 3'-стороны от РАМ.

Последовательность протоспейсера может включать в себя последовательность нуклеиновой кислоты, присутствующей в целевом полинуклеотиде, с которой может связываться нацеливающий на нуклеиновую кислоту сегмент направляющей нуклеиновой кислоты. К примеру, последовательность протоспейсера может включать в себя последовательность, с которой направляющая нуклеиновая кислота должна быть комплементарной. Последовательность протоспейсера может содержать любой полинуклеотид, который может находиться, к примеру, в ядре или цитоплазме клетки или внутри органеллы клетки типа митохондрий или хлоропластов. Последовательность протоспейсера может содержать сайты расщепления для белков Cas. Последовательность протоспейсера может примыкать к сайтам расщепления для белков Cas.

Белок Cas может связывать целевой полинуклеотид на сайте в пределах или за пределами последовательности, с которой может связываться нацеливающая на нуклеиновую кислоту последовательность направляющей нуклеиновой кислоты. Участок связывания может включать положение той нуклеиновой

кислоты, у которой белок Cas может образовывать одностранный разрыв или двустранный разрыв.

Сайт-специфичное связывание целевой нуклеиновой кислоты белком Cas может происходить в местах, которые определяются комплементарностью спаривания оснований между направляющей нуклеиновой кислотой и целевой нуклеиновой кислотой. Сайт-специфичное связывание целевой нуклеиновой кислоты белком Cas может происходить в местах, определяемых коротким мотивом, который называют смежным с протоспейсером мотивом (PAM), в целевой нуклеиновой кислоте. PAM может фланкировать протоспейсер, к примеру, на 3'-конце последовательности протоспейсера. Например, сайт связывания Cas9 может быть от 1 до 25 или от 2 до 5 или от 19 до 23 пар оснований (например, 3 пар оснований) вверх или вниз от последовательности PAM. Сайт связывания Cas (например, Cas9) может быть на 3 пары оснований выше последовательности PAM. Сайт связывания Cas (например, Cpf1) может составлять 19 оснований на (+)-нити и 23 основания на (-)-нити.

Различные организмы могут содержать разные последовательности PAM. Различные Cas-белки могут распознавать разные последовательности PAM. Например, PAM у *S.pyogenes* может содержать последовательность 5'-XRR-3', где R может означать A или G, а X - любой нуклеотид, причем X находится непосредственно 3' от последовательности целевой нуклеиновой кислоты, на которую нацелена последовательность спейсера. Последовательность PAM у Cas9 *S.pyogenes* (SpyCas9) может быть 5'-XGG-3', где X означает любой ДНК-нуклеотид, который находится непосредственно 3' от последовательности протоспейсера некомплементарной нити целевой ДНК. PAM Cpf1 может быть 5'-TTX-3', где X означает любой ДНК-нуклеотид, который находится непосредственно 3' от последовательности распознавания CRISPR.

Целевая последовательность для направляющей нуклеиновой кислоты может быть идентифицирована методами биоинформатики, к примеру, путем обнаружения последовательностей в целевой последовательности, смежных с последовательностью PAM. Оптимальная целевая последовательность для направляющей нуклеиновой кислоты может быть идентифицирована экспериментальными методами, к примеру, путем тестирования ряда последовательностей гидовых нуклеиновых кислот для идентификации последовательности с наибольшей прицельной активностью и наименьшей неприцельной активностью. Местонахождение целевой последовательности может определяться требуемым экспериментальным результатом. Например, целевой протоспейсер может располагаться в промоторе с тем, чтобы активировать или репрессировать целевой ген. Целевой протоспейсер может находиться в кодирующей последовательности типа 5'-концевого конститутивно экспрессируемого экзона или последовательности, кодирующей известный домен. Целевой протоспейсер может представлять собой уникальную последовательность в геноме с тем, чтобы уменьшить неприцельные эффекты. Для определения и ранжирования потенциальных целевых протоспейсеров в данной области известны и могут применяться многие общедоступные алгоритмы.

В некоторых аспектах раскрытые здесь системы могут регулировать экспрессию по меньшей мере одного гена, связанного с генетическим заболеванием или медицинским состоянием. Широкий спектр генетических заболеваний подробно описан на веб-сайте National Institutes of Health в тематическом подпункте Genetic Disorders (веб-сайт по адресу health.nih.gov/topic/GeneticDisorders).

Как должно быть ясно, предполагается, что настоящая система может применяться для наведения на последовательность любого представляющего интерес полинуклеотида. Некоторые примеры заболеваний, которые можно эффективно лечить с помощью настоящей системы, представлены в табл. 3-5, а также там приведены примеры генов, которые в настоящее время ассоциируются с этими заболеваниями. Однако примеры таких генов не являются исчерпывающими.

Целевая нуклеиновая кислота может содержать одну или несколько последовательностей, которые по меньшей мере частично комплементарны одной или нескольким гидовым нуклеиновым кислотам. Целевая нуклеиновая кислота может представлять собой часть гена или целый ген, 5'-конец гена, 3'-конец гена, регуляторный элемент (например, промотор, энхансер), псевдоген, некодирующую ДНК, микросателлит, интрон, экзон, хромосомную ДНК, митохондриальную ДНК, смысловую ДНК, антисмысловую ДНК, ДНК нуклеотидов, ДНК хлоропластов или РНК в числе других нуклеиновых кислот. Целевая нуклеиновая кислота может представлять собой часть или всю ДНК плазмиды. ДНК плазмиды или ее часть может быть отрицательно суперспиральной. Целевая нуклеиновая кислота может быть *in vitro* или *in vivo*.

Целевая нуклеиновая кислота может содержать участок последовательности с низким содержанием GC. Целевая нуклеиновая кислота может быть отрицательно суперспиральной. В качестве неограниченного примера, целевая нуклеиновая кислота может иметь содержание GC не менее 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60 или 65% и более. Целевая нуклеиновая кислота может иметь содержание GC не более 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60 или 65% и более.

Участок с определенным содержанием GC может составлять отрезок целевой нуклеиновой кислоты, который гибридизуется с направляющей нуклеиновой кислотой. Участок с таким содержанием GC может быть длиннее или короче, чем тот отрезок, который гибридизуется с направляющей нуклеиновой кислотой. Участок с таким содержанием GC может быть по меньшей мере на 30, 40, 50, 60, 70, 80, 90 или 100 и более нуклеотидов длиннее или короче, чем тот отрезок, который гибридизуется с направляющей

нуклеиновой кислотой. Участок с таким содержанием GC может быть не более чем на 30, 40, 50, 60, 70, 80, 90 или 100 и более нуклеотидов длиннее или короче, чем тот отрезок, который гибридизуется с направляющей нуклеиновой кислотой.

В одном аспекте настоящего изобретения предусмотрен способ регулирования экспрессии целевого полинуклеотида в клетках, содержащих ядро. В некоторых воплощениях способ включает (а) экспонирование химерного внутриклеточного рецептора с антигеном, причем (i) рецептор содержит взаимодействующий с антигеном домен и исполнительный элемент, и (ii) рецептор подвергается модификации при экспозиции с антигеном; (b) транслокацию модифицированного рецептора в ядро; и (c) образование комплекса между исполнительным элементом и целевым полинуклеотидом. Химерный внутриклеточный рецептор, как уже описано здесь, может содержать ядерный рецептор либо его производное, вариант или фрагмент. В некоторых воплощениях химерный внутриклеточный рецептор связывает гормон.

После экспозиции с антигеном химерный внутриклеточный рецептор может подвергаться модификации. После модификации рецептора химерный внутриклеточный рецептор может транслоцироваться в ядро клетки. В ядре исполнительный элемент может образовывать комплекс с целевым полинуклеотидом.

Исполнительный элемент, как описано выше, может содержать нуклеазу (например, ДНК-нуклеазу и/или РНК-нуклеазу), модифицированную нуклеазу (например, ДНК-нуклеазу и/или РНК-нуклеазу), которая дефектна по нуклеазе или обладает меньшей нуклеазной активностью по сравнению с нуклеазой дикого типа, её вариант, производное или фрагмент, как уже описано здесь. Исполнительный элемент может регулировать экспрессию и/или активность гена или редактировать последовательность нуклеиновой кислоты (например, гена и/или продукта гена). В некоторых воплощениях исполнительный элемент содержит ДНК-нуклеазу типа генно-инженерной (например, программируемой или наводимой) ДНК-нуклеазы, чтобы запускать геномное редактирование целевой последовательности ДНК. В некоторых воплощениях исполнительный элемент содержит РНК-нуклеазу типа генно-инженерной (например, программируемой или наводимой) РНК-нуклеазы, чтобы запускать редактирование целевой последовательности РНК. В некоторых воплощениях исполнительный элемент обладает пониженной или минимальной нуклеазной активностью. Исполнительный элемент с пониженной или минимальной нуклеазной активностью может регулировать экспрессию и/или активность гена путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида. В некоторых воплощениях исполнительный элемент содержит безнуклеазный ДНК-связывающий белок, происходящий из ДНК-нуклеазы, который может вызывать активацию или репрессию транскрипции целевой последовательности ДНК. В некоторых воплощениях исполнительный элемент содержит безнуклеазный белок, связывающий РНК, происходящий из РНК-нуклеазы, который может вызывать активацию или репрессию транскрипции целевой последовательности РНК. Исполнительный элемент может регулировать экспрессию или активность гена и/или редактировать последовательность нуклеиновой кислоты, будь то экзогенной или эндогенной.

В исполнительных элементах может использоваться любая подходящая нуклеаза. К подходящим нуклеазам относятся, без ограничения, CRISPR-ассоциированные (Cas) белки или Cas-нуклеазы, в том числе CRISPR-ассоциированные (Cas) полипептиды I типа, CRISPR-ассоциированные (Cas) полипептиды II типа, CRISPR-ассоциированные (Cas) полипептиды III типа, CRISPR-ассоциированные (Cas) полипептиды IV типа, CRISPR-ассоциированные (Cas) полипептиды V типа и CRISPR-ассоциированные (Cas) полипептиды VI типа; нуклеазы с "цинковым пальцем" (ZFN); эффекторные нуклеазы типа активаторов транскрипции (TALEN); мегануклеазы; РНК-связывающие белки (RBP); CRISPR-ассоциированные РНК-связывающие белки; рекомбиназы; флиппазы; транспозазы; белки Argonaute. В некоторых воплощениях целевой полинуклеотид содержит геномную ДНК. В некоторых воплощениях целевой полинуклеотид содержит участок плазмиды, к примеру, плазмиды, несущей экзогенный ген. В некоторых воплощениях целевой полинуклеотид содержит РНК, например, мРНК. В некоторых воплощениях целевой полинуклеотид содержит эндогенный ген или генный продукт, его производное, вариант или фрагмент.

В некоторых воплощениях исполнительный элемент содержит белок Cas, который образует комплекс с направляющей нуклеиновой кислотой типа направляющей РНК. В некоторых воплощениях исполнительный элемент содержит белок, связывающий РНК (RBP), необязательно в комплексе с направляющей нуклеиновой кислотой типа направляющей РНК, который способен образовывать комплекс с белком Cas. В некоторых воплощениях исполнительный элемент содержит белок Cas, лишенный активности расщепления.

В некоторых аспектах настоящего изобретения предусмотрен способ избирательного модулирования транскрипции целевой нуклеиновой кислоты в клетках хозяина. Способ может включать а) введение в клетки хозяина: i) исполнительного элемента, содержащего химерный белок Cas (например, белок Cas, слитый с рецепторным или адаптерным белком), либо нуклеиновой кислоты, содержащей последовательность нуклеотидов, кодирующую химерный белок Cas, причем белок Cas ферментативно неактивен (например, это мертвый Cas, dCas9) или проявляет пониженную (например, эндодезоксирибонуклеазную, эндорибонуклеазную) активность; и ii) направляющей нуклеиновой кислоты или нуклеиновой кислоты, содержащей последовательность нуклеотидов, кодирующую гидовую нуклеиновую кислоту. На-

правляющая нуклеиновая кислота может содержать: i) первый сегмент (например, спейсерный участок, участок наведения на нуклеиновую кислоту), содержащий последовательность нуклеотидов, которая комплементарна целевой последовательности в целевой нуклеиновой кислоте (например, геномной ДНК, мРНК); и ii) второй сегмент (например, белок-связывающий сегмент или сегмент, связывающий белок Cas), который взаимодействует с белком Cas. Белок Cas может содержать: i) участок связывания направляющей нуклеиновой кислоты, который взаимодействует с направляющей нуклеиновой кислотой; и ii) участок, который не проявляет или проявляет слабую активность нуклеазы. Направляющая нуклеиновая кислота и мертвый белок Cas могут образовывать комплекс в клетках хозяина. Комплекс может избирательно модулировать транскрипцию целевой ДНК в клетках хозяина. Целевым полинуклеотидом может быть любой целевой полинуклеотид, описанный здесь, к примеру, любой целевой полинуклеотид, связанный с описанным здесь геном.

В различных воплощениях изложенных здесь аспектов рассматриваемые системы могут применяться для избирательной модуляции (например, снижения или повышения) транскрипции целевой нуклеиновой кислоты в клетках хозяина). Избирательная модуляция транскрипции целевой нуклеиновой кислоты может уменьшать или повышать транскрипцию целевой нуклеиновой кислоты, но не может существенно модулировать транскрипцию нецелевой нуклеиновой кислоты или неприцельной нуклеиновой кислоты, например, транскрипция нецелевой нуклеиновой кислоты может модулироваться на менее 1%, менее 5%, менее 10%, менее 20%, менее 30%, менее 40% или менее 50% по сравнению с уровнем транскрипции нецелевой нуклеиновой кислоты в отсутствие исполнительного элемента типа комплекса направляющей нуклеиновой кислоты с ферментативно неактивным или ферментативно ослабленным белком Cas. К примеру, избирательная модуляция (например, снижение или повышение) транскрипции целевой нуклеиновой кислоты может понижать или повышать транскрипцию целевой нуклеиновой кислоты по меньшей мере на 10%, по меньшей мере на 20%, по меньшей мере на 30%, по меньшей мере на 40%, по меньшей мере на 50%, по меньшей мере на 60%, по меньшей мере на 70%, по меньшей мере на 80%, по меньшей мере на 90% или более 90% по сравнению с уровнем транскрипции целевой нуклеиновой кислоты в отсутствие исполнительного элемента типа комплекса направляющей нуклеиновой кислоты с ферментативно неактивным или ферментативно ослабленным белком Cas.

В некоторых воплощениях изобретения предусмотрены способы повышения транскрипции целевой нуклеиновой кислоты. Транскрипция целевой нуклеиновой кислоты может повышаться, по меньшей мере, в 1,1 раз, по меньшей мере, в 1,2 раза, по меньшей мере, в 1,3 раза, по меньшей мере, в 1,4 раза, по меньшей мере, в 1,5 раз, по меньшей мере, в 1,6 раз, по меньшей мере, в 1,7 раз, по меньшей мере, в 1,8 раз, по меньшей мере, в 1,9 раз, по меньшей мере, в 2 раза, по меньшей мере, в 2,5 раза, по меньшей мере, в 3 раза, по меньшей мере, в 3,5 раза, по меньшей мере, в 4 раза, по меньшей мере, в 4,5 раза, по меньшей мере, в 5 раз, по меньшей мере, в 6 раз, по меньшей мере, в 7 раз, по меньшей мере, в 8 раз, по меньшей мере, в 9 раз, по меньшей мере, в 10 раз, по меньшей мере, в 12 раз, по меньшей мере, в 15 раз, по меньшей мере, в 20 раз, по меньшей мере, в 50 раз, по меньшей мере, в 70 раз или, по меньшей мере, в 100 раз по сравнению с уровнем транскрипции целевой ДНК в отсутствие исполнительного элемента типа комплекса направляющей нуклеиновой кислоты с ферментативно неактивным или ферментативно ослабленным белком Cas. При избирательном повышении транскрипции целевой нуклеиновой кислоты повышается транскрипция целевой нуклеиновой кислоты, но не может существенно повышаться транскрипция нецелевой ДНК. Например, транскрипция нецелевой нуклеиновой кислоты повышается, если вообще, менее чем в 5 раз, менее чем в 4 раза, менее чем в 3 раза, менее чем в 2 раза, менее чем в 1,8 раза, менее чем в 1,6 раза, менее чем в 1,4 раза, менее чем в 1,2 раза или менее чем в 1,1 раз по сравнению с уровнем транскрипции нецелевой ДНК в отсутствие исполнительного элемента типа комплекса направляющей нуклеиновой кислоты с ферментативно неактивным или ферментативно ослабленным белком Cas.

В некоторых воплощениях изобретения предусмотрены способы снижения транскрипции целевой нуклеиновой кислоты. Транскрипция целевой нуклеиновой кислоты может снижаться, по меньшей мере, в 1,1 раз, по меньшей мере, в 1,2 раза, по меньшей мере, в 1,3 раза, по меньшей мере, в 1,4 раза, по меньшей мере, в 1,5 раз, по меньшей мере, в 1,6 раз, по меньшей мере, в 1,7 раз, по меньшей мере, в 1,8 раз, по меньшей мере, в 1,9 раз, по меньшей мере, в 2 раза, по меньшей мере, в 2,5 раза, по меньшей мере, в 3 раза, по меньшей мере, в 3,5 раза, по меньшей мере, в 4 раза, по меньшей мере, в 4,5 раза, по меньшей мере, в 5 раз, по меньшей мере, в 6 раз, по меньшей мере, в 7 раз, по меньшей мере, в 8 раз, по меньшей мере, в 9 раз, по меньшей мере, в 10 раз, по меньшей мере, в 12 раз, по меньшей мере, в 15 раз, по меньшей мере, в 20 раз, по меньшей мере, в 50 раз, по меньшей мере, в 70 раз или, по меньшей мере, в 100 раз по сравнению с уровнем транскрипции целевой ДНК в отсутствие исполнительного элемента типа комплекса направляющей нуклеиновой кислоты с ферментативно неактивным или ферментативно ослабленным белком Cas. При избирательном снижении транскрипции целевой нуклеиновой кислоты снижается транскрипция целевой нуклеиновой кислоты, но не может существенно снижаться транскрипция нецелевой ДНК, например, транскрипция нецелевой нуклеиновой кислоты снижается, если вообще, менее чем в 5 раз, менее чем в 4 раза, менее чем в 3 раза, менее чем в 2 раза, менее чем в 1,8 раза, менее чем в 1,6 раза, менее чем в 1,4 раза, менее чем в 1,2 раза или менее чем в 1,1 раз по сравнению с уровнем транскрип-

ции нецелевой ДНК в отсутствие исполнительного элемента типа комплекса направляющей нуклеиновой кислоты с ферментативно неактивным или ферментативно ослабленным белком Cas.

Модуляция транскрипции может осуществляться путем слияния исполнительного элемента типа ферментативно неактивного белка Cas с гетерологичной последовательностью. Гетерологичной последовательностью может быть подходящий партнер по слиянию, например, полипептид, который обеспечивает активность, которая косвенным образом повышает, снижает или иным образом модулирует транскрипцию, действуя непосредственно на целевую нуклеиновую кислоту или на полипептид (например, гистон или другой ДНК-связывающий белок), связанный с целевой нуклеиновой кислотой. Неограничительные примеры подходящих партнеров по слиянию включают полипептиды, которые обеспечивают метилтрансферазную активность, деметилазную активность, ацетилтрансферазную активность, деацетилазную активность, киназную активность, фосфатазную активность, убиквитин-лигазную активность, активность деубиквитинирования, активность аденилирования, активность деаденилирования, активность SUMOилирования, активность деSUMOилирования, активность рибозилирования, активность де-рибозилирования, активность миристоилирования или активность демиристоилирования.

Подходящий партнер по слиянию может включать в себя полипептид, который непосредственно обеспечивает повышение транскрипции целевой нуклеиновой кислоты. Например, активатор транскрипции или его фрагмент, белок или его фрагмент, который рекрутирует активатор транскрипции, или реагирующий на небольшие молекулы или лекарства регулятор транскрипции. Подходящий партнер по слиянию может включать в себя полипептид, который непосредственно обеспечивает снижение транскрипции целевой нуклеиновой кислоты. Например, репрессор транскрипции или его фрагмент, белок или его фрагмент, который рекрутирует репрессор транскрипции, или реагирующий на небольшие молекулы или лекарства регулятор транскрипции.

Гетерологичная последовательность или партнер по слиянию может быть слит с С-концом, N-концом или внутренней частью (то есть другой частью, чем N- или С-конец) исполнительного элемента, к примеру, мертвого белка Cas. Неограничительные примеры партнеров по слиянию включают активаторы транскрипции, репрессоры транскрипции, гистон-лизинметилтрансферазы (KMT), гистон-лизиндеметилазы, гистон-лизиацетилтрансферазы (KAT), гистон-лизиндеацетилазы, ДНК-метилазы (модифицируют аденозин или цитозин), CTCF, периферийные рекрутинг-элементы (например, ламин А, ламин В) и докинг-элементы белков (например, FKBP/FRB).

Неограничивающие примеры активаторов транскрипции включают GAL4, VP16, VP64 и субдомен p65 (NF-κB).

Неограничивающие примеры репрессоров транскрипции включают Krüppel-associated box (KRAB или SKD), домен взаимодействия Mad-mSIN3 (SID) и репрессорный домен ERF (ERD).

Неограничивающие примеры гистон-лизинметилтрансфераз (KMT) включают представителей семейства KMT1 (например, SUV39H1, SUV39H2, G9A, ESET/SETDB1, Ctr4, Su(var)3-9), представителей семейства KMT2 (например, hSETA, hSET1B, MLL1-MLL5, ASH1 и гомологи (Trx, Trt, Ash1)), семейства KMT3 (SYMD2, NSD1), KMT4 (DOT1L и гомологи), семейства KMT5 (Pr-SET7/8, SUV4-20H1 и гомологи), KMT6 (EZH2) и KMT8 (например, RIZ1).

Неограничивающие примеры гистон-лизиндеметилаз (KDM) включают представителей семейства KDM1 (LSD1/BHC110, Spsd1/Swm1/Saf110, Su(var)3-3), семейства KDM3 (JHDM2a/b), семейства KDM4 (JMJD2A/JHDM3A, JMJD2B, JMJD2C/GASC1, JMJD2D и гомологи (Rph1)), семейства KDM5 (JARID1A/RBP2, JARID1B/PLU-1, JARID1C/SMCX, JARID1D/SMCY и гомологи (Lid, Jhn2, Jmj2)) и семейства KDM6 (например, UTX, JMJD3).

Неограничивающие примеры KAT включают представителей семейства KAT2 (hGCN5, PCAF и гомологи (dGCN5/PCAF, Gcn5), семейства KMT3 (CBP, p300 и гомологи (dCBP/NEJ)), KAT4, KAT5, KAT6, KAT7, KAT8 и KAT13.

В некоторых воплощениях исполнительный элемент, содержащий мертвый белок Cas или слитый мертвый белок Cas, направляется направляющей нуклеиновой кислотой в определенное место (т.е. последовательность) в целевой нуклеиновой кислоте и осуществляет locus-специфичную регуляцию типа блокирования связывания РНК-полимеразы с промотором (например, это может избирательно ингибировать функцию активатора транскрипции) и/или модификации местного статуса хроматина (например, при использовании слитой последовательности, которая может модифицировать целевую нуклеиновую кислоту или модифицировать полипептид, связанный с целевой нуклеиновой кислотой). В некоторых случаях эти изменения являются кратковременными (например, репрессия или активация транскрипции). В некоторых случаях эти изменения наследуются (например, при проведении эпигенетических модификаций в целевой ДНК или в белках, связанных с целевой ДНК, например, нуклеосомных гистонах).

В некоторых воплощениях направляющая нуклеиновая кислота может содержать белок-связывающий сегмент для рекрутинга гетерологичного полипептида к целевой нуклеиновой кислоте для модуляции транскрипции целевой нуклеиновой кислоты. Неограничивающие примеры гетерологичных полипептидов включают полипептиды, которые обеспечивают метилтрансферазную активность, деметилазную активность, ацетилтрансферазную активность, деацетилазную активность, киназную активность, фосфатазную активность, убиквитин-лигазную активность, активность деубиквитинирования, актив-

ность аденилирования, активность деаденилирования, активность SUMOилирования, активность деSUMOилирования, активность рибозилирования, активность дерибозилирования, активность миристоилирования или активность демиростоилирования. Направляющая нуклеиновая кислота может содержать белок-связывающий сегмент для рекрутинга активатора транскрипции, репрессора транскрипции либо их фрагментов.

В некоторых воплощениях модуляция экспрессии гена достигается при помощи направляющей нуклеиновой кислоты, служащей для наводки на регуляторный элемент целевой нуклеиновой кислоты, к примеру, транскрипционный элемент отклика (например, промоторы, энхансеры), вышележащие активирующие последовательности (UAS) и/или последовательности с неизвестной или известной функцией, которые предположительно способны контролировать экспрессию целевой ДНК.

В различных воплощениях изложенных здесь аспектов изобретения предусмотрена направляющая нуклеиновая кислота для применения в системе CRISPR/Cas. Направляющая нуклеиновая кислота (например, направляющая РНК) может связываться с белком Cas и направлять белок Cas в определенное место в пределах целевого полинуклеотида. Направляющая нуклеиновая кислота может содержать сегмент, нацеливающий на нуклеиновую кислоту, и сегмент, связывающий белок Cas.

Направляющая нуклеиновая кислота может означать такую нуклеиновую кислоту, которая может гибридизироваться с другой нуклеиновой кислотой, например, целевым полинуклеотидом в геноме клетки. Направляющей нуклеиновой кислотой может быть РНК, к примеру, направляющая РНК. Направляющей нуклеиновой кислотой может быть ДНК. Направляющая нуклеиновая кислота может содержать ДНК и РНК. Направляющая нуклеиновая кислота может быть одноцепочечной. Направляющая нуклеиновая кислота может быть двухцепочечной. Направляющая нуклеиновая кислота может содержать аналоги нуклеотидов. Направляющая нуклеиновая кислота может содержать модифицированные нуклеотиды. Направляющая нуклеиновая кислота может быть запрограммирована или составлена для сайт-специфичного связывания с последовательностью нуклеиновой кислоты.

Направляющая нуклеиновая кислота может содержать одну или несколько модификаций для придания нуклеиновой кислоте новых или улучшенных характеристик. Направляющая нуклеиновая кислота может содержать аффинную метку нуклеиновой кислоты. Направляющая нуклеиновая кислота может содержать синтетические нуклеотиды, синтетические аналоги нуклеотидов, производные нуклеотидов и/или модифицированные нуклеотиды.

Направляющая нуклеиновая кислота может содержать нацеливающий на нуклеиновую кислоту участок (например, спейсерный участок), к примеру, на или возле 5'-конца или 3'-конца, который комплементарен последовательности протоспейсера в целевом полинуклеотиде. Спейсер направляющей нуклеиновой кислоты может взаимодействовать с протоспейсером специфичным для последовательности образом посредством гибридизации (то есть спаривания оснований). Последовательность протоспейсера может располагаться с 5'- или 3'-стороны от смежного с протоспейсером мотива (PAM) в целевом полинуклеотиде. Нуклеотидная последовательность спейсерного участка может варьироваться и определяет то место в целевой нуклеиновой кислоте, с которым может взаимодействовать направляющая нуклеиновая кислота. Спейсерный участок направляющей нуклеиновой кислоты можно составить или модифицировать так, чтобы он гибридизировался с любой желательной последовательностью в целевой нуклеиновой кислоте.

Направляющая нуклеиновая кислота может содержать две отдельные молекулы нуклеиновой кислоты, что можно назвать двойной направляющей нуклеиновой кислотой. Направляющая нуклеиновая кислота может содержать одну молекулу нуклеиновой кислоты, что можно назвать одинарной направляющей нуклеиновой кислотой (например, sgRNA). В некоторых воплощениях направляющая нуклеиновая кислота является одинарной направляющей нуклеиновой кислотой, содержащей слитую РНК CRISPR (crRNA) и трансаактивирующую crРНК (tracrRNA). В некоторых воплощениях Направляющая нуклеиновая кислота представлена одинарной направляющей нуклеиновой кислотой, содержащей crРНК. В некоторых воплощениях направляющая нуклеиновая кислота представлена одинарной направляющей нуклеиновой кислотой, содержащей crРНК, но лишенной tracrРНК. В некоторых воплощениях направляющая нуклеиновая кислота представлена двойной направляющей нуклеиновой кислотой, содержащей неслитые crРНК и tracrРНК. Типичная двойная направляющая нуклеиновая кислота может содержать молекулу типа crРНК и молекулу типа tracrРНК. Типичная одинарная направляющая нуклеиновая кислота может содержать молекулу типа crРНК. Типичная одинарная направляющая нуклеиновая кислота может содержать слитые молекулы типа crРНК и типа tracrРНК.

crРНК может содержать нацеливающий на нуклеиновую кислоту сегмент (например, спейсерный участок) направляющей нуклеиновой кислоты и цепочку нуклеотидов, которая может составлять одну половину двухцепочечного дуплекса связывающего белок Cas сегмента направляющей нуклеиновой кислоты.

tracrРНК может содержать цепочку нуклеотидов, которая составляет вторую половину двухцепочечного дуплекса связывающего белок Cas сегмента gРНК. Цепочка нуклеотидов crРНК может быть комплементарна и может гибридизироваться с цепочкой нуклеотидов tracrРНК с образованием двухцепочечного дуплекса связывающего белок Cas домена направляющей нуклеиновой кислоты.

сгРНК и тасгРНК могут гибридизироваться с образованием направляющей нуклеиновой кислоты. сгRNA также может обеспечивать одноцепочечный нацеливающий на нуклеиновую кислоту сегмент (например, спейсерный участок), который гибридизуется с распознавательной последовательностью целевой нуклеиновой кислоты (например, протоспейсера). Последовательность сгRNA, включающую спейсерный участок, или молекулу тасгRNA можно составить так, чтобы она была специфичной для того вида, у которого будет использоваться направляющая нуклеиновая кислота.

В некоторых воплощениях нацеливающий на нуклеиновую кислоту участок направляющей нуклеиновой кислоты может составлять от 18 до 72 нуклеотидов в длину. Нацеливающий на нуклеиновую кислоту участок направляющей нуклеиновой кислоты (например, спейсерный участок) может иметь в длину от 12 нуклеотидов до 100 нуклеотидов. Например, нацеливающий на нуклеиновую кислоту участок направляющей нуклеиновой кислоты (например, спейсерный участок) может иметь в длину от 12 нуклеотидов (нт) до 80 нт, от 12 нт до 50 нт, от 12 нт до 40 нт, от 12 нт до 30 нт, от 12 нт до 25 нт, от 12 нт до 20 нт, от 12 нт до 19 нт, от 12 нт до 18 нт, от 12 нт до 17 нт, от 12 нт до 16 нт или от 12 нт до 15 нт. С другой стороны, нацеливающий на ДНК участок может иметь в длину от 18 нт до 20 нт, от 18 нт до 25 нт, от 18 нт до 30 нт, от 18 нт до 35 нт, от 18 нт до 40 нт, от 18 нт до 45 нт, от 18 нт до 50 нт, от 18 нт до 60 нт, от 18 нт до 70 нт, от 18 нт до 80 нт, от 18 нт до 90 нт, от 18 нт до 100 нт, от 20 нт до 25 нт, от 20 нт до 30 нт, от 20 нт до 35 нт, от 20 нт до 40 нт, от 20 нт до 45 нт, от 20 нт до 50 нт, от 20 нт до 60 нт, от 20 нт до 70 нт, от 20 нт до 80 нт, от 20 нт до 90 нт или от 20 нт до 100 нт. Длина наводящего на нуклеиновую кислоту участка может составлять по меньшей мере 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 и более нуклеотидов. Длина наводящего на нуклеиновую кислоту участка (например, спейсерной последовательности) может составлять не более 5, 10, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30 и более нуклеотидов.

В некоторых воплощениях нацеливающий на нуклеиновую кислоту участок направляющей нуклеиновой кислоты (например, спейсер) составляет 20 нуклеотидов в длину. В некоторых воплощениях нацеливающий на нуклеиновую кислоту участок направляющей нуклеиновой кислоты составляет 19 нуклеотидов в длину. В некоторых воплощениях нацеливающий на нуклеиновую кислоту участок направляющей нуклеиновой кислоты составляет 18 нуклеотидов в длину. В некоторых воплощениях нацеливающий на нуклеиновую кислоту участок направляющей нуклеиновой кислоты составляет 17 нуклеотидов в длину. В некоторых воплощениях нацеливающий на нуклеиновую кислоту участок направляющей нуклеиновой кислоты составляет 16 нуклеотидов в длину. В некоторых воплощениях нацеливающий на нуклеиновую кислоту участок направляющей нуклеиновой кислоты составляет 21 нуклеотид в длину. В некоторых воплощениях нацеливающий на нуклеиновую кислоту участок направляющей нуклеиновой кислоты составляет 22 нуклеотида в длину.

Нуклеотидная последовательность направляющей нуклеиновой кислоты, которая комплементарна нуклеотидной последовательности (целевой последовательности) целевой нуклеиновой кислоты, может иметь в длину, к примеру, по меньшей мере 12 нт, по меньшей мере 15 нт, по меньшей мере 18 нт, по меньшей мере 20 нт, по меньшей мере 25 нт, по меньшей мере 30 нт, по меньшей мере 35 нт или по меньшей мере 40 нт. Нуклеотидная последовательность направляющей нуклеиновой кислоты, которая комплементарна нуклеотидной последовательности (целевой последовательности) целевой нуклеиновой кислоты, может иметь в длину от 12 нуклеотидов (нт) до 80 нт, от 12 нт до 50 нт, от 12 нт до 45 нт, от 12 нт до 40 нт, от 12 нт до 35 нт, от 12 нт до 30 нт, от 12 нт до 25 нт, от 12 нт до 20 нт, от 12 нт до 19 нт, от 19 нт до 20 нт, от 19 нт до 25 нт, от 19 нт до 30 нт, от 19 нт до 35 нт, от 19 нт до 40 нт, от 19 нт до 45 нт, от 19 нт до 50 нт, от 19 нт до 60 нт, от 20 нт до 25 нт, от 20 нт до 30 нт, от 20 нт до 35 нт, от 20 нт до 40 нт, от 20 нт до 45 нт, от 20 нт до 50 нт или от 20 нт до 60 нт.

Последовательность протоспейсера может быть идентифицирована путем идентификации РАМ в представляющем интерес участке и выбора участка требуемого размера выше или ниже от РАМ в качестве протоспейсера. Соответствующая последовательность спейсера может быть составлена путем определения комплементарной последовательности протоспейсерного участка.

Последовательность спейсера может быть идентифицирована с помощью компьютерной программы (например, машиночитаемого кода). Компьютерная программа может использовать такие переменные, как прогнозируемая температура плавления, образование вторичной структуры и прогнозируемая температура отжига, идентичность последовательности, геномный контекст, доступность хроматина, % GC, частота появления в геноме, статус метилирования, наличие SNP и т.д.

Степень комплементарности между нацеливающей на нуклеиновую кислоту последовательностью (например, спейсерной последовательностью) и целевой нуклеиновой кислотой (например, протоспейсером) может составлять по меньшей мере 60%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95%, по меньшей мере 97%, по меньшей мере 98%, по меньшей мере 99% или 100%. Степень комплементарности между нацеливающей на нуклеиновую кислоту последовательностью и целевой нуклеиновой кислотой может составлять по меньшей мере 60%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95%, по меньшей мере 97%, по меньшей мере 98%, по меньшей мере 99% или 100% на протяжении примерно 20 смежных нуклеотидов.

Связывающий белок Cas сегмент направляющей нуклеиновой кислоты может содержать две цепочки нуклеотидов (например, crРНК и tracrРНК), комплементарные друг другу. Две цепочки нуклеотидов (например, crРНК и tracrРНК), которые комплементарны друг другу, могут ковалентно соединиться промежуточными нуклеотидами (например, линкером в случае одинарной направляющей нуклеиновой кислоты). Две цепочки нуклеотидов (например, crРНК и tracrРНК), которые комплементарны друг другу, могут гибридизоваться с образованием двухцепочечного дуплекса или шпильки РНК связывающего белок Cas сегмента, при этом образуется структура типа "стебель и петля". crРНК и tracrРНК могут ковалентно соединиться через 3'-конец crРНК и 5'-конец tracrРНК. С другой стороны, tracrРНК и crРНК могут ковалентно соединиться через 5'-конец tracrРНК и 3'-конец crРНК.

Связывающий белок Cas сегмент направляющей нуклеиновой кислоты может иметь в длину от 10 нуклеотидов до 100 нуклеотидов, например, от 10 нуклеотидов (нт) до 20 нт, от 20 нт до 30 нт, от 30 нт до 40 нт, от 40 нт до 50 нт, от 50 нт до 60 нт, от 60 нт до 70 нт, от 70 нт до 80 нт, от 80 нт до 90 нт или от 90 нт до 100 нт. Например, связывающий белок Cas сегмент направляющей нуклеиновой кислоты может иметь длину от 15 нуклеотидов (нт) до 80 нт, от 15 нт до 50 нт, от 15 нт до 40 нт, от 15 нт до 30 нт или от 15 нт до 25 нт.

Дуплекс дцРНК связывающего белок Cas сегмента направляющей нуклеиновой кислоты может иметь в длину от 6 пар оснований (п.о.) до 50 п.о. Например, дуплекс дцРНК связывающего белок Cas сегмента может иметь длину от 6 до 40 п.о., от 6 п.о. до 30 п.о., от 6 до 25 п.о., от 6 до 20 п.о., от 6 до 15 п.о., от 8 п.о. до 40 п.о., от 8 п.о. до 30 п.о., от 8 п.о. до 25 п.о., от 8 п.о. до 20 п.о. или от 8 п.о. до 15 п.о. Например, дуплекс дцРНК связывающего белок Cas сегмента может иметь длину от 8 до 10 п.о., от 10 до 15 п.о., от 15 до 18 п.о., от 18 до 20 п.о., от 20 п.о. до 25 п.о., от 25 п.о. до 30 п.о., от 30 п.о. до 35 п.о., от 35 п.о. до 40 п.о. или от 40 п.о. до 50 п.о. В некоторых воплощениях дуплекс дцРНК связывающего белок Cas сегмента может иметь длину в 36 пар оснований. Степень комплементарности между нуклеотидными последовательностями, которые гибридизуются с образованием дуплекса дцРНК связывающего белок Cas сегмента, может составлять по меньшей мере 60%. Например, степень комплементарности между нуклеотидными последовательностями, которые гибридизуются с образованием дуплекса дцРНК связывающего белок Cas сегмента, может составлять по меньшей мере 65%, по меньшей мере 70%, по меньшей мере 75%, по меньшей мере 80%, по меньшей мере 85%, по меньшей мере 90%, по меньшей мере 95%, по меньшей мере 98% или по меньшей мере 99%. В некоторых случаях степень комплементарности между нуклеотидными последовательностями, которые гибридизуются с образованием дуплекса дцРНК связывающего белок Cas сегмента, составляет 100%.

Линкер (например, связывающий crРНК и tracrРНК в одинарной направляющей нуклеиновой кислоте) может иметь длину от 3 нуклеотидов до 100 нуклеотидов. Например, линкер может иметь длину от 3 нуклеотидов (нт) до 90 нт, от 3 нт до 80 нт, от 3 нт до 70 нт, от 3 нт до 60 нт, от 3 нт до 50 нт, от 3 нт до 40 нт, от 3 нт до 30 нт, от 3 нт до 20 нт или от 3 нт до 10 нт. Например, линкер может иметь длину от 3 нт до 5 нт, от 5 нт до 10 нт, от 10 нт до 15 нт, от 15 нт до 20 нт, от 20 нт до 25 нт, от 25 нт до 30 нт, от 30 нт до 35 нт, от 35 нт до 40 нт, от 40 нт до 50 нт, от 50 нт до 60 нт, от 60 нт до 70 нт, от 70 нт до 80 нт, от 80 нт до 90 нт или от 90 нт до 100 нт. В некоторых воплощениях линкер нацеливающей на ДНК РНК составляет 4 нт.

Гидовые нуклеиновые кислоты могут содержать модификации или последовательности, которые обеспечивают дополнительные желательные характеристики (например, модифицированную или регулируемую устойчивость, внутриклеточный таргетинг, отслеживание по флуоресцентной метке, сайты связывания для белков или белковых комплексов и т.п.). Примеры таких модификаций включают, к примеру, 5'-кэп (например, 7-метилгуанилатный кэп (m7G)); 3'-полиаденилированный хвост (т.е. хвост 3'-поли(A)); последовательность рибопереключателя (например, для обеспечения регулируемой устойчивости и/или регулируемой доступности для белков и/или белковых комплексов); последовательность контроля устойчивости; последовательность, которая образует дуплекс дцРНК (т.е. шпильку); модификации или последовательности, которые направляют РНК во внутриклеточную локализацию (например, в ядро, митохондрии, хлоропласты и т.п.); модификации или последовательности, которые обеспечивают отслеживание (например, прямое конъюгирование с флуоресцентной молекулой, конъюгирование с молекулой, которая облегчает флуоресцентное детектирование, с последовательностью, которая способствует флуоресцентному детектированию и т.д.); модификации или последовательности, которые обеспечивают сайты связывания для белков (например, белков, которые действуют на ДНК, включая активаторы транскрипции, репрессоры транскрипции, ДНК-метилтрансферазы, ДНК-деметилазы, гистон-ацетилтрансферазы, гистондеацетилазы и их комбинации).

Направляющая нуклеиновая кислота может содержать одну или несколько модификаций (например, модификации оснований, модификации остова), чтобы придать нуклеиновой кислоте новые или улучшенные характеристики (например, большую устойчивость). Направляющая нуклеиновая кислота может содержать аффинную метку нуклеиновой кислоты. Нуклеозид - это сочетание основания и сахара. Основанием у нуклеозида может быть гетероциклическое основание. Два самых распространенных класса таких гетероциклических оснований - пурины и пиримидины. Нуклеотиды - это нуклеозиды, которые дополнительно содержат фосфатную группу, ковалентно связанную с сахарной частью нуклеозида. Для

тех нуклеозидов, которые содержат пентофуранозильный сахар, фосфатная группа может соединяться с 2'-, 3'- или 5'-гидроксильной группой сахара. При формировании гидовых нуклеиновых кислот фосфатные группы могут ковалентно связывать соседние нуклеозиды друг с другом с образованием линейного полимерного соединения. В свою очередь, соответствующие концы этого линейного полимерного соединения могут дополнительно соединяться с образованием циклического соединения; однако обычно подходят линейные соединения. Кроме того, линейные соединения могут содержать внутренние комплементарные пары нуклеотидов, поэтому они могут укладываться таким образом, чтобы получилось полностью или частично двухцепочечное соединение. У направляющих нуклеиновых кислот фосфатные группы обычно составляют межнуклеозидный остов направляющей нуклеиновой кислоты. Связи у остова направляющей нуклеиновой кислоты могут представлять собой 3'-5'-фосфодиэфирную связь.

Направляющая нуклеиновая кислота может содержать модифицированный остов и/или модифицированные межнуклеозидные связи. Модифицированные остовы могут включать в себя такие, у которых сохраняются атомы фосфора в остове, и такие, у которых нет атомов фосфора в остове.

Подходящие модифицированные остовы гидовых нуклеиновых кислот, содержащие атомы фосфора, могут включать, к примеру, фосфоротиоаты, хиральные фосфоротиоаты, фосфородитиоаты, фосфотриэфиры, аминоалкилфосфотриэфиры, метил- и другие алкилфосфонаты типа 3'-алкиленфосфонатов и 5'-алкиленфосфонатов, хиральные фосфонаты, фосфинаты, фосфорамидаты, в том числе 3'-аминофосфорамидаты и аминоалкил-фосфорамидаты, фосфородиамидаты, тионофосфорамидаты, тиоалкилфосфонаты, тио-ноалкилфосфотриэфиры, селенофосфаты и боранофосфаты с нормальными 3'-5'-связями, аналоги с 2'-5'-связями и таковые с инвертированной полярностью, у которых одна или несколько межнуклеотидных связей представлены 3'-3'-, 5'-5'- и 2'-2'-связями. Подходящие гидовые нуклеиновые кислоты с инвертированной полярностью могут содержать одну 3'-3'-связь на самой 3'-концевой межнуклеотидной связи (т.е. один инвертированный остаток нуклеозида, в котором отсутствует основание нуклеозида или вместо него содержится гидроксильная группа). Также сюда входят различные соли (например, хлорида калия или хлорида натрия), смешанные соли и формы свободной кислоты.

Направляющая нуклеиновая кислота может содержать одну или несколько фосфоротиоатных и/или гетероатомных межнуклеотидных связей, в частности $-\text{CH}_2\text{-NH-O-CH}_2-$, $-\text{CH}_2\text{-N}(\text{CH}_3)\text{-O-CH}_2-$ (т.е. это метиленовый (метиляминовый) остов или MMI), $-\text{CH}_2\text{-O-N}(\text{CH}_3)\text{-CH}_2-$, $-\text{CH}_2\text{-N}(\text{CH}_3)\text{-N}(\text{CH}_3)\text{-CH}_2-$ и $-\text{O-N}(\text{CH}_3)\text{-CH}_2\text{-CH}_2-$ (при этом нативная фосфодиэфирная межнуклеотидная связь изображается как $-\text{O-P}(\text{=O})(\text{OH})\text{-O-CH}_2-$).

Направляющая нуклеиновая кислота может содержать морфолиновую (morpholino) структуру в остове. Например, нуклеиновая кислота может содержать 6-членное морфолиновое кольцо вместо кольца рибозы. В некоторых из этих воплощений фосфоро-диамидатная или другая нефосфодиэфирная межнуклеотидная связь заменяет фосфодиэфирную связь.

Направляющая нуклеиновая кислота может содержать полинуклеотидный остов, который образован связями между короткоцепочечными алкил- или циклоалкилнуклеозидами, между смешанными гетероатомными и алкил- или циклоалкилнуклеозидами или одной или несколькими связями между короткоцепочечными гетероатомными или гетероциклическими нуклеозидами. К таковым можно отнести морфолиновые остовы (частично образованные из сахарной части нуклеозидов); силоксановые остовы; сульфидные, сульфоксидные и сульфоновые остовы; формацетильные и тиоформацетильные остовы; метиленформацетильные и метилентиоформацетильные остовы; рибоацетильные остовы; алкенсодержащие остовы; сульфаматные остовы; метилениминовые и метиленигидразиновые остовы; сульфонатные и сульфонамидные остовы; амидные остовы; и другие, имеющие смешанные составные части N, O, S и CH_2 .

Направляющая нуклеиновая кислота может содержать миметик нуклеиновой кислоты. Термин "миметик" служит для обозначения таких полинуклеотидов, у которых только фуранозное кольцо либо и фуранозное кольцо, и межнуклеотидная связь заменены на нефуранозные группы, причем замена только фуранозного кольца также может называться сахарным суррогатом. Гетероциклические основания нуклеотидов или модифицированные гетероциклические основания нуклеотидов могут сохраняться для гибридизации с подходящей целевой нуклеиновой кислотой. Одной из таких нуклеиновых кислот может быть пептидная нуклеиновая кислота (PNA). У PNA сахарный остов полинуклеотида может быть заменен на амидосодержащий остов, в частности на аминоэтилглициновый остов. Нуклеотиды могут сохраняться и связываться прямо или косвенно с атомами азота в амидной части остова. У соединений PNA остов может содержать два или несколько соединенных аминоэтилглициновых звеньев, что даёт PNA амидосодержащий остов. Гетероциклические основания нуклеотидов могут прямо или косвенно связываться с атомами азота в амидной части остова.

Направляющая нуклеиновая кислота может содержать соединенные морфолиновые звенья (т.е. это морфолинонуклеиновая кислота) с гетероциклическими основаниями, присоединенными к морфолиновому кольцу. Соединительные группы могут связывать морфолиновые мономерные звенья в морфолинонуклеиновой кислоте. Неионные олигомерные соединения на основе морфолинов могут иметь меньше нежелательных взаимодействий с клеточными белками. Полинуклеотиды на основе морфолинов могут

быть неионными миметиками гидовых нуклеиновых кислот. Различные соединения в пределах класса морфолинов могут соединяться с помощью различных соединительных групп. Следующий класс полинуклеотидных миметиков можно обозначить как циклогексенилнуклеиновые кислоты (CeNA). Фуранозное кольцо, обычно присутствующее в молекуле нуклеиновой кислоты, может быть заменено циклогексенильным кольцом. Можно получить защищенные DMT фосфорамидитные мономеры CeNA и использовать их для синтеза олигомерных соединений методами химии фосфорамидитов. Включение мономеров CeNA в цепь нуклеиновой кислоты может повысить стабильность гибридов ДНК/РНК. Олигоаденилаты CeNA могут образовывать комплексы с комплементарными нуклеиновыми кислотами с такой же стабильностью, как у нативных комплексов. Следующая модификация может включать блокированные нуклеиновые кислоты (LNAs), у которых 2'-гидроксильная группа соединяется с 4'-атомом углерода в сахарном кольце, при этом образуется 2'-С,4'-С-оксиметиленовая связь, причем образуется бициклическая молекула сахара. Связывание может представлять собой метиленовую (-CH₂-) группу, соединяющую 2'-атом кислорода и 4'-атом углерода, где n равно 1 или 2. LNA и аналоги LNA могут проявлять очень высокую термостабильность дуплекса с комплементарной нуклеиновой кислотой (T_m от +3 до +10°C), устойчивостью к деградации 3'-экзонуклеазами и хорошие свойства растворимости. Направляющая нуклеиновая кислота может содержать одну или несколько замещенных молекул сахара. Подходящие полинуклеотиды могут содержать заместители для сахаров, выбранные из: OH; F; O-, S- или N-алкила; O-, S- или N-алкенила; O-, S- или N-алкинила; либо O-алкил-O-алкила, причем алкил, алкенил или алкинил может представлять собой замещенный или незамещенный C₁-C₁₀-алкил или C₂-C₁₀-алкенил или алкинил. Особенно подходят O((CH₂)_nO)_mCH₃, O(CH₂)_nOCH₃, O(CH₂)_nNH₂, O(CH₂)_nCH₃, O(CH₂)_nONH₂ и O(CH₂)_nON((CH₂)_nCH₃)₂, где n и m равны от 1 до 10. Заместитель для сахара может быть выбран из: низшего C₁-C₁₀-алкила, замещенного низшего алкила, алкенила, алкинила, алкарила, аралкила, O-алкарила или O-аралкила, SH, SCH₃, OCN, Cl, Br, CN, CF₃, OCF₃, SOCH₃, SO₂CH₃, ONO₂, NO₂, N₃, NH₂, гетероциклоалкила, гетероциклоалкарила, аминокламино, полиалкиламино, замещенного силила, расщепляющей РНК группы, репортерной группы, интеркалятора, группы для улучшения фармакокинетических свойств направляющей нуклеиновой кислоты или группы для улучшения фармакодинамических свойств направляющей нуклеиновой кислоты и других заместителей, имеющих близкие свойства. Подходящая модификация может включать 2'-метоксиэтокси (2'-O-CH₂CH₂OCH₃, также известна как 2'-O-(2-метоксиэтил) или 2'-МОЕ, т.е. это алкоксиалкоксигруппа). Другие подходящие модификации могут включать 2'-диметиламинооксиэтокси (т.е. это группа O(CH₂)₂ON(CH₃)₂, также известная как 2'-DMAOE) и 2'-диметиламиноэтоксиэтокси (также известная как 2'-O-диметиламиноэтоксиэтил или 2'-DMAEOE), т.е. это 2'-O-CH₂-O-CH₂-N(CH₃)₂.

Другие подходящие заместители для сахаров включают метокси (-O-CH₃), аминопропокси (-OCH₂CH₂CH₂NH₂), аллил (-CH₂-CH=CH₂), -O-аллил (-O-CH₂-CH=CH₂) и фтор (F). Заместители для 2'-сахаров могут находиться в положении арабинозы (вверху) или рибозы (внизу). Подходящая 2'-арабинозная модификация - 2'-F. Аналогичные модификации могут проводиться и по другим положениям в олигомерных соединениях, в частности, в 3'-положении сахара у 3'-концевого нуклеозида или у 2'-5'-связанных нуклеотидов и в 5'-положении 5'-концевого нуклеотида. Олигомерные соединения также могут содержать миметики сахаров типа циклобутиловых вместо пентофуранозильных сахаров.

Направляющая нуклеиновая кислота также может включать модификации или замены нуклеиновых оснований (которые часто называют просто "основаниями"). В настоящем изобретении "немодифицированные" или "природные" нуклеотиды могут включать в себя пуриновые основания (например, аденин (A) и гуанин (G)) и пиримидиновые основания (например, тимин (T), цитозин (C) и урацил (U)). Модифицированные нуклеиновые основания могут включать и другие синтетические и природные нуклеиновые основания, такие как 5-метилцитозин (5-me-C), 5-гидроксиметилцитозин, ксантин, гипоксантин, 2-аминоаденин, 6-метил- и другие алкильные производные аденина и гуанина, 2-пропил- и другие алкильные производные аденина и гуанина, 2-тиоурацил, 2-тиотимин и 2-тиоцитозин, 5-галоурацил и 5-галоцитозин, 5-пропинил- (-C≡C-CH₃) урацил и цитозин и другие алкильные производные пиримидиновых оснований, 6-азо-урацил, цитозин и тимин, 5-урацил (псевдоурацил), 4-тиоурацил, 8-гало, 8-амино, 8-тиол, 8-тиоалкил, 8-гидроксил и другие 8-замещенные аденины и гуанины, 5-гало, в частности 5-бром, 5-трифторметил и другие 5-замещенные урацилы и цитозины, 7-метилгуанин и 7-метиладенин, 2-F-аденин, 2-аминоаденин, 8-азагуанин и 8-азааденин, 7-деазагуанин и 7-деазааденин, 3-деазагуанин и 3-деазааденин. Модифицированные нуклеиновые основания могут включать такие трициклические пиримидины, как феноксазин-цитидин (1H-пиримидо(5,4-b)(1,4)бензо-ксазин-2(3H)-он), фенотиазин-цитидин (1H-пиримидо(5,4-b)(1,4)бензотиазин-2(3H)-он), такие "струбцины", как замещенный феноксазин-цитидин (например, 9-(2-аминоэтокси)-H-пиримидо(5,4-b)(1,4)бензоксазин-2(3H)-он), карбазол-цитидин (2H-пиримидо(4,5-b)индол-2-он), пиридоиндол-цитидин (H-пиридо(3',2':4,5)пирроло(2,3-d)пиримидин-2-он).

Гетероциклические основания нуклеотидов могут включать такие, у которых пуриновое или пиримидиновое основание заменено на другой гетероцикл, к примеру, 7-деазааденин, 7-деазагуанин, 2-аминопиридин или 2-пиридон. Нуклеиновые основания могут быть полезны для повышения средства связывания полинуклеотидных соединений. Они могут включать 5-замещенные пиримидины, 6-азапиримидины и N-2, N-6 и O-6-замещенные пурины, включая 2-аминопропиладенин, 5-пропинилурацил и 5-

пропинил-цитозин. Замены на 5-метилцитозин могут повысить стабильность дуплекса нуклеиновой кислоты на 0,6-1,2°C и могут быть подходящими заменами оснований (например, в сочетании с модификацией сахаров 2'-О-метоксиэтилом).

Модификация направляющей нуклеиновой кислоты может включать химическое присоединение к направляющей нуклеиновой кислоте одной или нескольких молекул или конъюгатов, которые могут усиливать активность, клеточное распределение или клеточное поглощение направляющей нуклеиновой кислоты. Эти молекулы или конъюгаты могут включать конъюгатные группы, ковалентно связанные с такими функциональными группами, как первичные или вторичные гидроксильные группы. Конъюгатные группы могут включать, без ограничения, интеркаляторы, репортерные молекулы, полиамины, полиамиды, полиэтиленгликоли, полиэфиры, группы, которые усиливают фармакодинамические свойства олигомеров, и группы, которые могут усиливать фармакокинетические свойства олигомеров. Конъюгатные группы могут включать, без ограничения, холестерин, липиды, фосфолипиды, биотин, феназин, фолат, фенантридин, антрахинон, акридин, флуоресцеины, родамины, кумарины и красители. Группы, которые усиливают фармакодинамические свойства, включают группы, которые улучшают поглощение (захват), повышают устойчивость к деградации и/или усиливают специфичную к последовательности гибридизацию с целевой нуклеиновой кислотой. Группы, которые могут усиливать фармакокинетические свойства, включают группы, которые улучшают поглощение (захват), распределение, метаболизм или выведение нуклеиновой кислоты. Конъюгатные молекулы могут включать, без ограничения, молекулы липидов типа холестерина, тиоэфиры (например, гексил-S-третилтиол) холевой кислоты, тиохолестерин, алифатические цепи (например, додекандиол или ундецильные остатки), фосфолипиды (например, дигексадецил-гас-глицерин или триэтиламмоний-1,2-ди-О-гексадецил-гас-глицеро-3-Н-фосфонат), цепи полиаминов или полиэтиленгликолев или адамантануксусной кислоты, пальмитиловые молекулы или молекулы октадециламина или гексиламинокарбонил-оксихолестерина.

Модификация может включать "домен трансдукции белка" или РТД (т.е. это проникающий в клетки пептид (СРР)). РТД может означать полипептид, полинуклеотид, углевод либо органическое или неорганическое соединение, которое способствует пересечению липидного бислоя, мицелл, клеточных мембран, мембран органелл или мембран пузырьков. РТД может присоединяться к другой молекуле, которая может варьироваться от небольшой полярной молекулы до большой макромолекулы и/или наночастицы и может способствовать пересечению молекулой мембраны, к примеру, переходу из внеклеточного пространства во внутриклеточное пространство либо из цитозоля внутрь органеллы. РТД может быть ковалентно связан с N-концом полипептида. РТД может быть ковалентно связан с C-концом полипептида. РТД может быть ковалентно связан с нуклеиновой кислотой. Типичные РТД могут включать, без ограничения, минимальный пептид домена трансдукции белка; полиаргининовую последовательность, содержащую ряд аргининов, достаточных для проникновения в клетку (например, 3, 4, 5, 6, 7, 8, 9, 10 или 10-50 аргининов), домен VP22, домен трансдукции белка Antennapedia дрозофилы, укороченный пептид кальцитонина человека, полилизин и транспортан, гомополимер аргинина из 3-50 остатков аргинина (SEQ ID NO: 87). РТД может представлять собой активируемый СРР (АСРР). АСРР может включать поликатионный СРР (например, Arg9 или "R9" (SEQ ID NO: 88)), соединенный через расщепляемый линкер с соответствующим полианионом (например, Glu9 или "E9" (SEQ ID NO: 89)), который может уменьшать общий заряд почти до нуля и тем самым подавлять адгезию и поглощение клетками. При расщеплении линкера полианион может высвободиться, локально демаскируя полиаргинин и присущую ему адгезивность, тем самым "активируя" АСРР для пересечения мембраны.

Направляющая нуклеиновая кислота может быть представлена в любой форме. К примеру, направляющая нуклеиновая кислота может быть представлена в виде РНК, как в виде двух молекул (например, отдельно *sc*РНК и *tracr*РНК), так и в виде одной молекулы (например, *sg*РНК). Направляющая нуклеиновая кислота может быть представлена в виде комплекса с белком Cas. Направляющая нуклеиновая кислота также может быть представлена в виде ДНК, кодирующей РНК. ДНК, кодирующая гидовую нуклеиновую кислоту, может кодировать единую гидовую нуклеиновую кислоту (например, *sg*РНК) или отдельные молекулы РНК (например, отдельно *sc*РНК и *tracr*РНК). В последнем случае ДНК, кодирующая гидовую нуклеиновую кислоту, может быть представлена в виде отдельных молекул ДНК, кодирующих *sc*РНК и *tracr*РНК, соответственно.

ДНК, кодирующая гидовую нуклеиновую кислоту, может быть стабильно встроена в геном клетки и, необязательно, функционально связана с промотором, активным в этой клетке. ДНК, кодирующая гидовую нуклеиновую кислоту, может быть функционально связана с промотором в экспрессирующей конструкции.

Направляющая нуклеиновая кислота может быть получена любым подходящим способом. Например, направляющая нуклеиновая кислота может быть получена путем транскрипции *in vitro*, к примеру, с помощью РНК-полимеразы T7. Направляющая нуклеиновая кислота также может быть получена синтетически методами химического синтеза.

Направляющая нуклеиновая кислота может содержать последовательность для повышения стабильности. Например, направляющая нуклеиновая кислота может содержать сегмент терминатора транскрипции (т.е. последовательность терминации транскрипции). Сегмент терминатора транскрипции мо-

жет иметь общую длину от 10 нуклеотидов до 100 нуклеотидов, например, от 10 нуклеотидов (нт) до 20 нт, от 20 нт до 30 нт, от 30 нт до 40 нт, от 40 нт до 50 нт, от 50 нт до 60 нт, от 60 нт до 70 нт, от 70 нт до 80 нт, от 80 нт до 90 нт или от 90 нт до 100 нт. Например, сегмент терминатора транскрипции может иметь длину от 15 нуклеотидов (нт) до 80 нт, от 15 нт до 50 нт, от 15 нт до 40 нт, от 15 нт до 30 нт или от 15 нт до 25 нт. Последовательность терминации транскрипции может функционировать в эукариотических клетках или прокариотических клетках.

В различных воплощениях изложенных здесь аспектов используются одновременно несколько исполнительных элементов в одних и тех же клетках. В некоторых воплощениях исполнительный элемент, содержащий белок Cas, может использоваться одновременно со вторым исполнительным элементом, содержащим нуклеазу с "цинковым пальцем" (ZFN), эффекторную нуклеазу типа активатора транскрипции (TALEN), мегануклеазу, белок, связывающий РНК (RBP), CRISPR-ассоциированный белок, связывающий РНК, рекомбиназу, флиппазу, транспозазу или белок Argonaute. В некоторых воплощениях исполнительный элемент, содержащий ZFN, может использоваться одновременно со вторым исполнительным элементом, содержащим белок Cas, нуклеазу с "цинковым пальцем" (ZFN), эффекторную нуклеазу типа активатора транскрипции (TALEN), мегануклеазу, белок, связывающий РНК (RBP), CRISPR-ассоциированный белок, связывающий РНК, рекомбиназу, флиппазу, транспозазу или белок Argonaute. В некоторых воплощениях исполнительный элемент, содержащий TALEN, может использоваться одновременно со вторым исполнительным элементом, содержащим белок Cas, нуклеазу с "цинковым пальцем" (ZFN), мегануклеазу, белок, связывающий РНК (RBP), CRISPR-ассоциированный белок, связывающий РНК, рекомбиназу, флиппазу, транспозазу или белок Argonaute. В некоторых воплощениях исполнительный элемент, содержащий мегануклеазу, может использоваться одновременно со вторым исполнительным элементом, содержащим белок Cas, нуклеазу с "цинковым пальцем" (ZFN), эффекторную нуклеазу типа активатора транскрипции (TALEN), белок, связывающий РНК (RBP), CRISPR-ассоциированный белок, связывающий РНК, рекомбиназу, флиппазу, транспозазу или белок Argonaute. В некоторых воплощениях исполнительный элемент, содержащий белок, связывающий РНК (RBP), может использоваться одновременно со вторым исполнительным элементом, содержащим белок Cas, нуклеазу с "цинковым пальцем" (ZFN), эффекторную нуклеазу типа активатора транскрипции (TALEN), мегануклеазу, CRISPR-ассоциированный белок, связывающий РНК, рекомбиназу, флиппазу, транспозазу или белок Argonaute. В некоторых воплощениях исполнительный элемент, содержащий CRISPR-ассоциированный белок, связывающий РНК, может использоваться одновременно со вторым исполнительным элементом, содержащим белок Cas, нуклеазу с "цинковым пальцем" (ZFN), эффекторную нуклеазу типа активатора транскрипции (TALEN), мегануклеазу, белок, связывающий РНК (RBP), рекомбиназу, флиппазу, транспозазу или белок Argonaute. В некоторых воплощениях исполнительный элемент, содержащий рекомбиназу, может использоваться одновременно со вторым исполнительным элементом, содержащим белок Cas, нуклеазу с "цинковым пальцем" (ZFN), эффекторную нуклеазу типа активатора транскрипции (TALEN), мегануклеазу, белок, связывающий РНК (RBP), CRISPR-ассоциированный белок, связывающий РНК, флиппазу, транспозазу или белок Argonaute. В некоторых воплощениях исполнительный элемент, содержащий флиппазу, может использоваться одновременно со вторым исполнительным элементом, содержащим белок Cas, нуклеазу с "цинковым пальцем" (ZFN), эффекторную нуклеазу типа активатора транскрипции (TALEN), мегануклеазу, белок, связывающий РНК (RBP), CRISPR-ассоциированный белок, связывающий РНК, рекомбиназу, транспозазу или белок Argonaute. В некоторых воплощениях исполнительный элемент, содержащий транспозазу, может использоваться одновременно со вторым исполнительным элементом, содержащим белок Cas, нуклеазу с "цинковым пальцем" (ZFN), эффекторную нуклеазу типа активатора транскрипции (TALEN), мегануклеазу, белок, связывающий РНК (RBP), CRISPR-ассоциированный белок, связывающий РНК, рекомбиназу, флиппазу, или белок Argonaute. В некоторых воплощениях исполнительный элемент, содержащий белок Argonaute, может использоваться одновременно со вторым исполнительным элементом, содержащим белок Cas, нуклеазу с "цинковым пальцем" (ZFN), эффекторную нуклеазу типа активатора транскрипции (TALEN), мегануклеазу, белок, связывающий РНК (RBP), CRISPR-ассоциированный белок, связывающий РНК, рекомбиназу, флиппазу, или транспозазу.

В некоторых воплощениях несколько исполнительных элементов используются одновременно в одних и тех же клетках, чтобы одновременно модулировать транскрипцию в разных местах на одной и той же ДНК-мишени или на разных ДНК-мишенях. В некоторых воплощениях исполнительный элемент содержит Cas-нуклеазу. Для наводки на различные нуклеиновые кислоты в нескольких комплексах CRISPR/Cas может использоваться белок Cas из одного источника или одного типа с несколькими гидовыми нуклеиновыми кислотами. С другой стороны, для наводки на несколько нуклеиновых кислот в нескольких комплексах CRISPR/Cas могут использоваться ортологические белки Cas (например, мертвые белки Cas9 из различных организмов типа *S.pyogenes*, *S.aureus*, *S.thermophilus*, *L.innocua* и *N.meningitidis*).

В некоторых воплощениях можно одновременно использовать несколько направляющих нуклеиновых кислот в одних и тех же клетках, чтобы одновременно модулировать транскрипцию в разных местах на одной и той же ДНК-мишени или на разных ДНК-мишенях. В некоторых воплощениях две или не-

сколько направляющих нуклеиновых кислот нацелены на один и тот же ген или транскрипт или локус. В некоторых воплощениях две или несколько направляющих нуклеиновых кислот нацелены на разные несвязанные локусы. В некоторых воплощениях две или несколько направляющих нуклеиновых кислот нацелены на разные, но связанные локусы.

Две или несколько направляющих нуклеиновых кислот могут одновременно присутствовать на одном и том же экспрессирующем векторе. Две или несколько направляющих нуклеиновых кислот могут находиться под одним и тем же транскрипционным контролем. В некоторых воплощениях в клетках мишени одновременно экспрессируются две или несколько (например, 3 или более, 4 или более, 5 или более, 10 или более, 15 или более, 20 или более, 25 или более, 30 или более, 35 или более, 40 или более, 45 или более или 50 или более) направляющих нуклеиновых кислот (из одного или из разных векторов). Экспрессируемые гидовые нуклеиновые кислоты могут по-разному распознаваться мертвыми Cas-белками (например, белками dCas9 из различных бактерий типа *S.pyogenes*, *S.aureus*, *S.thermophilus*, *L.innocua* и *N.meningitides*).

Для экспрессирования нескольких направляющих нуклеиновых кислот можно использовать искусственную систему обработки направляющих нуклеиновых кислот при помощи эндонуклеазы (например, для обработки направляющих нуклеиновых кислот можно использовать эндорибонуклеазу Csy4). Например, несколько направляющих РНК могут быть конкатенированы в tandemном порядке на транскрипте-предшественнике (например, экспрессируемом из промотора U6) и разделены Csy4-специфичной последовательностью РНК. Экспрессируемый совместно белок Csy4 может расщеплять транскрипт-предшественник на несколько направляющих РНК. Поскольку все гидовые РНК обрабатываются из транскрипта-предшественника, то их концентрации будут нормализованы для одинакового связывания dCas9.

Промоторы, которые можно использовать со способами и композициями изобретения, включают, к примеру, промоторы, активные в эукариотических клетках, клетках млекопитающих или человека. Промотор может быть индуцибельным или конститутивно активным промотором. С другой стороны, промотор может быть специфичным для ткани или клетки.

Неограничивающие примеры подходящих эукариотических промоторов (то есть промоторов, функционирующих в эукариотических клетках) могут включать такие промоторы, как самый ранний промотор цитомегаловируса (CMV), промотор тимидин-киназы вируса Herpes simplex (HSV), ранний и поздний промотор SV40, длинные концевые повторы (LTRs) из ретровирусов, промотор фактора-1 элонгации человека (EF1), гибридные конструкции, содержащие энхансер цитомегаловируса (CMV), слитый с промотором β-актина курицы (CAG), промотор вируса стволовых клеток мыши (MSCV), промотор локуса фосфоглицераткиназы-1 (PGK) и металлотионеина-I мыши. Промотором может быть промотор грибов. Промотором может быть промотор растений. Имеются базы данных по промоторам растений (например, PlantProm). Экспрессирующий вектор также может содержать сайт связывания рибосом для инициации трансляции и терминатор транскрипции. Экспрессирующий вектор также может содержать соответствующие последовательности для усиления экспрессии.

Для введения композиций и молекул (например, полипептидов и/или кодирующих полипептиды нуклеиновых кислот) изобретения в клетки хозяина можно использовать любой подходящий способ доставки. Композиции (например, исполнительный элемент типа Cas-белка, слитого или химерного; химерный рецептор, адаптер, направляющая нуклеиновая кислота) могут быть доставлены одновременно или раздельно по времени. Выбор метода генетической модификации может зависеть от типа подлежащих трансформации клеток и/или обстоятельств, при которых происходит трансформация (например, *in vitro*, *ex vivo* или *in vivo*).

Способ доставки может включать контактирование целевого полинуклеотида или введение в клетки (или популяцию клеток) одной или нескольких нуклеиновых кислот, содержащих последовательности нуклеотидов, кодирующие композиции по изобретению (например, исполнительный элемент типа Cas-белка или Cas-химеры, химерный рецептор, химерный адаптер, гидовую нуклеиновую кислоту и т.д.). Подходящие нуклеиновые кислоты, содержащие последовательности нуклеотидов, кодирующие композиции по изобретению, могут включать экспрессирующие векторы, причем экспрессирующий вектор, содержащий последовательность нуклеотидов, кодирующую одну или несколько композиций по изобретению (например, исполнительный элемент типа Cas-белка или Cas-химеры, химерный рецептор, химерный адаптер, направляющую нуклеиновую кислоту и т.д.) является рекомбинантным экспрессирующим вектором.

Неограничивающие примеры способов доставки или трансформации включают, к примеру, инфицирование вирусом или бактериофагом, трансфекцию, конъюгацию, слияние протопластов, липофекцию, электропорацию, осаждение фосфатом кальция, трансфекцию с помощью полиэтиленимина (PEI), трансфекцию с помощью DEAE-декстрана, трансфекцию при помощи липосом, технологию пушки для частиц, осаждение фосфатом кальция, прямые микроинъекции и доставку нуклеиновой кислоты при помощи проникающих в клетки пептидов и наночастиц.

В некоторых аспектах настоящего изобретения предусмотрены способы, включающие доставку одного или нескольких полинуклеотидов либо одного или нескольких описанных здесь векторов либо од-

ного или нескольких их транскриптов и/или одного или нескольких транскрибируемых из них белков в клетки хозяина. В некоторых аспектах изобретения также предусмотрены клетки, полученные такими способами, и организмы (как-то животные, растения или грибы), содержащие или полученные из таких клеток. В некоторых воплощениях в клетки вводится белок Cas и/или химерный рецептор и/или адаптер, в сочетании и необязательно в комплексе с направляющей последовательностью.

Полинуклеотид, кодирующий любой из приведенных здесь полипептидов (например, рецепторный полипептид, адаптерный полипептид, исполнительный элемент типа Cas-белка и т.п.), может быть оптимизирован по кодонам. Оптимизация кодонов может повлечь за собой мутацию чужеродной (например, рекомбинантной) ДНК с тем, чтобы она имитировала предпочтительность кодонов у предполагаемого организма или клеток хозяина, но кодировала один и тот же белок. Таким образом, кодоны могут изменяться, но кодируемый белок останется неизменным. Например, если бы намеченными клетками мишени были клетки человека, то для получения белка Cas подошел бы полинуклеотид, оптимизированный по кодонам человека. В качестве другого неограничительного примера, если бы намеченными клетками мишени были клетки мыши, то для Cas-белка подошел бы полинуклеотид, оптимизированный по кодонам мыши. Полинуклеотид, кодирующий полипептид типа исполнительного элемента (например, белок Cas), может быть оптимизирован по кодонам для многих клеток представляющих интерес хозяев. Клетками хозяина могут быть клетки из любого организма (например, бактериальные клетки, археальные клетки, клетки одноклеточных эукариотических организмов, клетки растений, клетки водорослей, напр. *Botryococcus braunii*, *Chlamydomonas reinhardtii*, *Nannochloropsis gaditana*, *Chlorella pyrenoidosa*, *Sargassum patens* C. Agardh и др., клетки грибов (например, дрожжевые клетки), клетки животных, клетки беспозвоночных животных (например, плодовой мушки, кишечнополостных, иглокожих, нематод и т.п.), клетки позвоночных животных (например, рыб, амфибий, рептилий, птиц, млекопитающих), клетки млекопитающих (например, свиней, коров, коз, овец, грызунов, крыс, мышей, обезьян, человека и т.п.) и т.д. В некоторых случаях оптимизация кодонов может и не требоваться. В некоторых случаях оптимизация кодонов может быть предпочтительной.

Для введения нуклеиновых кислот в клетки или целевые ткани млекопитающих могут применяться стандартные методы переноса генов на основе вирусов и не вирусов. Такие методы могут применяться для введения нуклеиновых кислот, кодирующих композиции по изобретению, в культуры клеток или в организм хозяина. Системы доставки типа невирусных векторов могут включать ДНК-плазмиды, РНК (например, транскрипты описанных здесь векторов), голые нуклеиновые кислоты и нуклеиновые кислоты в комплексе с носителем типа липосом. Системы доставки типа вирусных векторов могут включать ДНК- и РНК-вирусы, которые могут быть эписомальными либо встраиваться в геном после введения в клетки.

Методы невирусной доставки нуклеиновых кислот могут включать липофекцию, нуклеофекцию, микроинъекцию, биолистику, виросомы, липосомы, иммунолипосомы, поликатионы или конъюгаты липид:нуклеиновая кислота, голую ДНК, искусственные вирионы и усиленный агентом захват ДНК. Можно использовать катионные и нейтральные липиды, которые подходят для эффективной липофекции полинуклеотидов. Введение может быть в клетки (например, *in vitro* или *ex vivo*) или в ткани мишени (например, *in vivo*). Можно использовать препараты комплексов липид:нуклеиновая кислота, включая прицельные липосомы типа иммунолипидных комплексов.

Для наводки на определенные клетки в организме и транспортировки полезного груза вируса в ядро клетки можно использовать системы на основе РНК- или ДНК-вирусов. Вирусные векторы можно вводить непосредственно (*in vivo*) или же их можно использовать для обработки клеток *in vitro* и необязательно вводить модифицированные клетки (*ex vivo*). Вирусные системы могут включать ретровирусные, лентивирусные, аденовирусные, аденоассоциированные и векторы из вируса Herpes simplex для переноса генов. При переносе генов с помощью ретровирусов, лентивирусов и аденоассоциированных вирусов может происходить встраивание в геном хозяина, что может привести к долгосрочной экспрессии встроенного трансгена. Во многих различных типах клеток и тканях-мишенях может наблюдаться высокая эффективность трансдукции.

Тропизм ретровирусов может изменяться при включении чужеродных белков оболочки, что расширяет потенциальную целевую популяцию клеток-мишеней. Лентивирусные векторы - это ретровирусные векторы, которые могут трансдуцировать или инфицировать неделящиеся клетки и давать высокие титры вируса. Выбор ретровирусной системы переноса генов может зависеть от ткани-мишени. Ретровирусные векторы могут содержать цис-действующие длинные концевые повторы (LTR) с упаковочной емкостью до 6-10 т.о. чужеродной последовательности. Минимальные цис-действующие LTR могут быть достаточными для репликации и упаковки векторов, что можно использовать для встраивания терапевтического гена в клетки мишени для обеспечения непрерывной экспрессии трансгена. Ретровирусные векторы могут включать векторы на основе вируса лейкемии мыши (MuLV), вируса лейкемии гиббона (GaLV), вируса иммунодефицита обезьян (SIV), вируса иммунодефицита человека (HIV) и их комбинаций.

Можно использовать системы на основе аденовирусов. Аденовирусные системы могут привести к краткосрочной экспрессии трансгена. Аденовирусные векторы могут иметь высокую эффективность

трансдукции в клетках и не требуют деления клеток. С аденовирусными векторами можно получать высокие титры и уровни экспрессии. Для трансдукции клеток можно использовать векторы из аденоассоциированных вирусов ("AAV") с целевыми нуклеиновыми кислотами, например, при получении нуклеиновых кислот и пептидов *in vitro*, а также для процедур генной терапии *in vivo* и *ex vivo*.

Для формирования вирусных частиц, способных инфицировать клетки хозяина, можно использовать упаковочные клетки. Такие клетки могут включать клетки 293 (например, для упаковки аденовируса) и клетки Psi2 или клетки PA317 (например, для упаковки ретровируса). Вирусные векторы можно получать путем продуцирования клеточной линии, которая упаковывает нуклеиновую кислоту вектора в вирусные частицы. Векторы могут содержать минимальные вирусные последовательности, необходимые для упаковки и последующего встраивания у хозяина. Векторы могут содержать и другие вирусные последовательности, которые заменяются на экспрессирующую кассету для экспрессируемых полинуклеотидов. Недостающие вирусные функции могут возмещаться клетками упаковочной линии. Например, векторы из AAV могут содержать последовательности ITR из генома AAV, которые необходимы для упаковки и встраивания в геном хозяина. Вирусную ДНК можно упаковывать в такой линии клеток, которая содержит вспомогательную плазмиду, кодирующую другие гены AAV, а именно гер и кер, но без последовательности ITR. Линия клеток также может быть инфицирована аденовирусом в качестве вспомогательного. Вспомогательный вирус может способствовать репликации AAV-вектора и экспрессии генов AAV из вспомогательной плазмиды. Загрязнение аденовирусом можно уменьшить, например, термической обработкой, к которой аденовирус более чувствителен, чем AAV. Можно использовать и другие способы доставки нуклеиновых кислот в клетки, к примеру, как описано в US 2003/0087817, включенном сюда путем ссылки.

Клетки хозяина могут подвергаться кратковременной или некротической трансфекции одним или несколькими векторами, описанными здесь. Клетки можно трансфицировать в том виде, в каком они естественным образом встречаются у субъекта. Клетки можно взять или получить от субъекта и трансфицировать. Клетки могут происходить из клеток, взятых у субъекта, типа линии клеток. В некоторых воплощениях клетки, трансфицированные одним или несколькими векторами, описанными здесь, используются для установления новой клеточной линии, содержащей одну или несколько полученных из вектора последовательностей. В некоторых воплощениях клетки, подвергнутые кратковременной трансфекции композициями по изобретению (типа кратковременной трансфекции одним или несколькими векторами или трансфекции РНК) и модифицированные под действием исполнительного элемента типа комплекса CRISPR, применяются для создания новой линии клеток, содержащей клетки с модификацией, но без каких-либо других экзогенных последовательностей.

В способах по изобретению может использоваться любой подходящий вектор, совместимый с клетками хозяина. Неограничивающие примеры векторов для эукариотических клеток-хозяев включают pXT1, pSG5 (Stratagene™), pSVK3, pBPV, pMSG и pSVLSV40 (Pharmacia™).

В некоторых воплощениях последовательность нуклеотидов, кодирующая направляющую нуклеиновую кислоту и/или белок Cas или химеру, функционально связана с контрольным элементом, например, контрольным элементом транскрипции типа промотора. Контрольный элемент транскрипции может функционировать как в эукариотических клетках, например клетках млекопитающих, так и в прокариотических клетках (например, бактериальных или археальных клетках). В некоторых воплощениях последовательность нуклеотидов, кодирующая направляющую нуклеиновую кислоту и/или белок Cas или химеру, функционально связана с несколькими контрольными элементами, которые способствуют экспрессированию последовательности нуклеотидов, кодирующей направляющую нуклеиновую кислоту и/или белок Cas или химеру, в прокариотических и/или эукариотических клетках.

В зависимости от используемой системы хозяин/вектор, в экспрессирующем векторе может использоваться любой из целого ряда подходящих контрольных элементов транскрипции и трансляции, включая конститутивные и индуцибельные промоторы, элементы-энхансеры транскрипции, терминаторы транскрипции и т.д. (например, промотор U6, промотор H1 и т.п., см. выше) (например, см. Bitter et al. (1987) *Methods in Enzymology*, 153: 516-544).

В некоторых воплощениях композиции по изобретению (например, исполнительный элемент типа белка Cas или химеры Cas, химерный рецептор, химерный адаптер, направляющая нуклеиновая кислота и т.п.) могут быть представлены в виде РНК. В таких случаях композиции по изобретению (например, исполнительный элемент типа белка Cas или химеры Cas, химерный рецептор, химерный адаптер, направляющая нуклеиновая кислота и т.п.) могут быть получены прямым химическим синтезом или могут транскрибироваться *in vitro* из ДНК. Композиции по изобретению (например, исполнительный элемент типа белка Cas или химеры Cas, химерный рецептор, химерный адаптер, направляющая нуклеиновая кислота и т.п.) можно синтезировать *in vitro* с помощью фермента РНК-полимеразы (например, полимеразы T7, полимеразы T3, полимеразы SP6 и др.). После синтеза РНК может непосредственно контактировать с целевой ДНК или может быть введена в клетки любым подходящим методом введения нуклеиновых кислот в клетки (например, микроинъекции, электропорации, трансфекции и др.).

Нуклеотиды, кодирующие направляющую нуклеиновую кислоту (вводится в виде ДНК либо РНК) и/или белок Cas или химеру (вводятся в виде ДНК либо РНК), можно вводить в клетки, используя под-

ходящий метод трансфекции; например, см. Angel and Yanik (2010) PLoS One 5(7): e11756; и коммерчески доступные реагенты TransMessenger® фирмы Qiagen, набор Stemfect™ для трансфекции РНК фирмы Stement или набор TransIT® для трансфекции мРНК фирмы Mirus Bio LLC. Также см. Beumer et al. (2008) Efficient gene targeting in Drosophila by direct embryo injection with zinc-finger nucleases. PNAS 105(50): 19821-19826. Нуклеиновые кислоты, кодирующие композиции по изобретению (например, исполнительный элемент типа белка Cas или химеры Cas, химерный рецептор, адаптер, направляющую нуклеиновую кислоту и т.п.), могут быть представлены на ДНК-векторах или олигонуклеотидах. Существует много векторов, например плазмиды, космиды, миникольца, фаги, вирусы и т.д., пригодных для введения нуклеиновых кислот в клетки мишени. Векторы, содержащие нуклеиновые кислоты, могут содержаться в виде эписом, например в виде плазмид, космид, миникольцевой ДНК, вирусов типа цитомегаловируса, аденовируса и др., или же они могут встраиваться в геном клеток мишени посредством гомологической рекомбинации или случайного встраивания, например, ретровирусные векторы типа MMLV, HIV-1 и ALV.

Исполнительный элемент типа белка или химеры Cas, химерный рецептор и/или адаптер можно вводить в клетки в виде полипептида. Такой белок необязательно может быть слит с полипептидным доменом, повышающим растворимость продукта. Домен может быть связан с полипептидом через сайт расщепления определенной протеазы, например, последовательность TEV, которая расщепляется протеазой TEV. Линкер также может включать одну или несколько гибких последовательностей, например, от 1 до 10 остатков глицина. В некоторых воплощениях расщепление слитого белка проводится в буфере, который поддерживает растворимость продукта, например, в присутствии от 0,5 до 2М мочевины, в присутствии полипептидов и/или полинуклеотидов, повышающих растворимость, и т.п. Представляющие интерес домены включают эндосомолитические домены, например, домен НА вируса гриппа; и другие полипептиды, которые способствуют получению, например, домен IF2, домен GST, домен GRPE и др. Полипептид может быть сформулирован для улучшения стабильности. Например, пептиды можно ПЭ-Гилировать, при этом полиэтиленоксигруппа обеспечивает повышение срока жизни в кровотоке.

Композиции по изобретению (например, исполнительный элемент типа Cas-белка или Cas-химеры, химерный рецептор, химерный адаптер, направляющая нуклеиновая кислота и т.п.) можно сливать с проникающим доменом полипептида, чтобы усилить захват клетками. В неинтегрирующих полипептидах настоящего изобретения могут использоваться многие проникающие домены, включая пептиды, пептидомиметики и непептидные носители. К примеру, проникающий пептид может происходить из третьей α -спирали фактора транскрипции Antennapedia из *Drosophila melanogaster*, называемого пенетратином, который содержит аминокислотную последовательность RQKIWFQNRMKWKK (SEQ ID NO: 85). В качестве другого примера проникающий пептид может содержать аминокислотную последовательность основного участка tat HIV-1, которая может включать, к примеру, аминокислоты 49-57 природного белка tat. Другие проникающие домены включают полиаргининовые мотивы, к примеру, участок аминокислот 34-56 белка rev HIV-1, нонаргинин (SEQ ID NO: 88), октааргинин (SEQ ID NO: 90) и т.п. (например, см. Futaki et al. (2003) Curr Protein Pept, 2003 April, 4(2): 87-9 and 446; Wender et al. (2000) Proc. Natl. Acad. Sci. USA 2000 Nov. 21, 97(24): 13003-8; Published U.S. Patent Applications 2003/0220334; 2003/0083256; 2003/0032593, специально включенные сюда путем ссылки, на счет сведений о транслокации пептидов и пептоидов). Можно использовать последовательность нонаргинина (R9) (SEQ ID NO: 88) (Wender et al., 2000; Uemura et al., 2002). Участок, на котором проводится слияние, может быть выбран так, чтобы оптимизировать биологическую активность, секрецию или характеристики связывания полипептида.

Композиции по изобретению (например, исполнительный элемент типа белка Cas или химеры Cas, химерный рецептор, адаптер, направляющая нуклеиновая кислота и т.п.) могут быть получены *in vitro* либо вырабатываться эукариотическими клетками или прокариотическими клетками, причем они могут подвергаться дополнительной обработке путем разворачивания, например тепловой денатурацией, восстановлением DTT и др., а также могут быть свернуты опять.

Композиции по изобретению (например, исполнительный элемент типа белка Cas или химеры Cas, химерный рецептор, адаптер, направляющая нуклеиновая кислота и т.п.) могут быть получены путем синтеза *in vitro*. Можно использовать различные коммерческие установки для синтеза, к примеру автоматизированные синтезаторы фирмы Applied Biosystems, Inc., Beckman и др. При помощи синтезаторов можно заменять природные аминокислоты на неприродные аминокислоты. Конкретную последовательность и способ получения можно определить по удобству, экономичности, требуемой чистоте и пр.

Композиции по изобретению (например, исполнительный элемент типа белка Cas или химеры Cas, химерный рецептор, адаптер, направляющую нуклеиновую кислоту и т.п.) также можно выделять и очищать стандартными методами рекомбинантного синтеза. Можно приготовить лизат из экспрессирующего хозяина и очистить его методом HPLC, эксклюзионной хроматографии, гель-электрофореза, аффинной хроматографии либо иным методом очистки. Композиции могут содержать, к примеру, по меньшей мере 20% мас. требуемого продукта, по меньшей мере 75% мас., по меньшей мере 95% мас., а в терапевтических целях, к примеру, по меньшей мере 99,5% мас. по отношению к примесям, связанным с методом получения продукта и его очисткой. Проценты могут основываться на общем белке.

Композиции по изобретению (например, исполнительный элемент типа белка Cas или химеры Cas, химерный рецептор, адаптер, направляющую нуклеиновую кислоту и т.п.) при введении в виде нуклеиновых кислот или полипептидов можно предоставлять клеткам на время от 30 мин до 24 ч, например 1, 1,5, 2, 2,5, 3, 3,5 4, 5, 6, 7, 8, 12, 16, 18, 20 ч или на любой другой период от 30 мин до 24 ч, который можно повторять с частотой от каждого дня до каждых 4 дней, например каждые 1,5 дня, каждые 2 дня, каждые 3 дня или с любой другой частотой от каждого дня до 4 дней. Композиции можно предоставлять данным клеткам один или несколько раз, например один раз, два раза, три раза или более трех раз, а клетки можно инкубировать с агентами в течение некоторого промежутка времени после каждого контакта, например, 16-24 ч, после чего можно заменить среду свежей средой и культивировать клетки дальше.

В тех случаях, когда клеткам предоставляется два или несколько различных наводящих комплексов (например, две разные гидовые нуклеиновые кислоты, которые комплементарны различным последовательностям в одной и той же или в разных целевых ДНК), комплексы можно предоставлять одновременно (например, в виде двух полипептидов и/или нуклеиновых кислот) или добавлять одновременно. С другой стороны, их можно добавлять последовательно, например, сначала добавить первый нацеливающий комплекс, а затем второй нацеливающий комплекс и т.д. или наоборот.

К целевой ДНК или клеткам можно подавать эффективное количество композиций по изобретению (например, исполнительного элемента типа белка Cas или химеры Cas, химерного рецептора, адаптера, направляющей нуклеиновой кислоты и т.п.). Эффективным количеством может быть такое количество, которое вызывает, к примеру, по меньшей мере 2-кратное или большее изменение (повышение или уменьшение) степени регуляции мишени, отмечаемое между двумя гомологичными последовательностями относительно отрицательного контроля, например, клеток, обработанных пустым вектором или посторонним полипептидом. Эффективное количество или доза могут вызывать, к примеру, 2-кратное изменение, 3-кратное изменение, 4-кратное изменение, 7-кратное, 8-кратное, 10-кратное, 50-кратное, 100-кратное, 200-кратное, 500-кратное, 700-кратное, 1000-кратное, 5000-кратное или 10000-кратное изменение регуляции целевого гена. Степень регуляции целевого гена можно измерить любым подходящим способом.

Контакт клеток с композицией по изобретению может происходить в любой среде для культивирования и при любых условиях культивирования, которые способствуют выживанию клеток. Например, клетки могут быть суспендированы в любой подходящей питательной среде, которая будет удобной, типа модифицированной Iscove среды DMEM или RPMI 1640 с добавлением фетальной телячьей сыворотки или инактивированной нагреванием козьей сыворотки (5-10%), L-глутамина, тиола, в частности 2-меркаптоэтанол, и антибиотиков, например пенициллина и стрептомицина. Культура может содержать такие факторы роста, к которым клетки восприимчивы. Факторами роста, как определено здесь, являются молекулы, способные стимулировать выживание, рост и/или дифференцировку клеток, либо в культуре, либо в интактной ткани, посредством специфических эффектов на трансмембранные рецепторы. Факторы роста могут включать полипептиды и неполипептидные факторы.

В многочисленных воплощениях выбранная система доставки нацелена на определенные типы тканей или клеток. В некоторых случаях наводка системы доставки на ткани или клетки осуществляется путем связывания системы доставки со специфичными маркерами тканей или клеток типа поверхностных белков клетки. Вирусные и невирусные системы доставки могут быть адаптированы к целевой ткани или представляющим интерес типам клеток.

Фармацевтические композиции, содержащие описанные здесь молекулы (например, полипептиды и/или нуклеиновые кислоты, кодирующие полипептиды) или иммунные клетки, можно вводить для профилактического и/или терапевтического лечения. В терапевтических применениях композиции можно вводить субъектам, которые уже страдают заболеванием, в количестве, достаточном для излечения или по крайней мере частичной остановки симптомов заболевания или же для лечения, излечения, улучшения или ослабления заболевания. Эффективные количества для такого применения могут варьироваться в зависимости от тяжести и течения заболевания, предшествующей терапии, состояния здоровья пациента, веса и реакции на лекарства, а также суждения лечащего врача.

Можно вводить несколько терапевтических средств в любом порядке или одновременно. Если одновременно, то несколько терапевтических средств можно представить в единой унифицированной форме или в нескольких формах, к примеру, в виде нескольких отдельных таблеток. Молекулы могут быть упакованы вместе или по отдельности, в одной упаковке или в нескольких упаковках. Одно или все терапевтические средства можно назначать в нескольких дозах. Если не одновременно, то промежуток времени между несколькими дозами может составлять даже целый месяц.

Описанные здесь молекулы можно вводить до, во время или после возникновения заболевания, а сроки введения композиций, содержащих соединения, могут варьироваться. Например, фармацевтические композиции можно использовать в качестве профилактики и вводить непрерывно субъектам со склонностью к заболеваниям, чтобы предотвратить возникновение заболевания. Молекулы и фармацевтические композиции можно вводить субъекту во время или как можно скорее после появления симптомов. Введение молекул можно начинать в пределах первых 48 ч после появления симптомов, в пределах

первых 24 ч после появления симптомов, в пределах первых 6 ч после появления симптомов или в пределах первых 3 ч после появления симптомов. Первоначальное введение может осуществляться практически любым способом, как-то любым описанным здесь способом, используя любые описанные здесь лекарственные формы. Молекулы можно вводить, как только это будет практически возможно после явного или предполагаемого возникновения заболевания и на протяжении времени, необходимого для лечения заболевания, как, к примеру, от 1 до 3 месяцев. Продолжительность лечения может варьироваться для каждого пациента.

Молекулы могут быть упакованы в биологический компартмент. Биологический компартмент, содержащий молекулы, можно вводить субъектам. К биологическим компартментам могут относиться, без ограничения, вирусы (лентивирусы, аденовирусы), наносферы, липосомы, квантовые точки, наночастицы, микрочастицы, нанокапсулы, везикулы, частицы полиэтиленгликоля, гидрогели и мицеллы.

Например, биологический компартмент может включать липосомы. Липосома может быть самосборочной структурой, содержащей один или несколько липидных бислоев, каждый из которых может содержать два монослоя, содержащих противоположно ориентированные молекулы амфипатических липидов. Амфипатические липиды могут содержать полярную (гидрофильную) головку, ковалентно связанную с одной или двумя или несколькими неполярными (гидрофобными) ацильными или алкильными цепями. Энергетически неблагоприятные контакты между гидрофобными ацильными цепями и окружающей водной средой вынуждают молекулы амфипатических липидов устроиться таким образом, чтобы полярные головки были ориентированы на поверхность бислоя, а ацильные цепи ориентированы вовнутрь бислоя, эффективно ограждая ацильные цепи от контакта с водной средой.

Примеры предпочтительных амфипатических соединений, используемых в липосомах, включают фосфоглицериды и сфинголипиды, типичные примеры которых включают фосфатидилхолин, фосфатидилэтаноламин, фосфатидилсерин, фосфатидил-инозитол, фосфатидную кислоту, фосфатидилглицерин, пальмитоилолеоил-фосфатидил-холин, лизофосфатидилхолин, лизофосфатидилэтаноламин, димиристоилфосфатидил-холин (DMPC), дипальмитоилфосфатидилхолин (DPPC), диолеоилфосфатидилхолин, дистеароилфосфатидилхолин (DSPC), дилиолеоилфосфатидилхолин и яичный сфинго-миелин или любые их комбинации.

Биологический компартмент может включать наночастицы. Наночастицы могут иметь диаметр от 40 нанометров (нм) до 1,5 микрометров (мкм), от 50 нм до 1,2 мкм, от 60 нм до 1 мкм, от 70 нм до 800 нм, от 80 нм до 600 нм, от 90 нм до 400 нм, от 100 нм до 200 нм.

В некоторых случаях с увеличением размера наночастиц может замедляться или увеличиваться скорость высвобождения, а с уменьшением размера наночастиц скорость высвобождения может повышаться.

Содержание альбумина в наночастицах может составлять от 5% до 85% (об./об.), от 10% до 80%, от 15% до 80%, от 20% до 70% альбумина (об./об.), от 25% до 60%, от 30% до 50% или от 35% до 40%. Фармацевтическая композиция может содержать до 30, 40, 50, 60, 70 или 80% наночастиц или больше. В некоторых случаях молекулы нуклеиновой кислоты по изобретению могут быть связаны с поверхностью наночастиц.

Биологический компартмент может включать вирусы. Вирусы могут служить системой доставки для фармацевтических композиций изобретения. Типичные вирусы включают лентивирусы, ретровирусы, аденовирусы, вирусы простого герпеса I или II, парвовирусы, вирусы ретикулоэндотелиоза и аденоассоциированные вирусы (AAV). Фармацевтические композиции по изобретению можно доставлять в клетки с помощью вирусов. Вирусом можно инфицировать и трансдуцировать клетки *in vivo*, *ex vivo* или *in vitro*. При доставке *ex vivo* и *in vitro* нуждающемуся в терапии субъекту можно вводить трансдуцированные клетки.

Фармацевтические композиции могут быть упакованы в системы вирусной доставки. Например, композиции могут быть упакованы в вирионы с помощью упаковочной системы, свободной от хелперного вируса HSV-1.

Вирусные системы доставки (например, вирусы, содержащие фармацевтические композиции изобретения) можно вводить путем прямой инъекции, стереотаксической инъекции, в желудочки мозга, при помощи инфузионных миниасосов, посредством конвекции, через катетеры, путем внутривенной, парентеральной, внутрибрюшинной и/или подкожной инъекции, в клетки, ткани или органы нуждающихся в этом субъектов. В некоторых случаях клетки могут быть трансдуцированы *in vitro* или *ex vivo* с помощью вирусных систем доставки. Трансдуцированные клетки можно вводить заболевшему субъекту. Например, можно трансдуцировать стволовые клетки с помощью системы вирусной доставки, содержащей фармацевтическую композицию, а затем эти стволовые клетки можно имплантировать пациенту для лечения заболевания. В некоторых случаях доза трансдуцированных клеток при введении субъекту может составлять 1×10^5 клеток/кг, 5×10^5 клеток/кг, 1×10^6 клеток/кг, 2×10^6 клеток/кг, 3×10^6 клеток/кг, 4×10^6 клеток/кг, 5×10^6 клеток/кг, 6×10^6 клеток/кг, 7×10^6 клеток/кг, 8×10^6 клеток/кг, 9×10^6 клеток/кг, 1×10^7 клеток/кг, 5×10^7 клеток/кг, 1×10^8 клеток/кг или больше в одной разовой дозе.

Введение биологических компартментов в клетки может происходить путем инфицирования виру-

сом или бактериофагом, трансфекции, конъюгации, слияния протопластов, липофекции, электропорации, осаждения фосфатом кальция, трансфекции с помощью полиэтиленimina (PEI), трансфекции с помощью DEAE-декстрана, трансфекции при помощи липосом, технологии пушки для частиц, осаждения фосфатом кальция, прямой микроинъекции, доставки нуклеиновой кислоты при помощи наночастиц и т.д.

Описанные здесь молекулы (например, полипептиды и/или нуклеиновые кислоты) могут присутствовать в композиции в пределах от 1 мг до 2000 мг; от 5 мг до 1000 мг, от 10 мг до 25 мг до 500 мг, от 50 мг до 250 мг, от 100 мг до 200 мг, от 1 мг до 50 мг, от 50 мг до 100 мг, от 100 мг до 150 мг, от 150 мг до 200 мг, от 200 мг до 250 мг, от 250 мг до 300 мг, от 300 мг до 350 мг, от 350 мг до 400 мг, от 400 мг до 450 мг, от 450 мг до 500 мг, от 500 мг до 550 мг, от 550 мг до 600 мг, от 600 мг до 650 мг, от 650 мг до 700 мг, от 700 мг до 750 мг, от 750 мг до 800 мг, от 800 мг до 850 мг, от 850 мг до 900 мг, от 900 мг до 950 мг или от 950 мг до 1000 мг.

Описанные здесь молекулы (например, полипептиды и/или нуклеиновые кислоты) могут присутствовать в композиции в количестве примерно 1 мг, 2 мг, 3 мг, 4 мг, 5 мг, 10 мг, 15 мг, 20 мг, 25 мг, 30 мг, 35 мг, 40 мг, 45 мг, 50 мг, 55 мг, 60 мг, 65 мг, 70 мг, 75 мг, 80 мг, 85 мг, 90 мг, 95 мг, 100 мг, 125 мг, 150 мг, 175 мг, 200 мг, 250 мг, 300 мг, 350 мг, 400 мг, 450 мг, 500 мг, 550 мг, 600 мг, 650 мг, 700 мг, 750 мг, 800 мг, 850 мг, 900 мг, 950 мг, 1000 мг, 1050 мг, 1100 мг, 1150 мг, 1200 мг, 1250 мг, 1300 мг, 1350 мг, 1400 мг, 1450 мг, 1500 мг, 1550 мг, 1600 мг, 1650 мг, 1700 мг, 1750 мг, 1800 мг, 1850 мг, 1900 мг, 1950 мг или 2000 мг.

Описанные здесь молекулы (например, полипептиды и/или нуклеиновые кислоты) могут присутствовать в композиции так, чтобы получилось по меньшей мере 0,1, 0,5, 1, 1,5, 2, 2,5, 3, 3,5, 4, 4,5, 5, 5,5, 6, 6,5, 10 или больше единиц активности на 1 мг молекулы. Активность может представлять собой регуляцию экспрессии генов. В некоторых воплощениях общее количество единиц активности при введении молекул субъекту составляет по меньшей мере 25000, 30000, 35000, 40000, 45000, 50000, 60000, 70000, 80000, 90000, 110000, 120000, 130000, 140000, 150000, 160000, 170000, 180000, 190000, 200000, 210000, 220000, 230000, 250000 или больше единиц. В некоторых воплощениях общее количество единиц активности при введении молекул субъекту составляет не более 25000, 30000, 35000, 40000, 45000, 50000, 60000, 70000, 80000, 90000, 110000, 120000, 130000, 140000, 150000, 160000, 170000, 180000, 190000, 200000, 210000, 220000, 230000, 250000 или больше единиц.

В некоторых воплощениях субъекту вводится по меньшей мере 10000 единиц активности из расчета на 50 кг массы тела. В некоторых воплощениях субъекту вводится по меньшей мере 10000, 15000, 25000, 30000, 35000, 40000, 45000, 50000, 60000, 70000, 80000, 90000, 110000, 120000, 130000, 140000, 150000, 160000, 170000, 180000, 190000, 200000, 210000, 220000, 230000, 250000 или больше единиц активности молекулы из расчета на 50 кг массы тела. В некоторых воплощениях терапевтически эффективная доза составляет по меньшей мере 5×10^5 , 1×10^6 , 2×10^6 , 3×10^6 , 4×10^6 , 5×10^6 , 6×10^6 , 7×10^6 , 8×10^6 , 9×10^6 , 1×10^7 , $1,1 \times 10^7$, $1,2 \times 10^7$, $1,5 \times 10^7$, $1,6 \times 10^7$, $1,7 \times 10^7$, $1,8 \times 10^7$, $1,9 \times 10^7$, 2×10^7 , $2,1 \times 10^7$, 3×10^7 или больше единиц активности молекулы. В некоторых воплощениях терапевтически эффективная доза составляет не более 5×10^5 , 1×10^6 , 2×10^6 , 3×10^6 , 4×10^6 , 5×10^6 , 6×10^6 , 7×10^6 , 8×10^6 , 9×10^6 , 1×10^7 , $1,1 \times 10^7$, $1,2 \times 10^7$, $1,5 \times 10^7$, $1,6 \times 10^7$, $1,7 \times 10^7$, $1,8 \times 10^7$, $1,9 \times 10^7$, 2×10^7 , $2,1 \times 10^7$, 3×10^7 или больше единиц активности молекулы.

В некоторых воплощениях терапевтически эффективная доза составляет по меньшей мере 10000, 15000, 20000, 22000, 24000, 25000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 125000, 150000, 200000 или 500000 единиц на 1 кг массы тела. В некоторых воплощениях терапевтически эффективная доза составляет не более 10000, 15000, 20000, 22000, 24000, 25000, 30000, 40000, 50000, 60000, 70000, 80000, 90000, 100000, 125000, 150000, 200000 или 500000 единиц на 1 кг массы тела.

В некоторых воплощениях активность молекулы при введении субъекту составляет 10000, 11000, 12000, 13000, 14000, 20000, 21000, 22000, 23000, 24000, 25000, 26000, 27000, 28000, 30000, 32000, 34000, 35000, 36000, 37000, 40000, 45000, 50000 или больше ед./мг молекулы. В некоторых воплощениях активность молекулы при введении субъекту составляет не более 10000, 11000, 12000, 13000, 14000, 20000, 21000, 22000, 23000, 24000, 25000, 26000, 27000, 28000, 30000, 32000, 34000, 35000, 36000, 37000, 40000, 45000, 50000 или больше ед./мг молекулы.

В различных воплощениях изложенных здесь аспектов могут быть получены фармакокинетические и фармакодинамические данные. Существуют различные экспериментальные методы получения таких данных. Соответствующие компоненты фармакокинетического и фармакодинамического профиля, описывающие определенную композицию, могут варьироваться вследствие вариаций в метаболизме препарата у людей. Фармакокинетические и фармакодинамические профили могут основываться на определении средних параметров у группы испытуемых. Группа субъектов включает любое разумное количество субъектов, подходящее для определения репрезентативного среднего, к примеру, 5 субъектов, 10 субъектов, 15 субъектов, 20 субъектов, 25 субъектов, 30 субъектов, 35 субъектов или больше. Среднее значение определяется путем вычисления среднего значения по всем измерениям субъекта для каждого измеряемого параметра. Для получения требуемого фармакокинетического или фармакодинамического профиля типа желательного или эффективного профиля крови можно модулировать дозу, как описано здесь.

Фармакокинетические параметры могут быть любые параметры, подходящие для описания молекул. Например, C_{\max} может составлять, к примеру, не менее 25 нг/мл; не менее 50 нг/мл; не менее 75 нг/мл; не менее 100 нг/мл; не менее 200 нг/мл; не менее 300 нг/мл; не менее 400 нг/мл; не менее 500 нг/мл; не менее 600 нг/мл; не менее 700 нг/мл; не менее 800 нг/мл; не менее 900 нг/мл; не менее 1000 нг/мл; не менее 1250 нг/мл; не менее 1500 нг/мл; не менее 1750 нг/мл; не менее 2000 нг/мл; или любое другое значение C_{\max} , подходящее для описания фармакокинетического профиля описанных здесь молекул.

Значение T_{\max} у описанных здесь молекул может составлять, к примеру, не более 0,5 ч, не более 1 ч, не более 1,5 ч, не более 2 ч, не более 2,5 ч, не более 3 ч, не более 3,5 ч, не более 4 ч, не более 4,5 ч, не более 5 ч или любое другое значение T_{\max} , подходящее для описания фармакокинетического профиля описанных здесь молекул.

Значение $AUC_{(0-\text{inf})}$ у описанных здесь молекул может составлять, к примеру, не менее 50 нгч/мл, не менее 100 нгч/мл, не менее 150 нгч/мл, не менее 200 нгч/мл, не менее 250 нгч/мл, не менее 300 нгч/мл, не менее 350 нгч/мл, не менее 400 нгч/мл, не менее 450 нгч/мл, не менее 500 нгч/мл, не менее 600 нгч/мл, не менее 700 нгч/мл, не менее 800 нгч/мл, не менее 900 нгч/мл, не менее 1000 нгч/мл, не менее 1250 нгч/мл, не менее 1500 нгч/мл, не менее 1750 нгч/мл, не менее 2000 нгч/мл, не менее 2500 нгч/мл, не менее 3000 нгч/мл, не менее 3500 нгч/мл, не менее 4000 нгч/мл, не менее 5000 нгч/мл, не менее 6000 нгч/мл, не менее 7000 нгч/мл, не менее 8000 нгч/мл, не менее 9000 нгч/мл, не менее 10000 нгч/мл или любое другое значение $AUC_{(0-\text{inf})}$, подходящее для описания фармакокинетического профиля описанных здесь молекул.

Концентрация в плазме описанных здесь молекул через 1 час после введения может составлять, к примеру, не менее 25 нг/мл, не менее 50 нг/мл, не менее 75 нг/мл, не менее 100 нг/мл, не менее 150 нг/мл, не менее 200 нг/мл, не менее 300 нг/мл, не менее 400 нг/мл, не менее 500 нг/мл, не менее 600 нг/мл, не менее 700 нг/мл, не менее 800 нг/мл, не менее 900 нг/мл, не менее 1000 нг/мл, не менее 1200 нг/мл или любое другое значение концентрации в плазме описанных здесь молекул.

Фармакокинетические параметры могут быть любые параметры, подходящие для описания фармацевтических композиций изобретения. Например, фармакодинамический профиль может проявлять снижение факторов, связанных с воспалением, к примеру, через 2, 4, 8, 12 или 24 ч.

В различных воплощениях изложенных здесь аспектов способы по изобретению выполняются на субъектах. Субъектами могут быть люди. Субъектами могут быть млекопитающие (например, крысы, мыши, коровы, собаки, свиньи, овцы, лошади). Субъектами могут быть позвоночные или беспозвоночные. Субъектами могут быть лабораторные животные. Субъектами могут быть пациенты. Субъекты могут страдать заболеванием. Субъект может проявлять симптомы заболевания. Субъект может не проявлять симптомы заболевания, но все же иметь заболевание. Субъект может находиться на медицинском обслуживании у провайдера (например, субъект госпитализирован и проходит лечение у врача). Субъектами могут быть растения или сельскохозяйственные культуры.

Примеры

Следующие примеры приводятся для иллюстрации различных воплощений изобретения и не должны ограничивать настоящее изобретение никоим образом. Настоящие примеры, наряду с описанными здесь способами, сейчас являются репрезентативными для предпочтительных воплощений, являются показательными и не предназначаются для ограничения объема изобретения.

Пример 1. Сконструированные рекомбинантные химерные рецепторы, пришитые к эффекторным белкам, которые могут модулировать экспрессию генов посредством редактирования генома или регуляции транскрипции.

На фиг. 17 представлен сконструированный химерный антигеновый рецептор, содержащий домен, модулирующий ген. Сконструированный искусственный рекомбинантный рецептор включает, в линейном порядке, внеклеточный домен (ECD), трансмембранный домен (TM) и внутриклеточный домен (ICD). ECD обладает активностью специфического связывания лиганда. TM пронизывает клеточную мембрану, а ICD обладает функцией манипуляции генома. TM и ICD соединяются последовательностью пептидного линкера, а последовательность пептида может распознаваться протеазой. Рецептор экспрессируется на клетке и может связывать свой лиганд 1701. При связывании лиганда адаптерный белок, пришитый (слитый) к протеазе, привлекается к рецептору посредством белок-белкового взаимодействия, кластеризации или опосредованного каркасом взаимодействия. После того, как адаптер-протеаза ассоциируется со сконструированным химерным антигеновым рецептором, протеаза высвобождает домен, модулирующий ген, из рецептора 1705. Затем домен, модулирующий ген, может свободно перемещаться в ядро для регулирования или редактирования генов-мишеней 1709. ICD может включать в себя фактор транскрипции и модулировать экспрессию генов или эпигенетику (Nu, ядро).

Рецептор может быть слит с доменом, модулирующим ген, через линкер (фиг. 18). Линкер может располагаться между трансмембранным доменом и доменом, модулирующим ген, (внутриклеточным исполнительным доменом) сконструированного рецептора. Линкер может содержать один или несколько сигналов экспорта эндоплазматического ретикулула, одну или несколько последовательностей пептидных линкеров, один или несколько доменов фолдинга пептидов, один или несколько расщепляющих доменов протеаз, один или несколько доменов клеточной локализации например, сигнал ядерной локализа-

ции или сигнал митохондриальной локализации, или же любые их комбинации. В некоторых случаях компоненты линкера могут располагаться в любом порядке.

Сконструированный химерный антигеновый рецептор может быть разработан так, чтобы он связывал поверхностные антигены клетки (фиг. 19А). В некоторых случаях сконструированный химерный антигеновый рецептор и адаптер-протеаза располагаются на клеточной поверхности принимающей клетки, а лиганд находится на поверхности сигнализирующей клетки. Рецептор может быть сконструирован так, чтобы он связывался с растворимыми антигенами, находящимися в локальном окружении клетки (фиг. 19В). С другой стороны, сконструированный химерный антигеновый рецептор может связываться с сигнальными молекулами или лигандами внеклеточного матрикса (ЕСМ) (фиг. 19С). В других воплощениях сконструированный химерный антигеновый рецептор может димеризоваться со взаимодействующим рецептором, связанным с цитозольной протеазой (фиг. 19D). В других случаях взаимодействующий рецептор не пришит к протеазе, а связанный с лигандом рецептор может рекрутировать полипептид адаптера-протеазы (фиг. 19Е). В других воплощениях химерные рецепторы образуют комплекс с другими рецепторами (природными рецепторами, эндогенными рецепторами или синтетическими рецепторами), которые находятся в одной и той же клетке (фиг. 19F).

Полипептиды домена, модулирующего ген, сконструированного химерного антигенового рецептора могут быть основаны на таких рецепторах, как Notch, GPCRs, интегрин, кадгерин, рецепторы смерти и химерные антигеновые рецепторы (фиг. 20А). Эти полипептиды могут экспрессироваться вместе со слитым белком адаптер-протеаза типа пресенилин-протеаза, β -аррестин-протеаза, паксиллин-протеаза, β -катенин-протеаза или FADD-протеаза. Если домен, модулирующий ген, содержит белок CRISPR типа Cas9 или dCas9, то этот домен может связываться с направляющей РНК (gRNA, например, sgRNA). Протеаза может высвобождать комплекс домен, модулирующий ген,-sgRNA, который может транслоцироваться в ядро и связываться с последовательностью ДНК, комплементарной sgRNA (фиг. 20В). Последовательность ДНК может находиться в регуляторном участке или промоторе целевого гена.

Пример 2. Сконструированные рекомбинантные химерные рецепторы на основе GPCR, интегрина и Notch.

В этом примере описаны три класса сконструированных систем рекомбинантных химерных рецепторов. Эти рецепторы включают сопряженные с G-белком рецепторы (GPCRs), интегрин и Notch, которые естественным образом детектируют различные типы лигандов и сигналов, относящихся к микроокружению рака. Например, GPCRs, известные как рецепторы из семи трансмембранных доменов, рецепторы 7ТМ, семиспиральные рецепторы, серпантинные рецепторы и связанные с G-белками рецепторы (GPLR), составляют большое семейство белковых рецепторов, которые воспринимают молекулы вне клетки и активируют пути передачи сигналов, а в конечном счете клеточные реакции внутри клетки.

A1. Рекомбинантные химерные GPCR, содержащие dCas9.

Сопряженные с G-белком рецепторы встречаются только у эукариот. Лиганды, которые связываются и активируют эти рецепторы, очень разнообразны и включают светочувствительные соединения, запахи, феромоны, гормоны и нейротрансмиттеры. Лиганды GPCR варьируются от небольших молекул и пептидов до крупных белков.

Примечательно, что сопряженные с G-белком рецепторы участвуют во многих заболеваниях, а также являются мишенью приблизительно 40% всех современных лекарственных препаратов. Главные естественные пути передачи сигналов с участием сопряженных с G-белком рецепторов могут быть сложными, включая либо сигнальный путь цАМФ, либо сигнальный путь фосфатидилинозитола.

Когда лиганд связывается с GPCR, он может вызвать конформационное изменение в GPCR, что позволяет ему действовать как фактор обмена гуаниновых нуклеотидов (GEF). GPCR может затем активировать ассоциированный G-белок путем обмена связанного с ним ГДФ на ГТФ. Затем α -субъединица G-белка вместе со связанным ГТФ может диссоциировать от β - и γ -субъединиц и далее будет влиять на внутриклеточные сигнальные белки или непосредственно воздействовать на функциональные белки в зависимости от типа α -субъединицы (G α s, G α i/o, G α q/11, G α 12/13).

Важно отметить, что GPCRs вовлекаются в широкий спектр физиологических процессов. Некоторые примеры их физиологической роли включают: зрительное чувство, вкусовое чувство (вкус), обоняние, регуляцию поведения и настроения, передачу сигналов вегетативной нервной системы и модуляцию гомеостаза. Среди многих GPCR три класса особенно важны для клеточной иммунотерапии рака. Первый класс GPCRs может регулировать активность иммунной системы и воспаление. Например, хемокиновые GPCR-рецепторы связываются с лигандами, которые опосредуют межклеточную коммуникацию между клетками иммунной системы. Гистаминовые GPCR-рецепторы связывают воспалительные медиаторы и активируют различные типы клеток мишени при воспалительном ответе. Toll-подобные GPCR-рецепторы (TLR) задействованы в иммуномодуляции и непосредственно участвуют в подавлении иммунных реакций Т-клеток, что тесно связано с недавними открытиями того, что PDL-1 в раковых клетках может подавлять Т-клетки посредством взаимодействия с рецепторами PD-1 на Т-клетках. Второй класс GPCRs может модулировать восприятие плотности клеток. Недостаточность контактного ингибирования является отличительной чертой инвазивных раковых клеток. Имеются данные о том, что раковые клетки

могут манипулировать восприятием собственной плотности клеток так, что раковые клетки могут расти до высокой плотности и затем блокировать функцию иммунных клеток. Возможность дешифровки реакции восприятия плотности с помощью искусственных Т-клеток должна входить в схему лечения раковых клеток. Третий класс GPCRs участвует в росте и метастазировании некоторых типов опухолей.

Для создания синтетических GPCRs, способных усилить распознавание Т-клетками ракового микроокружения, сначала тестировали два GPCR: рецептор хемокинов С-Х-С 4-го типа (CXCR4) и рецептор лизофосфатидной кислоты (LPA). CXCR4 является α -хемокиновым рецептором, который специфичен для стромального фактора-1 (SDF1, также называется CXCL12) - молекулы, обладающей сильной хемотаксической активностью для лимфоцитов. Экспрессия CXCR4 слабая или отсутствует во многих здоровых тканях, но сильная во многих типах рака, включая рак молочной железы, рак яичников, меланому и рак простаты. Сильная экспрессия этого рецептора в раковых клетках связана с высокой концентрацией CXCL12 в раковых клетках, причем экспрессия CXCL12 положительно коррелирует с CXCR4-положительными клетками. LPA связывает молекулу сигнального липида - лизофосфатидной кислоты (LPA). Вследствие способности LPA стимулировать пролиферацию клеток, нарушения LPA-сигнализации связаны с раком различным образом. К примеру, дисрегуляция аутоактина или рецепторов LPA может приводить к гиперпролиферации, что может способствовать онкогенезу и метастазированию.

Синтетические GPCR настоящего изобретения могут быть созданы путем замены домена G-белка у GPCR доменом dCas9-активатора (фиг. 21A). Например, dCas9-VPR (состоящий из трех частей активатора, содержащий VP64, p65AD и Rta) может быть слит с GPCR через расщепляемую TEV последовательность 2111. dCas9-VPR также может быть слит с двумя копиями сигнала ядерной локализации (NLS) для усиления ядерной локализации dCas9-VPR после расщепления TEV. Протеаза TEV может быть слита с β -аррестинном, белком, который условно привлекается к GPCR при связывании лиганда 2121. Воплощение настоящего изобретения, представленное на фиг. 21, показывает, что сконструированный активатор GPCR-dCas9 может экспрессироваться в Т-клетках 2111 и использоваться для лечения рака типа рака яичников 2101. Повышение внеклеточной концентрации LPA (лиганда GPCR) может активировать сконструированный LPAR1-dCas9 2131, который затем рекрутирует β -аррестин, слитый с протеазой TEV 2121. Протеаза может высвобождать dCas9 в ядро путем отщепления TEV-специфичного пептида. dCas9 также может образовывать комплекс с направляющей РНК, например, программируемой sgRNA. В ядре dCas9-sgRNA может активировать гены 2141 типа IL-2, которые могут усилить уничтожение клеток 2151.

На фиг. 23A и 23B представлены результаты трех независимых экспериментов по тестированию активности химерного CXCR4-dCas9-VPR-mCherry 2301 в клетках хозяина, например клетках HEK293. После связывания лиганда 2311 β_2 -аррестин-протеаза привлекается к активированному рецептору, а протеаза высвобождает dCas9-VPR-mCherry из рецептора 2321. Экспрессионная конструкция для β_2 -аррестин-протеазы представлена как 2351. Репортерная конструкция представлена как 2341. В каждом эксперименте dCas9-VPR работает в паре с одинарной направляющей РНК (sgRNA), которая специфически активирует экспрессию люциферазы, которая в свою очередь генерирует люминесценцию. В этом эксперименте использовали sgRNA-BFP 2331. В некоторых экспериментах в качестве контроля использовали scFv-mCherry-VPR 2305 и/или dCas9-10xGCN4 2307.

CXCR4-dCas9-VPR-mCherry 2361 подвергали условной экспрессии из промотора, индуцируемого доксициклином (TRE3G) с тем, что только добавление доксициклина (Dox) и присутствие коактиватора может запустить экспрессию рекомбинантного рецептора. В отсутствие Dox не было люминесценции (фиг. 23B). В качестве положительного контроля добавляли "свободный" dCas9-VPR (то есть не слитый с CXCR4), при этом, как и ожидалось, вырабатывалась люминесценция. При добавлении Dox экспрессировался химерный CXCR4-dCas9 и регистрировалась люминесценция. Из фиг. 23B видно, что люминесценция была чувствительна к возрастающим концентрациям CXCL12, лиганда CXCR4. Описанные здесь химерные белки GPCR-dCas9 можно экспрессировать в таких иммунных клетках, как Т-клетки и макрофаги. Кроме того, можно провести эксперименты по хемотаксису, чтобы установить, может ли CXCL12 служить хемотактантом для Т-клеток, экспрессирующих химерные белки GPCR-dCas9.

Пример аминокислотной последовательности полипептида CXCR4-dCas9-VPR-mCherry представлен в SEQ ID NO: 3.

Обращаясь к фиг. 25C, в клетках HEK293T, экспрессирующих sgRNA, β_2 -аррестин-протеазу ("A") и LPAR1-dCas9-VPR ("B"), CXCR4-dCas9-VPR ("C") либо hM3D-dCas9-VPR ("D"), присутствие лиганда (например, LPA, SDF1 или CNO) приводило к повышению уровня репортерного белка GFP по сравнению с отсутствием лиганда (фиг. 23D). Связывание лиганда соответственно с каждым химерным рецептором приводило к активации рецептора и рекрутированию адаптера-протеазы. Рекрутирование к рецептору адаптера-протеазы приводило к отщеплению dCas9-VPR от рецептора по сайту расщепления для TEV. Высвобожденный dCas9-VPR затем наводился с помощью sgRNA на целевой полинуклеотид и изменял уровень экспрессии репортера GFP.

A2. Рекомбинантный химерный GPCR-рецептор, содержащий расщепляющий элемент.

В альтернативной конфигурации синтетические GPCRs настоящего изобретения содержат расщепляющий элемент, а адаптерный полипептид содержит dCas9 с эффекторным доменом. Например, как показано на фиг. 19B, dCas9-VPR (состоящий из трех частей активатор, содержащий VP64, р65AD и Rta) 2161 может быть слит с адаптерным белком 2171 типа β_2 -аррестина через расщепляемую TEV последовательность 2181. β_2 -аррестин-dCas9-VPR также может содержать две копии сигнала ядерного экспорта (NES) для усиления ядерного экспорта β_2 -аррестина-dCas9-VPR. Протеаза TEV 2191 в этой конфигурации может быть слита с GPCR 2100 (GPCR-протеаза). Когда β_2 -аррестин-dCas9-VPR привлекается к активированной GPCR-протеазе, протеаза может расщеплять TEV-расщепляемую последовательность и высвобождать dCas9-VPR из полипептида β_2 -аррестин-dCas9-VPR.

Например, сконструированный β_2 -аррестин-dCas9-VPR может экспрессироваться вместе с синтетической GPCR-протеазой типа GPCR LPAR1. Повышение внеклеточной концентрации LPA (лиганда GPCR) может активировать сконструированную LPAR1-протеазу, которая затем рекрутирует β_2 -аррестин-dCas9-VPR. Протеаза может высвобождать dCas9 в ядро путем отщепления TEV-специфичного пептида. dCas9 также может образовывать комплекс с направляющей РНК, например, программируемой sgRNA. В ядре dCas9-sgRNA может активировать гены типа гена флуоресцентного репортера.

В качестве другого примера, сконструированный β_2 -аррестин-dCas9-VPR может экспрессироваться вместе с синтетической GPCR-протеазой типа GPCR CXCR4. Повышение внеклеточной концентрации SDF1 (лиганда GPCR) может активировать сконструированную CXCR4-протеазу, которая затем рекрутирует β_2 -аррестин-dCas9-VPR. Протеаза может высвобождать dCas9 в ядро путем отщепления TEV-специфичного пептида. dCas9 также может образовывать комплекс с направляющей РНК, например, программируемой sgRNA. В ядре dCas9-sgRNA может активировать гены типа гена флуоресцентного репортера.

В качестве другого примера, сконструированный β_2 -аррестин-dCas9-VPR может экспрессироваться вместе с синтетической GPCR-протеазой типа как hM3D (версия DREADD GPCR hM3). Повышение внеклеточной концентрации клозапин-N-оксида (CNO, лиганда GPCR) может активировать сконструированную hM3D-протеазу, которая затем рекрутирует β_2 -аррестин-dCas9-VPR. Протеаза может высвобождать dCas9 в ядро путем отщепления TEV-специфичного пептида. dCas9 также может образовывать комплекс с направляющей РНК, например, программируемой sgRNA. В ядре dCas9-sgRNA может активировать гены типа гена флуоресцентного репортера.

Обращаясь к фиг. 23E и 23F, уровни репортера GFP в присутствии hM3D-протеазы (фиг. 23E, "A") + β_2 -аррестина-dCas9-VPR (фиг. 23E, "C") + sgRNA (sgTET, который нацелен на промотор TRE3G репортерного гена GFP) в присутствии лиганда (CNO) сопоставимы с уровнями репортера GFP в присутствии hM3D-протеазы (фиг. 23E, "A") + dCas9-VPR (фиг. 23E, "E") + sgRNA (фиг. 23F, "A + B + CNO" по сравнению с "D"). В качестве положительного контроля, уровни репортера GFP в присутствии hM3D-протеазы (фиг. 23E, "A") + β_2 -аррестина-TetR-VPR (фиг. 23E, "B") + sgRNA (в присутствии лиганда, CNO) сравнимы с уровнями репортера GFP в присутствии hM3D-протеазы (фиг. 23E, "A") + TetR-VPR (фиг. 23E, "D") + sgRNA (фиг. 23F, "A + C + CNO" по сравнению с "E" (TetR связывается непосредственно с промотором репортерного гена). Связывание лиганда (CNO) с hM3D-протеазой приводило к активации рецептора и рекрутированию адаптера-dCas9-VPR. Рекрутирование β_2 -аррестина-dCas9-VPR к активированному рецептору приводило к отщеплению dCas9-VPR от адаптера по сайту расщепления для TEV. Высвобожденный dCas9-VPR затем наводился с помощью sgRNA на целевой полинуклеотид и изменял уровень экспрессии репортера GFP.

В. Рекомбинантный химерный интегрин, содержащий dCas9 с эффекторными доменами.

Здесь также представлены сконструированные интегрины, слитые с dCas9 с эффекторными доменами, например dCas9 с активаторными доменами (фиг. 22A). Интегрин, содержащий различные пары α - и β -субъединиц, может воспринимать разнообразные сигналы матрикса типа фибронектина, коллагена, ламинина и др. dCas9-VPR-mCherry сливали через TEV-расщепляемую последовательность пептидного линкера с α - либо β -субъединицей любого интегрин, например интегрин $\alpha_5\beta_1$ 2201. Протеазу TEV соединяли с паксиллином, белком, который условно привлекается к парным интегринам при связывании лиганда 2221. В клетки хозяина 2231 вводили sgRNA, которая может связываться с репортерным геном. Кроме того, вводили ген репортера транскрипции (H2B-GFP) для визуализации направляемой sgRNA контролируемой dCas9-VPR активации 2241. В качестве контроля на экспрессию рецептора использовали интегрин, связанный с scFv, mCherry и VPR 2211. Другие контроли в этом эксперименте включали контроль dCas9-10xGCN4 2251. Как видно из фиг. 22B, dCas9-активаторный домен, связанный с β -субъединицей (фиг. 22D, "B"), без лиганда, давал минимальный сигнал GFP (геномная копия GFP использовалась в качестве репортерного сигнала для регистрации количества "высвобождаемого" в ядро dCas9-VPR). В присутствии фибронектина проявлялась высокая экспрессия GFP, свидетельствуя о том, что химерный белок интегрин-dCas9 активирует транскрипцию при связывании лиганда. В присутствии коллагена химерный белок интегрин-dCas9 не активировал транскрипцию репортерного гена. Наряду с активацией интегринов при связывании лиганда, активация интегрин при поверхностном контакте так-

же приводит к высвобождению dCas9-VPR и экспрессии GFP. Система интегрин-dCas9-VPR+паксиллин-протеаза + sgPHK при экспрессии в адгезированных клетках дает более высокий уровень флуоресцентного репортерного белка по сравнению с суспензионными культурами (фиг. 22С). Когда dCas9-VPR связан с α -субъединицей (фиг. 22D, "С"), активация интегрин приводит к минимальной экспрессии GFP (фиг. 22Е, "А+С") по сравнению с тем, когда dCas9-VPR связан с β -субъединицей (фиг. 22Е, "А+В"), так как паксиллин специфически связывается с хвостами β -интегрин.

Пример аминокислотной последовательности полипептида β_1 -интегрин-dCas9-VPR-mCherry представлен в SEQ ID NO: 4.

С. Рекомбинантный химерный Notch, содержащий dCas9 с эффекторными доменами.

Также представлены химерные рецепторы Notch, содержащие dCas9 с активаторными доменами (фиг. 24А-24D). В одном воплощении Notch непосредственно сливали с dCas9-активатором без линкерной последовательности расщепляемого TEV пептида. Во втором воплощении Notch непосредственно сливали с dCas9-активатором через последовательность расщепляемого пептида TEV. Протеазу TEV соединяли с белком адаптера Notch, пресенилина-1 (PS-1). Пресенилин 1 является одним из четырех главных белков в комплексе пресенилина, который опосредуют регулируемые протеолитические события у нескольких белков в клетке, включая гамма-секретазу. Для каждого из химерных рецепторов создавали экспрессионные конструкции и вводили в клетки хозяина типа клеток НЕК293. Пример аминокислотной последовательности полипептида Notch-dCas9-VPR-mCherry представлен в SEQ ID NO: 5.

Пример 3. Контроль регуляции генов с помощью сконструированных химерных рецепторов Notch, слитых с дефектными по нуклеазе белками Cas9.

В этом примере описан новый подход к клеточной терапии - создание модифицированных иммунных клеток, которые могут распознавать пораженное болезнью микроокружение и отвечать точным терапевтическим действием. Иммунные клетки пациента являются перспективными реагентами для этих терапевтических средств, так как они используют врожденные сложные системы, которые могут воспринимать и реагировать на определенные клетки и локальное окружение клетки.

В отличие от использования нуклеазы Cas9 для редактирования генов для активации или репрессии транскрипции можно использовать дефектный по нуклеазе белок Cas9 (dCas9) без генетического изменения последовательности генома. Белок dCas9 в паре с небольшой направляющей РНК (sgRNA) может распознавать и регулировать гены, содержащие комплементарные последовательности. Вкупе с активаторами или репрессорами транскрипции эта система может представлять высокопрограммируемый подход для точного контроля генов. Хотя dCas9 представляет мощный подход для модуляции клеточной активности, он не обладает способностью воспринимать сигналы типа внешних или сигналов окружающей среды.

Было показано, что клеточные сигнальные пути представлены теми молекулами, которые можно перенастроить на реагирование на несобственные сигналы путем обмена и рекомбинации их собственных сигнальных компонентов. В частности, сигнальный путь Notch-Delta особенно привлекателен для искусственной перенастройки вследствие его простого механизма действия, быстрой динамики и важной роли в регуляции линейки иммунных клеток и прогрессирования рака. Notch и Delta - это семейства однопролетных трансмембранных белков у многоклеточных животных. Их взаимодействие ведет к протеолитическому высвобождению внутриклеточного домена Notch (NICD), который транслоцируется в ядро и активирует целевые гены. Для изменения молекул Notch применялись подходы белковой инженерии путем замены NICD на синтетические факторы транскрипции (GAL4-AD) для активации альтернативных генов пути Notch либо путем замены внеклеточного домена Notch (NECD) на распознавательные мотивы против определенных рецепторов-антигенов.

А. Конструирование модульного искусственного химерного рецептора Notch для запускаемой CD47, опосредованной CRISPR регуляции транскрипции.

В этом примере описаны химерные рецепторы, в которых сочетаются белки Notch и CRISPR/Cas для получения настраиваемой ортогональной системы сигнализации для распознавания клеток мишени (через Notch) и программируемой реакции клеток (через CRISPR/dCas9). Это высокопрограммируемая платформа распознавания рака, в которой используется опосредованная контактами сигнализация Notch-Delta и управляемая направляющей РНК CRISPR-инженерия генома. Такие химерные рецепторы Notch (фиг. 24А-24С) могут применяться для активации любых целевых генов путем замены NICD Notch на компоненты Cas9. Посредством взаимодействия Notch-Delta можно управлять широким спектром сложных клеточных реакций. Описанные здесь химерные рецепторы антигенов могут применяться для модуляции транскрипции и клеточной активности Т-клеток в ответ на локальное межклеточное микроокружение.

Можно получить модульный искусственный химерный рецептор Notch для запускаемой CD47, опосредованной CRISPR регуляции транскрипции. Рецепторы Notch дикого типа содержат два модуля: внеклеточный домен (NECD) и внутриклеточный домен (NICD). Для получения химерного искусственного рецептора Notch (caN) домен NICD можно заменить на dCas9, слитый с активатором или репрессором транскрипции (а также слитый с синим флуоресцентным белком BFP для визуализации). К примеру,

dCas9 может быть слит с активаторным доменом VP64 или VPR. Слитый dCas9 может транслоцироваться в ядро при взаимодействии Notch-Delta (фиг. 24B). Конструкцией, кодирующей химерный искусственный рецептор Notch, можно трансформировать клетки. Для проверки транслокации в ядро слитого dCas9 в качестве NICD можно совместно экспрессировать такую sgRNA, которая нацелена на репортерную конструкцию GAL4 UAS-GFP. После совместного культивирования с другой линией клеток, экспрессирующих Delta, можно по флуоресценции dCas9-VP64-BFP отслеживать трансактивацию Notch-Delta во времени и транслокацию из плазматической мембраны в ядро.

NECD Notch содержит 29-36 tandemных повторов типа эпидермального фактора роста (EGF), из которых 11-12 повторов типа EGF способствуют продуктивным взаимодействиям с Delta. Для неканонического распознавания отдельные участки NECD могут быть заменены когнатным одноцепочечным вариативным фрагментом (scFv). Различные участки EGF-повторов можно заменить на MABL, scFv против CD47 человека (hCD47). Для измерения транскрипционной активности слитого рецептора можно присоединить к ECD минимальный активатор GAL4esn. В частности, можно использовать варианты MABL(ECD)-GAL4esn, используя экспрессию репортера UAS-GFP при совместном культивировании с клетками, экспрессирующими hCD47 (фиг. 24C). Химерных антигеновых рецепторы Notch могут рекрутировать пресенилин с протеазой TEV после связывания рецептора с Delta или желательным лигандом.

Для получения нового связывающего CD47 рецептора scFv-Notch можно объединить модули NICD и NECD. Ожидаемый комбинаторный эффект заключается в том, что hCD47 активирует рецептор CD47scFv-Notch-dCas9, вызывая отщепление dCas9-VP64 и его транслокацию в ядро, которая может направляться направляющей sgRNA для активации экспрессии репортера. Можно получить различные варианты, используя варианты dCas9 и различные sgRNA. Химерные рецепторы можно трансфицировать в клетки таких линий, как CHO, HEK293 и Jurkat.

Notch-dCas9 с активаторами, представленные на фиг. 24A-24C, тестировали на клетках HEK293 (фиг. 24E), Т-клетках (клетки Jurkat, фиг. 24F) и макрофагах (THP-1; фиг. 24G) на их активность и функции. В этих экспериментах Notch-dCas9-активатор экспрессировался в каждой линии клеток. При взаимодействии Delta-Notch активировалась экспрессия флуоресцентного репортера из генома. Клетки проявляли высокий уровень экспрессии белка-репортера при обработке рекомбинантных клеток Delta и низкую экспрессию без обработки. Результаты показывают, что искусственный химерный рецептор полностью функционален в иммунных клетках. Для проверки влияния таких рецепторов на активацию пролиферации клеток в клетки хозяина вводили короткие гидовые РНК, специфичные к генам клеточного апоптоза (каспазы-8 (CASP8)) или генам клеточного цикла (циклина D1 (CCND1) и ингибитора циклин-зависимой киназы 1B (CDKN1B)), вместе с химерными рецепторами. Результаты представлены на фиг. 26A-26D.

В другом эксперименте Notch-dCas9-активатор экспрессировали в клетках вместе с такой направляющей РНК (sgUAS, SEQ ID NO: 1, gactccgacctctagtgt), которая может связываться с вышележащей активирующей последовательностью (UAS) вблизи от промотора гена флуоресцентного репортера (фиг. 27A). Химерный искусственный рецептор Notch содержит домены ECM и TM Notch дикого типа, "мертвый" по нуклеазной активности Cas9 (dCas) и трехчастный эффекторный домен, состоящий из белков VP64, p65 и Rta (VPR). Лиганд Delta, который связывается с Notch, может быть либо иммобилизован на поверхности, либо презентируван другой клеткой. Связывание Notch-Delta ведет к отщеплению и транслокации dCas9-VPR в ядро. dCas9 и одинарная направляющая РНК (sgUAS), нацеленная на регуляторный элемент UAS (промотор), вызывают экспрессию помеченного флуоресцентным цитрином гистона-2В (H2B). На фиг. 27B представлена флуоресценция одиночных клеток при проточной цитометрии клеток из фиг. 27A. У клеток, которые культивировались на поверхности, покрытой Delta, отмечалась гиперэкспрессия H2B-цитрина по сравнению с клетками, культивированными на поверхности без Delta.

В. Характеристики поведения химерного антигенового рецептора Notch (caN), включая параметры, которые формируют функцию отклика на сигнал *in silico* и *in vitro*.

Простота активации Notch-Delta позволяет настраивать реакцию на сигналы. С помощью математического моделирования и экспериментов можно исследовать, как взаимное цис-ингибирование между Notch и Delta влияет на функцию отклика, когда присутствует и эндогенный, и синтетический Notch. Это можно сделать путем экспрессии caN-dCas9-EGFP и Delta-mCherry в одних и тех же клетках с использованием промоторов различной силы.

Для изучения того, как caN перестраивает клеточные реакции (например, фагоцитарную активность, P) на внешние сигналы (например, Delta или CD47), можно рассмотреть одну ситуацию, при которой эндогенный Notch в принимающей клетке (R на фиг. 25A-25C) и Delta в сигнализирующей клетке (Dt на фиг. 25A-25C) вызывают репрессию реакции P принимающей клетки (фиг. 25A-25C). Для простоты можно принять одинаковые константы скорости без ограничения общности. Модели показывают, что P гиперболически уменьшается с возрастанием D_T в стационарном состоянии. При добавлении рецептора caN-dCas9 (N в принимающей клетке), который теперь подавляет экспрессию эндогенного Notch (R) с большим коэффициентом кооперативности (ηR , критический параметр, который модулируется dCas9 с репрессорным доменом), чем эндогенная репрессия P под действием R, можно перевести выходной сигнал в бимодальный ответ с возрастанием уровня D_T (фиг. 25B). Наконец, еще одним критическим пара-

метром, определяющим порог, при котором Р-ответ меняется из репрессии на активацию в ответ на D_T , является дополнительный уровень ингибирования цис-Delta. (D_C), который конкурирует с D_T при связывании с N (но не с R) (фиг. 25C).

Критические параметры, установленные *in silico*, можно оптимизировать с помощью экспериментов *in vitro*. В частности, можно модулировать уровни экспрессии различных рецепторных компонентов (D_C , D_T , R и N на фиг. 25A-25C) с применением конструкций конститутивных промоторов с известной силой действия. Для настройки силы репрессирования можно выбрать различные dCas9-эффекторы. Активационную и репрессорную силу вариантов рецептор-dCas9-эффектор можно протестировать с помощью известных и эффективных sgRNAs, регулирующих определенные эндогенные гены. Этот метод также позволяет модифицировать R с минимальными неприцельными эффектами.

С. Конструирование макрофагов со связывающими CD47 рецепторами caN-dCas9, активирующими фагоцитарные реакции против раковых клеток с высоким уровнем CD47.

Связывающий CD47 слитый с dCas9 белок химерного антигенового рецептора Notch, описанный здесь, может применяться в таких областях онкологии, как прогрессирование рака и иммунное уклонение. Для проверки эффективности связывающего CD47 рецептора caN-dCas9 в опосредовании или регулировании иммунного уклонения при раке можно экспрессировать описанный ранее рецептор CD47scFv-Notch-dCas9 в макрофагах, полученных из THP-1. CD47 человека можно экспрессировать в клетках CHO. Для разработки и проверки теоретических моделей на этих клетках в итеративном режиме можно провести анализ поверхностной экспрессии SIRP α (эндогенный ответ) и caN-dCas9 в макрофагах и hCD47 в целевых клетках CHO методом проточной цитометрии.

Для проверки того, будет ли изменяться естественная или нативная фагоцитарная реакция у макрофагов с химерным антигеновым рецептором Notch, можно провести сравнение фагоцитарной реакции у кодирующих caN-dCas9 макрофагов THP-1 при трансдукции вместе с sgPHK против SIRP α или без неё. Можно разработать sgSIRP α и dCas9-репрессор соответствующей силы действия и ввести в итерационную модель. Преобразованные макрофаги можно инкубировать с эпителиальными клетками аденокарциномных линий человека (A549), экспрессирующими низкие, средние или высокие уровни hCD47 (клетки-мишени D_T). Конструкцию связывающего CD47 слитого белка caN-dCas9 можно оптимизировать так, чтобы клетки A549 со средним и высоким уровнем CD47 фагоцитировались преобразованными макрофагами, а клетки с низким уровнем CD47 фагоцитировались не содержащими sgSIRP α макрофагами.

В этом примере представлены новые рецепторы для иммунотерапии рака типа рационально сконструированных химерных рецепторов, содержащих домены Notch и dCas9. Эти рецепторы позволяют программировать контроль транскрипции в иммунных клетках путем определения специфичных для рака уровней рецепторов.

Пример 4. Контроль регуляции генов с помощью сконструированных химерных GPCR-рецепторов, слитых с дефектными по нуклеазе белками Cas9.

Получали химерный искусственный GPCR-рецептор, включающий в себя полноразмерный GPCR с С-концевым сайтом связывания β_2 -аррестина, полученным из AVPR2 (27 аминокислотных остатков), сайт расщепления протеазой и "мертвый" безнуклеазный Cas9 (dCas9) с трехчастным эффекторным доменом (VPR) (фиг. 28A). Химерная β_2 -аррестин-протеаза распознает и расщепляет сайт расщепления протеазы при опосредованной лигандом активации GPCR. dCas9-VPR транслоцируется в ядро и вместе с одинарной направляющей РНК (sgTet, SEQ ID NO: 2; gtagcttctctccactgata), которая специфически нацелена на промотор tetO, вызывает экспрессию репортерного гена люциферазы. Клетки HEK293, стабильно экспрессирующие управляемый tetO ген люциферазы и конструкцию с β_2 -аррестин-протеазой, трансфицировали одной из следующих экспрессирующей конструкций, кодирующих: dCas9-VPR, CXCR4 или химерный CXCR4-dCas9-VPR. Трансфицированные клетки обрабатывали CXCL12 (лиганд для CXCR4) в течение 18 часов при различных концентрациях. Для измерения степени экспрессии люциферазы добавляли люциферин с тем, что при расщеплении люциферазой испускается и детектируется люминесцентный сигнал (фиг. 28B).

Пример 5. Контроль регуляции генов с помощью сконструированных химерных интегриновых рецепторов, слитых с дефектными по нуклеазе белками Cas9.

Получали химерный искусственный интегриновый рецептор, включающий в себя полноразмерный S_1 -интегрин, сайт расщепления протеазой и "мертвый" безнуклеазный Cas9 (dCas9) с трехчастным эффекторным доменом (VPR) (фиг. 29A). Внеклеточный матрикс, который связывается с димерами $\alpha\beta$ -интегрин, может быть иммобилизован на поверхности или презентируван другой клеткой. Активация интегрин приводит к связыванию химерной паксиллин-протеазы с С-концом β -интегрин, расщеплению сайта расщепления протеазой и транслокации dCas9-VPR в ядро. Паксиллин распознает мотив HDRK (SEQ ID NO: 91) β -интегрин. dCas9 вместе с одинарной направляющей РНК, sgUAS (SEQ ID NO: 1; gtagctccgacctctagtgt), которая нацелена на промотор вышележащей активирующей последовательности (UAS), могут активировать экспрессию репортерного гена, например, гена, кодирующего помеченный флуоресцентным цитрином гистон 2В (H2B). Клетки HEK293, стабильно экспрессирующие управляемый

UAS ген H2B-цитрина и конструкцию с паксиллин-протеазой, трансфецировали экспрессирующей конструкцией, кодирующей химерный β_1 -интегрин-dCas9-VPR и α -интегрин. Трансфецированные клетки культивировали на покрытой ЕСМ поверхности. Данные проточной цитометрии показывают, что клетки, контактирующие с покрытой ЕСМ поверхностью, гиперэкспрессировали H2B-цитрин по сравнению с клетками, которые не контактировали с ЕСМ (фиг. 29B).

Пример 6. Несколько сконструированных химерных рецепторов, слитых с доменами, модулирующими ген, для логики, каскада и сетей для программируемых функций.

При сцеплении двух рецепторов вместе так, чтобы первый рецептор при связывании лиганда активировал экспрессию второго рецептора, образуется вентиль "AND". Таким образом, только присутствие обоих лигандов может вызвать клеточный ответ. Также возможна схема "А, но не В". Например, чтобы определить, содержит ли раковая клетка первый антиген, но лишена второго другого антигена, рецептор А можно слить с активатором, а рецептор В - с репрессором. Это позволит интегрировать множественные сигналы, связанные с раком, для более конкретной инженерии сигналов. Эта система также может применяться для проверки конструкции "генетической памяти", которая может изменять эпигенетическое состояние клеток при активации.

Логический элемент "AND", описанный здесь, представляет собой устройство, которое поддерживает совместную стимуляцию. Такие природные процессы, как активация лимфоцитов, требуют стимуляции для развития эффективного иммунного ответа. Костимуляция Т-клеток важна для пролиферации, дифференцировки и выживания Т-клеток. Активация Т-клеток без костимуляции может привести к апноэ Т-клеток, делеции Т-клеток или развитию иммунотолерантности. Например, чтобы стать полностью активированными, Т-клетки могут зависеть от двух сигналов: первый сигнал является антиген-специфичным и включает в себя Т-клеточные рецепторы (TCR), взаимодействующие с молекулами пептидов МНС на мембране антиген-презентирующих клеток (APC), а второй сигнал не специфичен к антигенам и возникает при взаимодействии между костимулирующими молекулами, экспрессированными на мембране APC и Т-клеток. Костимулирующие молекулы могут быть заменены описанными здесь сконструированными рекомбинантными химерными рецепторами, издающими произвольные костимулирующие сигналы для TCR. Например, один костимулирующий рецептор, экспрессируемый Т-клетками, представлен CD28, который взаимодействует с CD80 (B7.1) и CD86 (B7.2) на мембране APC. Другой костимулирующий рецептор, экспрессируемый Т-клетками, представлен ICOS (индуцибельный костимулятор), который взаимодействует с ICOS-L. CD28 и ICOS могут быть подвергнуты инженерии для поддержки процесса костимуляции TCRs.

По схеме "расщепленного AND" (фиг. 30A) полипептид dCas9 можно разделить на нефункциональные компоненты или домены, которые могут быть индивидуально пришиты к требуемым рецепторам (например, Notch и GPCR) для распознавания своих соответствующих лигандов. Активация химерных рецепторов обоими лигандами может вызвать расщепление и восстановление функции dCas9 и последующую активацию гена. Этот способ модуляции гена не работает, если ни один или хотя бы один из лигандов не связывается с химерными рецепторами.

По схеме "каскадного AND" (фиг. 30B) первый химерный рецептор (например, Notch) при связывании со своим лигандом (например, Delta) индуцирует управляемую TetO экспрессию второго химерного рецептора, слитого с dCas9, через тетрациклиновый трансактиватор tTa. Второй химерный рецептор (например, GPCR) может распознавать свой лиганд, а dCas9 в свою очередь может отщепляться и индуцировать управляемую направляющей sgRNA экспрессию целевого гена (например, H2B-цитрина). По схеме "каскадного AND" связывание первого лиганда и второго лиганда со своими соответствующими рецепторами важно для возникновения сигнального каскада.

Пример 7. Преобразование внеклеточных сигналов в регуляцию генома.

Были разработаны химерные рецепторы типа Notch1-Cas9 человека, раскрытые здесь, которые при связывании со внеклеточным лигандом Delta высвобождают мембраносвязанный Cas9 для манипулирования экспрессией генов у млекопитающих. Эти химерные рецепторы проявляют устойчивую Delta-зависимую, опосредованную Cas9 экспрессию репортерного гена и модуляцию функции эндогенного гена. Такая технология химерных рецепторов с Cas9 дает возможность зависимость от внеклеточных сигналов конструирования генома, что может быть полезным для анализа сигнальных путей и обеспечения чувствительных к микроокружению клеточных функций.

Ассоциированный с кластеризованными, регулярно разделенными промежутками короткими палиндромными повторами (CRISPR) микробный белок 9 (Cas9) применялся для прицельного редактирования, активации или репрессирования практически любых генов. Одинарная направляющая РНК (sgRNA) связывается с последовательностями ДНК, комплементарными Cas9, который затем катализирует двухцепочечный разрыв посредством своих доменов RuvC и HNH. Мутации в этих доменах и слияние с эффекторными доменами может перенастроить каталитически мертвый Cas9 (dCas9) на выполнение прицельной регуляции генома в клетках млекопитающих.

Регулирование генома через CRISPR-Cas9 является процессом типа "изнутри наружу", в котором генотипические изменения могут влиять на поведение клеток. И наоборот, внеклеточное микроокружение также может влиять на поведение клеток через процесс сигнализации "снаружи-внутри". Поэтому

была предпринята ранее не исследованная стратегия для соединения обоих клеточных процессов (фиг. 31А). Прямое слияние Cas9 с однопролетным трансмембранным рецептором Notch образует основу химерного рецептора Notch, первого в своем роде инструмента, который направляет внеклеточные сигналы на регуляцию генома на основе CRISPR. Рецептор Notch в некоторых случаях является привлекательным кандидатом из-за простоты своего механизма. Связывание своего когнатного лиганда Delta со внеклеточным доменом Notch (NECD) позволяет комплексу γ -секретазы расщеплять внутриклеточный домен Notch (NICD), который затем транслоцируется в ядро и функционирует в качестве фактора транскрипции. Ранее анализ видоспецифичных гомологов Notch и их доменов показал модульность как внеклеточных, так и внутриклеточных доменов Notch.

Создавали слитую конструкцию NC1 путем замены NICD из Notch1 человека на dCas9 *Streptococcus pyogenes*, слитый с трехчастным активационным доменом (VP64, p65 и Rta) (фиг. 31В) и с тремя копиями последовательности сигнала ядерной локализации (NLS). Конструкцию поместили флуоресцентной меткой mCherry для определения её субклеточного распределения. Затем исследовали, как слитая конструкция NC1 отвечает на внеклеточные лиганды Delta при трансфекции в клетки млекопитающих. Использовали репортерную линию клеток яичников китайского хомячка (CHO), которая ранее была разработана для анализа сигнализации Notch-Delta (Sprinzak D. et al. Cis-interactions between Notch and Delta generate mutually exclusive signalling states. *Nature* 465, 86-90, doi:10.1038/nature08959 (2010)). Линия клеток CHO содержала индуцибельный промотор UAS (вышележащей активирующей последовательности), контролирующей ген помеченного цитрином (YFP) репортера, гистона 2В (H2В). Посредством лентивирусной трансдукции стабильно экспрессировали направляющую sgRNA (sgUAS), которая позволяет расщепленному dCas9 связываться с промотором UAS (см. табл. 6). Химерную конструкцию hNECD-dCas9-VPR трансфецировали в клетки CHO, которые затем культивировали в течение 4 дней на голой поверхности планшета или на поверхности, покрытой насыщающим количеством Delta (DLL4 с усиливающими мутациями (Luca V.C. et al. Structural biology. Structural basis for Notch1 engagement of Delta-like 4. *Science* 347, 847-853, doi:10.1126/science.1261093 (2015)). По флуоресценции связанного с ядром H2В-цитрина определяли, будет ли NC1 индуцировать экспрессию H2В при контакте с адсорбированным на поверхности Delta (фиг. 31С). В некоторых клетках отмечалось, что возникающие молекулы NC1 застревают в эндоплазматическом ретикулуме (ER) и преждевременно активируют экспрессию H2В, возможно, отчасти из-за наличия 3 мотивов NLS (фиг. 33).

Осуществление запускаемой внеклеточным сигналом опосредованной dCas9 регуляции генома в этом примере включало балансирование мембранной локализации химерных рецепторов до расщепления и транслокации dCas9 в ядро после расщепления (фиг. 33). Для оптимизации их функции у нескольких вариантов hNECD-dCas9 варьировали количество копий NLS и известных сигналов мембранного созревания (MMS) (фиг. 31В). Вариант NC2, который не содержал NLS, тоже индуцировал экспрессию H2В в отсутствие лиганда Delta, свидетельствуя о том, что сам размер внутриклеточного домена, включающего dCas9-VPR-mCherry (243 кДа; размер hNECD составляет 188 кДа) мог вызвать его преждевременное расщепление. Поэтому последовательность NICD в 22 а.к. заново вводили в конструкцию NC3; эта конструкция работала не лучше, чем NC1 или NC2 (фиг. 34). Также в С-концевую часть гибридной химеры вводили последовательность MMS (последовательность: RSQQEAAAKKFF (SEQ ID NO: 30)) из LMAN1, трансмембранного белка, вовлеченного в сортировку и рециркуляцию белка, получая варианты NC4 и NC5. NC4 содержит один синтетический NLS, а NC5 - нет. Только NC5 проявлял Delta-зависимую активацию H2В (фиг. 31D), свидетельствуя о том, что синтетические NLS нарушают правильность мембранного созревания, а такие MMS, как в NC5, способствуют надлежащему созреванию. Наконец, NC5 и sgUAS были стабильно встроены в репортерные клетки и был выделен один клон на основании минимальной базальной экспрессии H2В в отсутствие Delta. Наблюдалась трехкратная активация H2В в ответ на Delta (фиг. 31Е и 31F), поэтому конструкция NC5 использовалась для дальнейших исследований.

В отсутствие синтетических NLS в конструкциях NC2 и NC5 отщепленный dCas9-VPR-mCherry все-таки проявлял транслокацию в ядро, что могло быть частично связано с предсказанным собственным мотивом NLS (iNLS) по остаткам 647-670 у Cas9 из *S. pyogenes*. Эта последовательность аминокислотных остатков iNLS (VMKQLKRRRYTGW GRLSRKLINGI (SEQ ID NO: 31)) попадала на мотив спираль-линкер-спираль в кристаллической структуре Cas9 из *S. pyogenes*. Поверхностная диспозиция обеих спиралей свидетельствует об их доступности для импортин-зависимого ядерного транспорта (фиг. 35А). Мутация 6 остатков аргинина или лизина этого iNLS на аланин (VMAQLKAAA~~Y~~TGWGRLSAALINGI (SEQ ID NO: 32)) в конструкции dCas9-VPR-mCherry ухудшает активацию EGFP (с 18-кратной до 4-кратной) в репортерных клетках HEK293. А у небольшого процента (~16%) клеток EGFP активировался, свидетельствуя о том, что этот мутантный вариант iNLS остается функциональным и/или что еще остались другие неканонические скрытые NLS (фиг. 35В). Разрушение этого мотива iNLS посредством мутаций остатков аргинина и лизина на аланины VMAQLKAAA~~Y~~TGWGRLSAALINGI (SEQ ID NO: 32) в конструкции dCas9-VPR-mCherry сильно нарушает ядерную локализацию и активацию репортера EGFP в эмбриональных почечных клетках челове-

ка (HEK293T), как это видно на репрезентативных снимках конфокальной флуоресцентной микроскопии на фиг. 35C. Справа представлены репрезентативные гистограммы репортера EGFP и процент клеток, считающихся "включенными (ON)", у которых интенсивность репортера была выше, чем у нетрансфицированных "выключенных (OFF)" клеток. Активация EGFP частично восстанавливалась при обратном добавлении синтетических NLS в N-концевую часть dCas9-VPR с мутантным iNLS, свидетельствуя о том, что мутации в iNLS dCas9 не изменяют его функцию связывания ДНК (фиг. 35D).

Модульность химер дополнительно исследовали, тестируя другой ортолог Cas9 в аналогичном репортерном методе (H2B-цитрин, управляемый промотором 12xCSL, который также позволяет связывание NICD дикого типа). В конструкцию NC5 вводили dCas9 *Staphylococcus aureus*. Полученный при этом вариант NC5 с dCas9 *S. aureus* вместе со специфичной к *S. aureus* sgRNA (sasCSL) тоже проявлял уровень Delta-зависимой активации, аналогичный Notch1 дикого типа человека.

Активация рецепторов NC5 иммобилизованным Delta-лигандом ведет к отщеплению и транслокации в ядро dCas9-VPR. dCas9-VPR в комплексе со специфичной к последовательности sgRNA (например, sgTET) способствует связыванию комплекса с промотором (например, TET) и активации гена EGFP (фиг. 31G). На фиг. 31H представлены гистограммы интенсивности репортера EGFP в репортерных клетках HEK293T, стабильно экспрессирующих tet-индуцибельный ген EGFP и наводящую sgRNA (sgTET). Клетки, трансфицированные dCas9-VPR или рецептором NC5 (с ингибитором γ -секретазы или без него, DAPT), культивировали с иммобилизованным Delta или без него в течение 3 дней. Также указан процент активированных клеток (ON%). Клетки считаются включенными (ON) при интенсивности выше базальной интенсивности всех клеток без конструкции. На фиг. 31I представлены контурные графики (где между каждой парой контурных линий приходится одинаковое количество клеток) активации EGFP в репортерных клетках HEK293T, трансфицированных рецептором NC5 и культивируемых при различных концентрациях иммобилизованного Delta в течение 3 дней. EGFP не активировался в клетках, культивируемых без Delta, и в обработанных DAPT, культивируемых при наивысшей концентрации Delta. Указана средняя кратность изменений по сравнению с обработкой DAPT.

Сочетание способности направлять естественные внеклеточные сигналы с различительной способностью опосредованного sgRNA связывания Cas9 практически с любой геномной мишенью приоткрывает экспоненциальное количество сигнальных путей и функций (фиг. 32A). Раскрытые здесь химерные рецепторы Notch использовали для запуска нового Delta-зависимого способа остановки клеточного цикла. Клетки HEK293 стабильно трансдуцировали упакованными в лентивирусы sgRNAs, специфически нацеленными на ингибитор циклинзависимой киназы 1B (CDKN1B), эндогенный регулятор клеточного цикла. Было показано, что гиперэкспрессия CDKN1B вызывает остановку в G0/G1 (фиг. 32B). NC5 трансфицировали в клетки HEK293, стабильно экспрессирующие sgRNAs для CDKN1B (sgCDKN1B), а затем эти клетки вносили на голые или покрытые Delta поверхности. После 4-дневного культивирования на Delta наблюдалось >2-кратное повышение CDKN1B (по сравнению с голой поверхностью, нетрансфицированные контроли) в клетках, экспрессирующих NC5 (фиг. 32C).

Застрявшими в фазе G0/G1 оказались 38% обработанных Delta, трансфицированных клеток NC5 и sgCDKN1B⁺ и только 2-16% клеток во всех других исследованных условиях. Предполагается, что "протечка" NC5 в клетках sgCDKN1B⁺, культивируемых без Delta, могла быть частично обусловлена эндогенным DLL4 в клетках HEK293, который мог индуцировать экспрессию CDKN1B при контакте между клетками (фиг. 32C, условие "NC5 + sgRNA"), давая фенотипы, близкие клеткам, находившимся в контакте с адсорбированным на поверхности DLL4 с большим сродством, но в несколько меньшей пропорции.

Подход с химерным рецептором Notch, приведенный в этом примере, позволяет модулировать экспрессию генов и клеточные функции в ответ на внеклеточные сигналы. Сочетая модульную замену NICD на приведенные здесь ортологи Cas9, обладающие способностью прицельного воздействия на геном (например, различные распознающие PAM мотивы), с недавней модульной заменой NICD на одноцепочечные варибельные фрагменты, распознающие другие внеклеточные лиганды, химерные рецепторы Notch смогут подключаться к большему числу внеклеточных сигналов помимо естественных лигандов Notch и с более широким охватом генома. Предполагается, что химерные рецепторы Notch должны быть широко применимы для изучения и манипулирования эндогенными сигнальными путями с выходом на отдельные или множественные нисходящие компоненты, а также для клеточной терапии с зондированием внеклеточного микроокружения.

Методы.

Получение генетических конструкций. Для сборки всех конструкций, описанных в этом примере, применялись стандартные методы молекулярного клонирования (табл. 6 и 8). Все химерные рецепторы Notch клонировали в векторе pcDNA3 под управлением промотора CMV с маркером устойчивости к гентимицину. Все конструкции sgRNA клонировали в векторе pHR с промотором U6 и с маркером устойчивости к пурамицину либо с флуоресцентными маркерами.

Получение стабильных клеточных линий. Все клеточные линии, приведенные на фиг. 31A-31F, основывались на линии клеток T-Rex-CHO-K1 (Invitrogen). Получали стабильные клоны клеток UAS-H2B-цитрин и 12xCSL-H2B-цитрин. Дополнительные линии клеток получали при лентивирусной трансдук-

ции (всех sgRNAs и SV40-EGFP) или плазмидной трансфекции с помощью реагента TransIT-LT1 (Mirus Bio) согласно инструкциям производителя, с последующим отбором на соответствующем антибиотике и поддержанием. Стабильно трансфицированные клоны, экспрессирующие NC5 и sgUAS в линии UAS-H2B-цитрин, выделяли методом FACS после 2-недельного отбора на пурамицине и генетицине. Все линии клеток CHO культивировали в среде альфа-МЕМ с солями Эрла (Irvine Scientific) с добавлением 10% FBS специально для системы Tet (Clontech), по 100 ед./мл пенициллина и стрептомицина, 2 мМ L-глутамин (Gibco) и 10 мкг/мл бластицидина (VWR International) при 37°C с 5% CO₂ в увлажняемом инкубаторе.

Получение лентивируса и трансдукция. Для упаковки лентивируса использовали клетки HEK293T (ATCC). Клетки содержали в среде DMEM High Glucose with Gluta-MAX™ (Thermo Fisher) с добавлением 10% FBS специально для системы Tet (Clontech) и по 100 ед./мл пенициллина и стрептомицина (Gibco) при 37°C с 5% CO₂. В 1-й день клетки высевали при 2,0-3,0×10⁵ клеток/мл в 6-луночные планшеты (Corning). На 2-й день клетки были конфлюентными на 50-70% во время трансфекции. Для каждой лунки смешивали 1,51 мкг/мл плазмидного вектора pHR, 1,32 мкг dR8.91 и 165 нг pMD2.G (Addgene) в 250 мкл среды Opti-МЕМ с пониженным уровнем сыворотки (Gibco) с 7,5 мкл реагента TransIT-LT1 и инкубировали при комнатной температуре 15-30 мин. Раствор трансфекционного комплекса вносили поровну в культуры HEK293T по каплям. На 3-й день среду заменяли свежей средой для последующей трансдукции (например, альфа-МЕМ для линий клеток CHO). На 4-й день собирали лентивирус из супернатанта стерильным шприцем и фильтровали через фильтр из поливинилиденфторида на 0,45 мкм (Millipore) в криофлаконы для хранения при -80°C или немедленной трансдукции намеченных культур клеток.

Профильтрованные супернатанты лентивируса смешивали 1:1 с соответствующей свежей средой, чтобы заменить среду целевых клеток для трансдукции. Адгезированные культуры клеток подвергали трансдукции при конфлюентности в 50%. Добавляли Polybrene (Millipore) при 5 мкг/мл. Для встраивания sgRNA через 2 дня после трансдукции заменяли среду на среду с соответствующим антибиотиком для отбора; обычно методом проточной цитометрии отмечалось, что через 2 дня после трансдукции по меньшей мере 90% клеток содержат стабильно встроенные sgRNAs. Антибиотики сохраняли при последующем культивировании клеток. Трансдуцированные линии клеток использовали для экспериментов после того, как при отборе на антибиотиках погибали все нетрансдуцированные контрольные клетки и после одного пересева. Клетки не проверяли на загрязнение микоплазмой.

Экспериментальные методы. Адсорбирование Delta на поверхности: описанные в этом примере эксперименты проводили с клетками, засеянными в 24- или 48-луночные планшеты (Corning) при 1,5×10⁴ клеток/мл. На поверхности планшетов адсорбировали Delta при насыщающей концентрации в 2,15 мкг/мл, используя вариант DLL4 с повышенным сродством (E12) с меткой из 8 гистидинов (SEQ ID NO: 92), и инкубировали 2 ч при 37°C перед высевом клеток. Для экспериментов с клетками CHO также адсорбировали фибронектин хомяка при 5 мкг/мл (Innovative Research) вместе с Delta.

Методы проточной цитометрии. Клетки обрабатывали трипсином и анализировали на флуоресценцию репортера или иммунофлуоресценцию белков на анализаторе Scanford FACScan (Becton Dickinson) по стандартной методике. Для внутриклеточных белков клетки обрабатывали трипсином и фиксировали в 4% растворе параформальдегида в течение 15 мин, а затем инкубировали 2 ч при комнатной температуре с соответствующим первичным антителом в 0,3% растворе Triton-X100 (Sigma) с 2% нормальной ослиной сыворотки (Thermo Fisher) в PBS. Использовали первичное антитело против CDKN1B (D69C12, Cell Signaling) в рекомендованных изготовителем концентрациях. Затем образцы клеток инкубировали 30 мин с соответствующими видоспецифичными вторичными ослиными антителами с Alexa 647 (Life Technologies) при разведении 1:1000. В качестве отрицательного контроля использовали клетки, окрашенные только лишь вторичными антителами.

Таблица 6

Плазмидные конструкции

Фигура	Конструкция	Промотор	Продукт	Отбор клеток
Фиг. 31С, фиг. 31D	pHR-sgUAS-puro-t2a-BFP	U6 и EF1a	sgUAS и BFP	пурамицин (5 мкг/мл)
	pHR-sasgCSL-mCherry	U6 и CMV	sasgCSL и mCherry	сортинг по флуоресценции
	pHULK-EF1a-hNotch1-tagBFP	EF1a	hNotch1-tagBFP	гентицин (600 мкг/мл)
	pcDNA3-hNECD-SpdCas9-VPR-mCherry	CMV	NC5 <i>S. pyogenes</i>	гентицин (600 мкг/мл)
	pcDNA3-hNECD-SadCas9-VPR-mCherry	CMV	NC5 <i>S. aureus</i>	гентицин (600 мкг/мл)
Фиг. 31Е, фиг. 31F	pHR-EGFP	SV40	EGFP	сортинг по флуоресценции
	pHR-sgEGFP	U6	sgEGFP	пурамицин (5 мкг/мл)
	pcDNA3-hNECD-SpCas9-BFP	CMV	hNECD-SpCas9-BFP	гентицин (600 мкг/мл)
Фиг. 32	pHR-sgCDKN1B-puro	U6 и CMV	sgCDKN1B	пурамицин (5 мкг/мл)
Фиг. 34	pcDNA3-SpdCas9-VPR-mCherry	CMV	dCas9-VPR-mCherry <i>S. pyogenes</i>	гентицин (600 мкг/мл)
	pcDNA3-SpdCas9-mut-iNLS-VPR-mCherry	CMV	dCas9-VPR-mCherry <i>S. pyogenes</i> с мутантным iNLS	гентицин (600 мкг/мл)
Фиг. 35	pHR-Tet-EGFP	TetO	EGFP	сортинг по флуоресценции

Таблица 7

Стабильные линии клеток

Фигура	Стабильная линия клеток	Исходная линия	Конструкция при трансфекции/трансдукции	Отбор на антибиотиках
Фиг. 31С	sgUAS + UAS-H2B	UAS-H2B-цитрин	pHR-sgUAS-puro-t2a-BFP	бластицидин (10 мкг/мл), зеоцин (400 мкг/мл), пурамицин (5 мкг/мл)
	sasgCSL + CSL-H2B	12xCSL-H2B-цитрин	pHR-sasgCSL-mCherry	бластицидин (10 мкг/мл), зеоцин (400 мкг/мл)
Фиг. 31D	<i>S. pyogenes</i> NC5	sgUAS-BFP + UAS-H2B	pcDNA3-hNECD-SpdCas9-VPR-mCherry	бластицидин (10 мкг/мл), зеоцин (400 мкг/мл), гентицин (600 мкг/мл)
	<i>S. aureus</i> NC5	sasgCSL + CSL-H2B	pcDNA3-hNECD-SadCas9-VPR-mCherry	бластицидин (10 мкг/мл), зеоцин (400 мкг/мл), гентицин (600 мкг/мл)
	hNotch1	12xCSL-H2B-цитрин	pHULK-EF1a-hNotch1-tagBFP	бластицидин (10 мкг/мл), зеоцин (400 мкг/мл), гентицин (600 мкг/мл)
Фиг. 31Е, фиг. 31F	EGFP	T-Rex-CHO-K1	pHR-EGFP	бластицидин (10 мкг/мл)
	sgEGFP + EGFP	EGFP	pHR-sgEGFP-puro	бластицидин (10 мкг/мл), пурамицин (5 мкг/мл)
Фиг. 32	sgCDKN1B-HEK	HEK293T	pHR-sgCDKN1B-puro	пурамицин (5 мкг/мл)

Отдельные последовательности

SEQ ID NO:	Название	Последовательность
33	стандартный скелет sgRNA <i>S. pyogenes</i>	NNNNNNNNNNNNNNNNNNNNNGTTTAAAGAGC TATGCTGGAAACAGCATAGCAAGTTTAAATA AGGCTAGTCCGTTATCAACTGAAAAAGTGG CACCGAGTCGGTGCTTTTTTT
34	стандартный скелет <i>sa</i> sgRNA <i>S. aureus</i>	NNNNNNNNNNNNNNNNNNNNNNNGTTTATAG TACTCTGGAAACAGAATCTACTAAAACAAG GCAAAATGCCGTGTTTATCTCGTCAACTGT TGGCGAGATTTTT
35	sgUAS (число сайтов связывания для про- мотора UAS: 4)	GTACTCCGACCTCTAGTGT
36	<i>sa</i> sgCSL (число сайтов связывания для промо- тора 12xCSL: 6)	GGTGCCCTTCCGCCCATTTTCCC
37	sgEGFP	GACCAGGATGGGCACCACCC
38	sgCDKN1B	GGCTGGCGAGCGCGGCCTTA

Для регуляции генов путем редактирования генов в некоторых случаях может использоваться каталитически активный Cas9. На фиг. 36A представлено Delta-зависимое разрезание ДНК у варианта NC5: hNECD, слитого с Cas9 *S. pyogenes* с активной нуклеазой дикого типа (hNECD-Cas9). На конститутивно экспрессируемый трансген EGFP была нацелена sgEGFP в качестве sgRNA. На фиг. 36B представлена эффективность hNECD-Cas9 в клетках CHO со стабильно встроенным геном EGFP с промотором SV40 и нацеливающей sgRNA (например, sgEGFP). Среди положительных по hNECD-Cas9 и sgEGFP клеток, которые подвергались воздействию Delta при 4-дневном культивировании, отмечалась значительная доля EGFP-отрицательных клеток. Сверху представлены репрезентативные графики плотности по боковому рассеянию света (SSC) по отношению к EGFP при различных условиях после 4-дневного культивирования. Клетки считались EGFP-отрицательными, если их интенсивность попадала ниже порога, установленного в 99% EGFP-положительных клеток CHO без sgEGFP, но трансфицированных hNECD-Cas9. Внизу представлен количественный анализ EGFP-отрицательных клеток (n=3). Среднее \pm SEM. ***p<0,001.

На фиг. 36C представлена типичная схема для двух sgRNAs (короткие прямоугольники слева и справа от "sgCXCR4"), которые стабильно экспрессируются в клетках HEK293T и служат для наведения на 5'-нетранслируемый участок (UTR) и интрон-1 у CXCR4. Отметка масштаба - 1000 п.о. На фиг. 36D сверху представлены типичные результаты по эндонуклеазе T7E1 для анализа степени индуцируемой Delta опосредованной hNECD-Cas9 модификации гена CXCR4 в клетках HEK293T при детектировании по количеству расщепленных T7E1 продуктов в гелях SDS-PAGE. Внизу представлены типичные результаты по частоте мутаций indel у CXCR4 при оценке по соотношению расщепленных и нерасщепленных продуктов (n=3 независимых эксперимента). Среднее \pm SEM. *p<0,05 по сравнению со всеми другими условиями. На фиг. 36E представлены типичные результаты по количественному определению экспрессии белка CXCR4 в клетках HEK293T методом иммунофлуоресцентного окрашивания на основе проточной цитометрии после 4-дневного культивирования в указанных условиях (n=3 независимых эксперимента). Среднее \pm SEM. *p<0,05 по сравнению со всеми другими условиями.

Пример 8. Разработка минимального варианта рецептора Notch (NC5).

В некоторых случаях может потребоваться разработать минимальный вариант рецептора NC5 для применения *in vivo*. Особый интерес представляет NECD, который состоит из ~1700 аминокислот (по сравнению с dCas9 из ~1300 аминокислот). Домен NECD состоит из множественных повторов типа эпидермального фактора роста (EGF); функциональный анализ показал, что повторы 11 и 12 типа эпидермального фактора роста (EGF) являются важным звеном связывания Delta. Филогенетически повторы EGF 11,12 очень консервативны. Совмещение повторов EGF 11,12 и их гомологов у человека, шпорцевой лягушки *Xenopus*, полосатого данио и дрозофилы показало наличие сильно консервативных консенсусных остатков (фиг. 37A). На фиг. 37B представлена схема полноразмерного NECD и серии делеционных вариантов, составленных для систематического определения функционально минимального варианта химерного рецептора NC5. Представлены гистограммы интенсивности репортера EGFP в клетках HEK293T, трансфицированных соответствующими минимальными вариантами рецептора NC5 (гистограммы, справа) при культивировании в присутствии Delta (в течение 3 дней, если не указано иначе). Делеция всех 36 повторов EGF приводит к полной потере Delta-активации (AEGF, нулевой). У дрозофилы сообщалось о функционально минимальном варианте Notch при сохранении только EGF 11,12 (например, EGF(10-12] при обозначении интервалов согласно Rebay I. et al. Cell 67, 687-699 (1991)). Однако вариант NECD человека с EGF(10-12] проявлял снижение активации по сравнению с hNECD дикого типа (13,5%, фиг. 37B). Это означает, что другие повторы EGF в Notch человека могут способствовать стабильности взаимодействия Notch-Delta. Повторное введение 1-2 EGF-повторов, фланкирующих EGF

11,12, работало хуже, чем только лишь EGF 11,12, приводя к полной потере активации. Напротив, добавление дополнительных 3-5 EGF-повторов дает лучшее восстановление эффективности активации, причем вариант EGF(7-14] проявлял наибольшую степень активации (83,5%). Примечательно, что вариант EGF(7-14] составляет лишь одну треть длины NECD дикого типа.

Пример 9. Таргетинг генома и построение новых путей.

Одновременная активация двух генов в клетках HEK293T минимальными вариантами рецептора NC5 осуществлялась с помощью нескольких sgRNA (по 2 на ген) - наводящих на CXCR4 (sgCXCR4) и наводящих на CD95 (sgCD95). В присутствии "свободного" dCas9-VPR и sgCXCR4 или sgCD95 наблюдалось 5,6-кратное или 3,5-кратное повышение активности по сравнению с нецелевой sgRNA (например, sgNT). Одновременная активация CXCR4 и CD95 приводила к повышению в 3,7 и 4,6 раза по сравнению с sgNT, соответственно (фиг. 38). При использовании минимального варианта EGF(7-14] рецептора NC5 и обеих sgRNA не наблюдалось значительного повышения экспрессии обоих генов в клетках при культивировании без Delta, только с DAPT или с Delta и DAPT по сравнению с dCas9-VPR + sgNT. Когда эти клетки культивировали в присутствии Delta, то наблюдалась одновременная активация CXCR4 (повышение в 2,2 раза) и CD95 (повышение в 1,7 раз) по сравнению с клетками с dCas9-VPR + sgNT (фиг. 38).

Минимальные варианты рецептора NC5 использовали для запуска Delta-зависимой остановки клеточного цикла (фиг. 39A). Воздействовали на CDKN1B (ингибитор циклинзависимой киназы 1B) в клетках HEK293T с помощью sgRNAs (sgCDKN1B). Гиперэкспрессия CDKN1B приводит к остановке клеток в фазе G0/G1 (фиг. 39B). В клетках со "свободным" dCas9-VPR и sgCDKN1B повышающая регуляция CDKN1B сопровождается наращиванием G0/G1, а минимальное повышение CDKN1B наблюдается в клетках с dCas9-VPR и нецелевой sgRNA (sgNT, фиг. 39C и 39D). Индуцированная Delta повышающая регуляция CDKN1B и остановка клеток в фазе G0/G1 устраняется DAPT (фиг. 39E и 39F).

Хотя предпочтительные воплощения настоящего изобретения были представлены и описаны здесь, специалистам в данной области должно быть ясно, что такие воплощения приводятся только в качестве примера. Теперь у специалистов в данной области должны возникать многочисленные вариации, изменения и замены, не отходящие от изобретения. Следует понимать, что при осуществлении изобретения могут использоваться различные альтернативы описанных здесь воплощений изобретения. Предполагается, что объем изобретения определяется нижеприведенной формулой изобретения, причем она охватывает способы и структуры, входящие в рамки этой формулы и её эквивалентов.

Неформальный список последовательностей

SEQ ID NO:1

sgUAS

gtactccgacctctagtg

SEQ ID NO:2

sgTET

gtacggttctctatcactgata

SEQ ID NO:3

CXCR4-dCas9-VPR-mCherry [CXCR4-хвост V2-сайт расщепления TEV-dCas9-VPR-NLS-mCherry-рег HA-линкер]

MKTHIALSYIFCLVFADYKDDDDASIDMEGSIYTSNYTEEMGSGDYDSMKEPCFREANANFNKIF
 LPTIYSIIFLTGIVGNGLVILVMGYQKLRSMYRLHLSVADLLFVITLFWAVDAVANWYFGNFKCA
 VHVIYTVNLYSSVLILAFISLDRYLAIHVATNSQRPRKLLAEKVYVGVWIPALLTIPDFIFANVSEADRY
 ICDRFYPNDLVVVFQFQHIMVGLILPGIVILSCYCIISKLSHSHKQKALKTTVILILAFFACWLPYYIGI
 SIDSFILLEIKQGFENTVHKWISITEALAFFHCCLNPILYAFLGAKFKTSAQHALTSVSRGSSKILSKGKR
 GGHSSVSTESSESSFHSSIDTGGRTPPSLGPQDESCCTASSSLAKDTSSTGENLYFQLEMDKKYSIGLAIGTN
 SVGWAIVITDEYKVPKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRNRCYQLQEIF
 SNEMAKVDDSFHRLLEESFLVEEDKKHERHPFGNIVDEVAYHEKYPTIYHLRKKLV DSTKADLRLIYLAL
 AHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPG
 EKKNGFLGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADFLAAKNLSDAILL
 SDILRVNTEITKAPLSASMIKRYDEHHQDLTLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYK
 FIKPILEKMDGTEELLVKLNREDLLRQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTRIP
 YYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVLPHKSLLEYEFT
 VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS
 LGTYHDLKIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGR
 LSRKLINGIRDKQSGKTILDFLKSDFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIK
 KGILQTVKVVDELVKVMGRHKPENIVEMARENQTTQKQKNSRERMKRIEIGIKELGSQILKEHPVENTQ
 LQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRSDKNRKGSDNVPSEEV
 VKMKMNYWRQLLNAKLITQRKFDNLTKAERGLSELDKAGFIKQQLVETRQITKHVAQILD SRMNTKYDE
 NDKLIREVKVITLKSCLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVGTA LIKKYPKLESEFYGDYKYV
 YDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
 SMPQVNI VKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKG GFDSPVAYSVLVAKVEK GKSKKL
 KSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIILPKYSLFELNGRKRMLASAGELQKGNELALPS
 KYVNFLYLASHYEKLGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDK VLSAYNKHRDKPI
 REQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYSTKEVLDATLIHQSIITGLYETRIDLSQLGGDAYPYDVP
 DYASLGSQDGI GSGSNGSSLDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLDM
 LGSPKRRKRVGSQYLPDTPDRHRIEERKRYETFKSIMKKS PFGPTDPRPPRRRIA VPSRSSASVPKPAPQ
 PYPFTSSLSTINYDEFPTMVFPSPGQISQASALAPAPPQVLPQAPAPAPAMVSALAQAPAPVPLAPGPPQA

VAPPAPKPTQAGEGTLSEALLQLQFDEDEDLALLGNSTDPVFTDLASVDNSEFQQLLNQGIPVAPHTTEP
MLMEYPEAITRLVTGAQRPPDPAPAPLGPGLPNGLLSDGDEDFSSSIADMDFSALLSQISSGSGSRSRDSREG
MFLPKPEAGSAISDVFEGRVQCQPKRIRPFHPPGSPWANRPLASLAPTPTGPVHEPVGSLTPAPVPQPLDPA
PAVTPPEASHLLEDPEETSQAVKALREMADEVIPQKEEAACGQMDLSHPPRGLHDELTTTLESMTEDLN
LDSPLTPELNEILDFTLNDECLLHAMHISTGLSIFDTSFLMLVSKGEEDNMAIIEFMRFKVHMEGSVNGHEF
EIEGEGEGRPYEGTQAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDYKLSFPEGFKWERVM
NFEDEGGVVTVTQDSSLQDGEFIYKVKLRGTNFPDGPVMQKKTMGWEASSERMYPEDGALKGEIKQRLK
LKDGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDYTIIVEQYERAEGRHSTGGMDELYK

SEQ ID NO:4

CXCR4

MKTIIALSIIIFCLVFADYKDDDDASIDMEGSIYTSNDNYTEEMSGSDYDSMKEPCFREANANFNK
IFLPTIYSIIIFLTGIVGNGLVILVMGYQKKLRSMTDKYRLHLSVADLLFVITLPPFWAVDAVANWYFGNFL
CKAVHVIYTVNLYSSVLILAFISLDRYLAIIVHATNSQRPRKLLAEKVYVYVGVWIPALLLTIPDFIFANVS
EADDRYICDRFYPNLWVVVFQFQHIMVGLILPGIVILSCYCIISKLSHSHKGHQKRKALKTTVILILAF
FACWLPYYIGISIDSFILLEIKQGCDFENTVHKWISITEALAFPHCLNPIIYAFILGAKFKTSAQHALT
SVSRGSSLKILSKGKRGGHSSVSTESSESSFHSS

SEQ ID NO:5

Хвост V2

GRTPPSLGPQDESCTTASSSLAKDTSS

SEQ ID NO:6

Сайт расщепления TEV

ENLYFQL

SEQ ID NO:7

dCas9

MDKKYSIGLAIGTNSVGVAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLK
RTARRRYTRRRNRCYLQEIFSNEMAKVDDSEFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTI
YHLRKKLVDSITDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASG
VDAKAILSARLSKSRLENLIAQLPGEKKNLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDL
DNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIH
LGELHAILRRQEDFYPFKDNREKIEKILTFRIPIYVVGPLARGNSRFAMWTRKSEETITPWNFEVVDKG
ASAQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKT
NRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTL
FEDREMIERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMQ
LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKGIILQTVKVVDELVKVMGRHKPENIVEMAR

ENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSD
YDVDAIVPQSFLKDDSIDNKVLTRSDKNRSGSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAER
GGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVTALIKKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNF
FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VVKTEVQTGGFSKESILPKRN
SDKLIARKKDWDPKPYGGFDSPTVAVSVLVAKVEKGSKLLKSVKELLGITIMERSSEKNPIDFLEAK
GYKEVKKDLIIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLGKSPEDNEQK
QLFVEQHKHYLDEIEEQISEFSKRVI LADANLDKVL SAYNKHRDKPIREQAENI IHLFTLTNLGAPAAFK
YFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDLSQLGGD

SEQ ID NO:8

VPR (NLS жирным, линкер подчеркнут)

DALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSD**PKKKRKV**GSQYLP
DTDDRHRIEEKRRKTYETFKSIMKSPFSGPTDRPPPRRIAVPSRSSASVPKPAPQYPFTSSLSTINY
DEFPTMVFPSGQISQASALAPAPPQVLPQAPAPAPAPAMVSALAQAAPVPVLAAGPPQAVAPPAPKPTQ
AGEGTLSEALLQLQFDEDLGALLGNSTDPAVFTDLASVDNSEFQQLLNQGI PVAPHTTEPMLMEYPEAI
TRLVTGAQRPPDPAPAPLGPAGLPNGLLSGDEDFSSIADMDFSALLSQISSGSGSGSRDSREGMFLPKPE
AGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVT
PEASHLLEDPDEETSQAVKALREMADTVIPQKEEAAICGQMDLSHPPPRGHLDELTTLESMTEDLNLDLSD
PLTPELNEILDFTFLNDECLLHAMHISTGLSIFDTSLF

SEQ ID NO:9

NLS

PKKKRKV

SEQ ID NO:56

mCherry

MVSKGEEDNMAIIKEFMRFKVHMEGVSNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDI
LSPQFMYGSKAYVKHPADIPDYLKLSFPEGFKWERVMNFEDGGVVTQTQDSSLQDGEFIYKVKLRGTNFP
SDGPMQKKTMGWEASSERMPEDGALKGEIKQRLKLDKGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLD
ITSHNEDYTIVEQYERAEGRHSTGGMDELYK

SEQ ID NO:10

Ter HA

YPYDVPDY

SEQ ID NO:11

Линкер

GIGSGSNGSS

SEQ ID NO:12

ITGB1-dCas9-VPR-mCherry [интегрин β 1-сайт расщепления TEV-dCas9-VPR-NLS-mCherry-
тег HA-линкер]

MNLQPIFWIGLISSVCCVFAQTDENRCLKANAKSCGECIQAGPNCGWCTNSTFLQEGMPT SARCD
 DLEALKKKGCPPDDIENPRGSKDIKKKNVNTNRSKGTAEKLPEDITQIQPQQLVLRRLRSGEFPQTFTLKF
 KRAEDYPIDLYYLMDSLYSMKDDLENVKSLGTDLMNEMRRITSDFRIGFGSFVEKTVMPYISTTPAKLRN
 PCTSEQNCTSPFSYKNVLSLTNKGEVFNELVGKQRI SGNLDSPEGGFDAIMQVAVCGSLIGWRNVTRLLV
 FSTDAGFHFAGDGKLGIVLPNDGQCHLENNMYTMSHYDYPSIAHLVQKLENNIQTIFAVTEEFQPVY
 KELKNLIPKSAVGTLSANSSNVIQLIIDAYNSLSSEVILENGKLESGVTISYKSYCKNGVNGTGENGRKC
 SNISIGDEVQFEISITSNKCPKKDSDSFKIRPLGFTTEEVEVILQYICVCECQSEGIPESPKCHEGNGTFE
 CGACRCNEGRVGRHCECSTDEVNSEDMDAYCRKENSSEICSNNGECVCGQCVCRRKRDNTNEIYSGKFCEC
 DNFNCDRNSGLICGGNGVCKCRVCECNPNYTGSA CDCSLDTSTCEASNGQICNGRGICEGVCCKCTDPKF
 QGQTCEMCQTCLGVCAEHKECVQCRAFNKGEKKTCTQEC SYFNITKVESRDKLPQVPQDPVSHCKEKD
 VDCCWFYFTYSVNGNNEVMHVVENPECPGPDIIPIVAGVVAGIVLIGLALLLIWKLMIIDHRRFAK
 FEKEKMNKAWDTGENPIYKSAVTTVVNPKYEGKENLYFQLASMDKKYSIGLAIGTNSVGVAVITDEYKVP
 SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSEFF
 HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRILIYALAHMIKFRGHF
 LI EGDLPNDNSVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLLENLIAQLPGEKKNGLF
 GNLIALSGLTPNFKSNFDLAEDAKQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKP
 ILEKMDGTEELLVKNLREDLLRQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRI
 PYYVGPLARGNSRFAWMTKSEETITPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYF
 TVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFN
 ASLGTYHDLKI IKDKDFLDNEENEDILEDIVLTLTFEDREMIERLKYAHLFDDKVMKQLKRRRYTG
 WGRLSRKLINGIRDKQSGKTI LDFLKSDFANRNFQMLIHDDSLTFKEDIQKAQVSGGDSLHEHIANLA
 GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGKELGSQILKE
 HPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLRSDKNRGKSD
 NVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIILDSR
 MNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG TALIKKYPKLESEF
 VYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR
 DFATVRKVL SMPQVNI VVKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
 VEKGSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAG
 ELQKGNELALPSKYVNFYLYLASHYEKLGSPEDNEQKQLFVEQHKHYLDEIEQISEF SKRVILADANLD
 KVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYET
 RIDLSQLGGDAYPYDVPDYASLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDF
 LDMLGSPKKRKRKVGSYLPD TDRHRIEKRKRTYETFKSIMKKS PFSGPTDPRPPPRRIAVPSRSSAS

VPKPAPQYPFTSSSLSTINYDEFPTMVFPSPGQISQASALAPAPPQVLPQAPAPAPAPAMVSALAQAAPAPV
FVLAPGPPQAVAPPAPKPTQAGEGLTSEALLQLQFDDDELGALLGNSTDPVFTDLASVDNSEFQQLLNQ
GIPVAPHTTEPMLMEYPEAITRLVTGAQRPPDPAPAPLGAAPLGNLLSGDEDFSSIADMDFSALLSQIS
SGSGSGSRDSREGMFLPKPEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEP
VGSLTAPVPQPLDPAPAVTPEASHLLEDPEETSQAVKALREMA DTVI PQKEEAAICGQMDLSHPPRG
HLDELTTTLESMTEDLNLDSPLTPELNEILDFTLNDECLLHAMHISTGLSIFDTSLFLMVSKGEEDNMAI
IKEFMRFKVHMEGVSNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPA
DIPDYLLKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTFNFPDGPVMQKKTMGWEASS
ERMPEDGALKGEIKRQLKLDGGHYDAEVKTTYKAKKPVQLPGAYNVNLIKLDITSHNEDYTIVEQYERA
EGRHSTGGMDELYKLE

SEQ ID NO:13

Интегрин β_1

MNLQPIFWIGLISSVCCVFAQTDENRCLKANAKSCGECIQAGPNCGWCTNSTFLQEGMPTSARCD
DLEALKKKKCPDDIENPRGSKDIKKKNVNTNRSGTAEKLPEDITQIQPQQVLVRLRSGEPTFTLKF
KRAEDYPIDLYLMDLSYSMKDDLENVKS LGTDLNMNEMRRITSDFRIGFGSFVEKTVMPYISTTPAKLRN
ECTSEQNCTSPFSYKNVLSLTNKGEVFNELVGKQRISGNLDSPEGGFDAIMQVAVCGSLIGWRNVTRLLV
FSTDAGFHAFAGDKLGGIVLPNDGQCHLENNMYTMSHYDYPSIAHLVQKLSENNIQTI FAVTEEFQPVY
KELKNLIPKSAVGTLSANSSNVIQLIIDAYNSLSSEVILENGKLESEGVTISYKSYCKNGVNGTGENGRKC
SNISIGDEVQFEISITSNKCPKKDSDFKIRPLGFTTEEVEVILQYICVCECQSEGIPESPKCHEGNGTFE
CGACRCNEGRVGRHCECSTDEVNSEDMDAYCRKENSSEICSNNGECVCGQCVCRKRDNTNEIYSGKFCEC
DNFNCDRSNGLICGGNGVCKCRVCECNPNYTGSAACDCSLDTSTCEASNGQICNGRGICEGVCKCTDPKF
QGQTCEMCQTC LGVCAEHKECVQCRAFNGEKKDTCTQEC SYFNITKVESRDKLPQVPVQDPVSHCKEKD
VDCCWFYFTYSVNGNNEVMVHVVENPECPTGPDIIPIVAGVVAGIVLIGLALLLIWKLMI IHDRREFAK
FEKEKMNAKWD TGENPIYKSAVTTVVNPKYEGK

SEQ ID NO:14

Сайт расщепления TEV

ENLYFQL

SEQ ID NO:15

dCas9

MDKKYSIGLAIGTNSVGVAVITDEYKVPSSKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLK
RTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTI
YHLRKKLV DSTKADLR LIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFTIQLVQTYNQLFEENPINASG
VDAKAIL SARLSKSRRENLI AQLPGEKKNGLFGNLI ALSLGLTPNFKSNFDLAEDAQLQLSKD TYDDDL
DNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIH

LGELHAILRRQEDFYFPLKDNREKIEKILTRIPYVVGPLARGNSRFAMWTRKSEETITPWNFEVVDKG
 ASAQSFIERMTNFDKNLPNEKVLPHKSLLEYEFTVYNELTKVKYVTEGMRKPAFLSAGEQKKAIVDLLFKT
 NRRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKI IKDKDFLDNEENEDILEDIVLTLTL
 FEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQ
 LIHDDSLTFKEDIQKAQVSGQDLSLHEHIANLAGSPAIKKGILOTVKVVDELVKVMGRHKPENIVIEMAR
 ENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSD
 YDVAIVPQSFLKDDSIDNKVLRSDKNRGSNDVPSSEVVKKMKNYWRQLLNAKLITQRKFDNLTKAER
 GGLSELKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKV
 REINNYHHHADAYLNAVGTALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNF
 FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIIVKKTVEVQTGGFSKESILPKRN
 SDKLIARKKDWDPKKYGGFDSPTVAYSVLVAKVEKGSKLLKSVKELLGITIMERSSFEKNPIDFLEAK
 GYKEVKKDLI IKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFYLYLASHYEKLGKSPEDNEQK
 QLFVEQHKHYLDEIEEQISEFSKRVLADANLDKVL SAYNKHRDKPIREQAENI IHLFTLTLNLAGAPAAFK
 YFDTTIDRKRYTSTKEVLDTLHQITGLYETRIDLSQLGGD

SEQ ID NO:16

VPR (NLS жирным, линкер подчеркнут)

DALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSPKRRKRVGSQYLP
 DTDDRHRIEEKRRKRYETFKSIMKKSPPFSGPTDRPPPRRIAVPSRSSASVPKPAPQFPFTSSSLSTINY
 DEFPTMVFPFGQISQASALAPAPPQVLPQAPAPAPAMVSALAQAPAPVPLAPGPPQAVAPPAPKPTQ
 AGEGLTSEALLQLQFDDEDLGNSTDPVFTDLASVDNSEFQQLLNQGI PVAPHTTEPMLMEYPEAI
 TRLVGTGAQRPPDPAPAPLGPGLPGLLGSDEDFSSIADMDFSALLSQISSGSGSGSRDSREGMFLPKPE
 AGSAISDVFEGREVCQPKRIRPFHPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLDPAPAVT
 PEASHLLEDPEETSQAVKALREMA DTVIPQKEEAAICGQMDLSHPPPRGHLDELTTLESMTEDLNLDS
 PLTPELNEILDFTLNDECLLHAMHISTGLSIFDTSLF

SEQ ID NO:17

NLS

PKKKRKV

SEQ ID NO:18

mCherry

MVSKGEEDNMAIIKEFMRFKVHMEGVSNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDI
 LSPQFMYGSKAYVKHPADIPDY LKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFIYKVKLRGTNFP
 SDGPFVMQKKTMGWEASSERMPEDGALKEIKQRLKLDKGGHYDAEVKTTYKAKKPVQLPGAYNVNLIKLD
 ITSHNEDYTIIVEQYERAEGRHSTGGMDELYKLE

SEQ ID NO:19

Ter HA
YPYDVPDYA

SEQ ID NO:20

Линкер
GSGSGS

SEQ ID NO:21

Notch-dCas9-VPR-mCherry [hNECD-TM-hNICD-dCas9-VPR-mCherry-тер HA-линкер]

MPPLLAPLLCLALLPALAARGPRCSQPGETCLNGGKCEAANGTEACVCGGAFVGFRCQDPNPCLS
TPCKNAGTCHVVDRRGVADYACSCALGFSGPLCLTPLDNACLTNPCRNGGTCDLLTLTEYKCRCPGWSG
KSCQQADRPCASNPCANGGQCLPFEASYICHCPPSFHGPTRQDVNECGQKPGLCRHGGTCHNEVGSYRCV
CRATHTGPNCEPFYVPCSPSPCQNGGTCRPTGDVTHECACLPGFTGQNCENIDDCPGNNCKNGGACVDG
VNTYNCRCPPPEWTGQYCTEDVDECQLMPNACQNGGTCNTHGGYNCVCVNGWTGEDCSENIDDCASAACF
HGATCHDRVASFYCECPHGRTGLLCHLNDAICISNPCNEGSNCDTNPVNGKAICTCPSGYTGPACSQDVDE
CSLGANPCEHAGKICINTLGSFECQCLQGYTGPRCEIDVNECVSNPCQNDATCLDQIGEFQCIOMPGEV
HCEVNTDECASSPCLHNGRCLDKINEFQCECPTGFTGHLQYDVDECASTPCKNGAKCLDGPNTYTCVCT
EGYTGTHCEVDIDECDDPCHYGSKDGVATFTCLCRPGYTGHHCEETNINECSSQPCRHHGGTCQDRDNAY
LCFCLKGTTGPNCEINLDDCASSPCDSGTCLDKIDGYECACEPGYTGSMCNINIDECAGNPCHNGGTCED
GINGFTCRCPEGYHDPCLSEVNECNSNPCVHGACRDSLNGYKDCDPCGWSGTNCDINNNECESNPCVNG
GTCKDMTSGYVCTCREGFSGPNCQTNINECASNPCLNQGTCIDDVAGYKCNCLLPYTGATCEVVLAPCAP
SPCRNGGECRQSEDEYEFSCVCPTGWQAGQTCEVDINECVLSPCRHGASCQNTHHGGYRCHCQAGYSGRNC
ETDIDDCRPNPCHNGGSCDGTGINTAFCDCLPGFRGTFCCEEDINECASDPCRNGANCTDCVDSYTCTCPAG
FSGIHENNTPDCTESSCFNGGTCVDGINSFTCLCPPGFTGSYQHDVNECDSQPCLHGGTCQDGGCSYR
CTCPQGYTGPNQCQLVHWCDSSEPCKNGGKCWQHTHTQYRCECPSGWTGLYCDVPSVSCVAAQRQGVDFVAR
LCQHGGGLCVDAGNTHHCRCQAGYTGSYCEDLVDECSPPCQNGATCTDYLGYSCKCVAGYHGVNCSEEI
DECLSHPCQNGGTCCLDLPTNYKCSQPRGTQGVHCEINVDDCNPPVDPVSRSPKCFNNGTCVDQVGGYSCT
CPPGFVGERCEGDVNECLSNPCDARGTQNCVQRVNDFHCECRAGHTGRRCESVINGCKGKPKKNGGTCVAV
ASNTARGFICKCPAGFEGATCENDARTCGSLRCLNGGTCISGPRSPCLCLGPFPTGPECQFPASSPCLGG
NPCYNQGTCEPTSESPFYRCLCPAKFNGLLCHILDYSFSGGAGRDI PPPLIEEACELPECQEDAGNKVCS
LQCNNHACGWDGGDCSLNFNDPWKNCTQSLQCWKYFSDGHCDSCNSAGCLFDGFDCQRAEQCNPLYDQ
YCKDHFSDGHCDQGCNSAECEWDGLDCAEHVPERLAAGTLVVVLMPEQLRNSSFHFLRELSRVLHTNV
VFKRDAHGQMIFFPYYGREEELRKHPIKRAAEGWAAPDALLGQVKASLLPGGSEGGRRRRELDPMDFVRS
IVYLEIDNRQCQVASSQCFQSATDVAAFLGALASLGLNIPYKIEAVQSETVEPPPPAQLHFMYVAAAAF
VLLFFVCGVLLSRKRRRASMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLAESFLVEEDKKHERHPIFG
NIVDEVAYHEKYPTIYHLRKKLVDSYDKADLRLLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQ
TYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAE

DAKLQLSKDITYDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQ
DLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFKIPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVVGPLARGNSRFAMWTRKSE
ETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLYEYFTVYNELTKVKYVTEGMRKPAFL
SGEQKKAIVDLLFKTRNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGYHDLKLIKDKDFLDNE
ENEDILEDIVLTLTLFEDREMIERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTIL
DFLKSDGFANRNFQMLIHDDSLTFKEDIQKAQVSGGDSLHEHIANLAGSPAIAKKGILQTVKVVDELVKV
MGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGKELGSQILKEHPVENTQLQNEKLYLYYLQNGR
DMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTNRSDKNRKGSDNVPSEEVVKMKNYWRQLLNAK
LITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKREINNYHHAHDAYLNAVVGTAI KKYPKLESEFVYGDYKVDVRKMI AKSEQEIG
KATAKYFFYSNIMNFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNI VKKTEV
OTGGFSKESILPKRNSDKLIARKKDWDPKYGGFDSPTVAYSVLVVAKVEKGSKKLKS VKELLGITIME
RSSFEKNPIDFLEAKGYKEVKKDLIKLPKYSLEFELENGRKRMLASAGELQKGNELALPSKYVNFYLLAS
HYEKLKSGSPEDNEQQLFVEQHKHYLDEIEQISEFSKRVI LADANLDKVL SAYNKHRDKPIREQAENII
HLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQSI TGLYETRIDLSQLGGDAYPYDVPDYASL
GSGDGI GSGSNGSSLDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDGSGS
GSQYLPDTPDRHRIEEKRRTYETFKS IMKKS PFSGP TDRPPRR IAVPSRSSASVPKPAQPYPFTSS
LSTINYDEFPTMVFPSGQISQASALAPAPPVLPQAPAPAPAPAMVSALAQAPAPVPVLAPGPQAVAPP
APKPTQAGEGTLSEALLQLQFDDEDL GALLGNSTDP AVFTDLASVDNSEFQQLLNQGI PVAPHTTEPMLM
EYPEAITRLVTGAQRPPDPAPAPL GAPGLPNGLLSGDEDFSSIADMDFSALLSQISSGSGS SRDSREGM
FLPKPEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQPLD
PAPAVTPEASHLLEDPEETSQAVKALREMA DTVI PQKEEAAICGQMDLSHP PPRGHDELTTTLESMTE
DLNLDSP LPELNEILD TFLNDECLLHAMHISTGLSIFDTSLFLMVSKGEEDNMAI I KEFMRFKVHMEGS
VNGHEFEIEGEGEGRPYEGTQTA KLKVTKGGPLPFAWDILSPQFMYGSKAYVKHPADIPDY LKLSFPEGF
KWERVMNFEDGGVVTVTQDSSLQDGEFYKVKLRGTNFPSDGPFVMQKKTMGWEASSERMY PEDGALKGEI
KQRLKLDKGGHYDAEVKTTYKAKKPVQLPGAYNVNIKLDITSHNEDY TIVEQYERAEGRHSTGGMDELYK
LE

SEQ ID NO:22

hNECD

MPPLLAPLLCLALLPALAARGPRCSQPGETCLNGGKCEAANGTEACVCGGAFVGPQCQDPNPCLS
TPCKNAGTCHVVD RRGVADYACSCALGFSGPLCLTPLDNA CLTNPCRNGGTC DLLTLTEYKCRCP PGWSG
KSCQQADPCASNPCANGGQLPFEASYICHCPPSFHGPTCRQDVNECGQKPGLCRHHGGTCHNEVGSYRCV
CRATH TGPNCERPVP CSPSPCQNGGTCRPTGDVTHECA CLPGFTGQNC EENIDDCPGNCKNGGACVDG
VNTYNCRC PPEWTGQYCTEDVDECQLMPNACQNGGTC HNTHGGYNCVCVNGW TGEDCSENIDDCASAA CF
HGATCHDRVASFYCECPHGRTG LLCHLN DACISNPCNEGSNCDTNPVNGKAICTCPSGYTGFACSQDVDE
CSLGANPCEHAGKCINTLGSFECQCLQGYTGPRCEIDVNECVSNPCQNDATCLDQIGEFQCICMPGYEGV

HCEVNTDECASSPCLHNGRCLDKINEFQCECPTGFTGHLQYDVDECASTPCKNGAKCLDGPNTYTCVCT
EGYTGTHCEVDIDECDPDPCHYGSKDGVATFTCLCRPGYTGHHHCETNINECSSQPCRHHGGTCQDRDNAY
LCFCLKGTGPNCEINLDDCASSPCDSGTCLDKIDGYECACEPGYTGSMCNINIDEACAGNPCHNGGTCD
GINGFTCRCPEGYHDPCLSEVNECNSNPCVHGACRDSLNGYKDCDCPGWSTNCDINNECESNPCVNG
GTCKDMTSGYVCTCREGFGSPNCQTNINECASNPCLNQGTCIDDVAGYKCNCLLPYTGATCEVVLAPCAP
SPCRNNGGECRQSEDIYESFSCVCPGTGWQAGQTCEVDINECVLSPCRHGASCQNTHHGGYRCHCQAGYSGRNC
ETDIDDCRPNPCHNGGSCDGTAFCDCLPGFRGTFCCEEDINECASDPCRNGANCTDCVDSYTCCTCPAG
FSGIHCENNTPDCTESSCFNGGTCVDGINSFTCLCPPGFTGSYCQHDVNECDSQPCLHGGTCQDGCOSYR
CTCPQGYTGPNCQNLVHWCDS SPCKNGGKCWQHTHTQYRCECPGWTGLYCDVPSVSCVAAQRQGV DVAR
LCQHGGGLCVDAGNTHHCRCQAGYTGSYCEDLVDECSPPCQNGATCTDYLGYSCKCVAGYHGVCSEEI
DECLSHPCQNGGTCLDLFNTYKCS CPRGTQGVHCEINVDDCNPPVDPVSRSPKCFNNGTCVDQVGGYSCT
CPPGFVGERCEGDVNECLSNPCDARGTQNCVQRVNDHFCECRAGHTGRRCESVINGCKGKPKNGGTCAV
ASNTARGFICKCPAGFEGATCENDARTCGSLRCLNGGTCISGPRSPTCLCLGPFTGPECQFPASSPCLGG
NPCYNQGTCEPTESESPFYRCLCPAKFNGLLCHILDYSFGGAGRDIPPLIEEACELPECQEDAGNKVCS
LQCNNHACGWDGGDCSLNFNDPWKNCTQSLQCWKYFSDGHCDSCNSAGCLFDGFDQRAEGQCNP LYDQ
YCKDHFSDGHCDQGCNSAECEWDGLDCAEHVPERLAAGTLVVVVLMPPEQLRNSSFHFLRELSRVLHTNV
VFKRDAHGQQMIFPYYGREEELRKHPKRAAEGWAAPDALLGQVKASLLPGGSEGGRRRRELDPM DVRGS
IVYLEIDNRQCQVASSQCFSATDVAAFLGALASLGS LNIPYKIEAVQSETVEPPPPAQ

SEQ ID NO:23

TM

LHFMYVAAAAFVLLFFVGC

SEQ ID NO:24

hNICD

VLLSRKRRR

SEQ ID NO:25

dCas9

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR
TARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIY
HLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDFLFIQLVQTYNQLFEENP INASGV
DAKAIL SARLSKSRLENLIAQLPGEKKNLFGNLIASLGLTPNFKSNFDLAE DAKLQLSKDTYDDDL
NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLKALVRQQLPEK
YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKNREDLLRKQRTFDNGSIPHQIHL
GELHAILRRQEDFYPF LKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGA
SAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN
RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHLLKIKDKDFLDNEENEDI EDIVLTLTLF

EDREMI EERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQQL
IHDDSLTFKEDIQKAQVSGQDLSLHEHTANLAGSPAIKKGIQTVKVVDELVKVMGRHKPENIVIEMARE
NQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDY
DVDAIVPQSFLKDDSIDNKVLRSDKNRGSNDNPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERG
GLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVR
EINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKSEQIEGKATAKYFFYSNIMNFF
KTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNS
DKLIARKKDWDPKKGFFDSPTVAYSVLVAVKEKGSKLLKSVKELLGITIMERSSEFKNPIDFLEAKG
YKEVKKDLIIKLPKYSLELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSPEDNEQKQ
LFVEQHKHYLDEIEEQISEFSCRVILADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFY
FDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDLSQLGGD

SEQ ID NO:26

VPR с подчеркнутыми последовательностями линкеров

DGIGSGSNGSSLDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSDALDDFDLMLGSG
GSGSQYLPDTPDRHRIEKRKRYETFKSIMKKSPFSGPTDPRPPRRRIAVPSRSSASVPKPAPQYFPT
SSLSTINYDEFPTMVFPSGQISQASALAPAPPVLPQAPAPAPAPAMVSALAQAPAPVPVLPAGPPQAVA
PPAPKPTQAGEGTLSEALLQLQFDEDLGALLGNSTDPAVFTDLASVDNSEFQQLLNQGI PVAPHTTEPM
LMEYPEAITRLVTGAQRPPDPAPAPLGAPGLPNGLLSGDEDFSSIADMDFSALLSQISSGSGSGRDSRE
GMFLPKPEAGSAISDVFEGREVCQPKRIRPFHPPGSPWANRPLPASLAPTPTGPVHEPVGSLTPAPVPQP
LDPAPAVTPEASHLLEDPEETSQAVKALREMA DTVIPQKEEAAICGQMDLSHPPRGLHDELTTTLESM
TEDLNLDSPFLTPELNEILDFTLNDECLLHAMHISTGLSIFDTSLF

SEQ ID NO:27

mCherry

MVSKGEEDNMAIIKEFMRFKVHMEGVSNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWDI
LSPQFMYGSKAYVKHPADIPDY LKLSFPEGFKWERVMNFEDGGVVTVTQDSSLQDGEFTIYKVKLRGTNFP
SDGPVMQKKTMGWEASSERMPEDGALKGEIKQRLKLDGGHYDAEVKTTYKAKKPVLPGAYNVNIKLD
ITSHNEDYTIVEQYERAEGRHSTGGMDELYKLE

SEQ ID NO:28

метка HA

YPYDVPDYA

SEQ ID NO:29

Линкер между меткой HA и VPR

DGIGSGSNGSS

ПЕРЕЧЕНЬ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

- <110> Совет попечителей Университета Лиланда Стэнфорда, мл.
<120> ХИМЕРНЫЕ БЕЛКИ И СПОСОБЫ РЕГУЛИРОВАНИЯ ЭКСПРЕССИИ ГЕНОВ
<130> 079445-1092326-000830JP
<140> JP 2018-536123
<141> 2018-07-10
<150> PCT/US2017/012885
<151> 2017-01-10
<150> 62/399,902
<151> 2016-09-26
<150> 62/351,522
<151> 2016-06-17

<150> 62/277,322

<151> 2016-01-11

<160> 92

<170> PatentIn version 3.5

<210> 1

<211> 19

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 1

gtactccgac ctctagtgt

19

<210> 2

<211> 21

<212> ДНК

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 2

gtacgttctc tatcactgat a

21

<210> 3

<211> 2575

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 3

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

Asp Tyr Lys Asp Asp Asp Asp Ala Ser Ile Asp Met Glu Gly Ile Ser
20 25 30

Ile Tyr Thr Ser Asp Asn Tyr Thr Glu Glu Met Gly Ser Gly Asp Tyr
35 40 45

Asp Ser Met Lys Glu Pro Cys Phe Arg Glu Glu Asn Ala Asn Phe Asn
50 55 60

Lys Ile Phe Leu Pro Thr Ile Tyr Ser Ile Ile Phe Leu Thr Gly Ile
65 70 75 80

Val Gly Asn Gly Leu Val Ile Leu Val Met Gly Tyr Gln Lys Lys Leu

048071

				85					90					95			
Arg	Ser	Met	Thr	Asp	Lys	Tyr	Arg	Leu	His	Leu	Ser	Val	Ala	Asp	Leu		
			100					105					110				
Leu	Phe	Val	Ile	Thr	Leu	Pro	Phe	Trp	Ala	Val	Asp	Ala	Val	Ala	Asn		
		115					120					125					
Trp	Tyr	Phe	Gly	Asn	Phe	Leu	Cys	Lys	Ala	Val	His	Val	Ile	Tyr	Thr		
	130					135					140						
Val	Asn	Leu	Tyr	Ser	Ser	Val	Leu	Ile	Leu	Ala	Phe	Ile	Ser	Leu	Asp		
145					150					155					160		
Arg	Tyr	Leu	Ala	Ile	Val	His	Ala	Thr	Asn	Ser	Gln	Arg	Pro	Arg	Lys		
				165					170						175		
Leu	Leu	Ala	Glu	Lys	Val	Val	Tyr	Val	Gly	Val	Trp	Ile	Pro	Ala	Leu		
			180					185					190				
Leu	Leu	Thr	Ile	Pro	Asp	Phe	Ile	Phe	Ala	Asn	Val	Ser	Glu	Ala	Asp		
		195					200					205					
Asp	Arg	Tyr	Ile	Cys	Asp	Arg	Phe	Tyr	Pro	Asn	Asp	Leu	Trp	Val	Val		
	210					215					220						
Val	Phe	Gln	Phe	Gln	His	Ile	Met	Val	Gly	Leu	Ile	Leu	Pro	Gly	Ile		
225					230					235					240		
Val	Ile	Leu	Ser	Cys	Tyr	Cys	Ile	Ile	Ile	Ser	Lys	Leu	Ser	His	Ser		
				245					250					255			
Lys	Gly	His	Gln	Lys	Arg	Lys	Ala	Leu	Lys	Thr	Thr	Val	Ile	Leu	Ile		
			260					265					270				
Leu	Ala	Phe	Phe	Ala	Cys	Trp	Leu	Pro	Tyr	Tyr	Ile	Gly	Ile	Ser	Ile		
		275					280					285					
Asp	Ser	Phe	Ile	Leu	Leu	Glu	Ile	Ile	Lys	Gln	Gly	Cys	Glu	Phe	Glu		
	290					295					300						
Asn	Thr	Val	His	Lys	Trp	Ile	Ser	Ile	Thr	Glu	Ala	Leu	Ala	Phe	Phe		
305					310					315					320		
His	Cys	Cys	Leu	Asn	Pro	Ile	Leu	Tyr	Ala	Phe	Leu	Gly	Ala	Lys	Phe		
				325					330					335			

048071

Lys Thr Ser Ala Gln His Ala Leu Thr Ser Val Ser Arg Gly Ser Ser
340 345 350

Leu Lys Ile Leu Ser Lys Gly Lys Arg Gly Gly His Ser Ser Val Ser
355 360 365

Thr Glu Ser Glu Ser Ser Ser Phe His Ser Ser Ile Asp Thr Gly Gly
370 375 380

Arg Thr Pro Pro Ser Leu Gly Pro Gln Asp Glu Ser Cys Thr Thr Ala
385 390 395 400

Ser Ser Ser Leu Ala Lys Asp Thr Ser Ser Thr Gly Glu Asn Leu Tyr
405 410 415

Phe Gln Leu Leu Glu Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile
420 425 430

Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val
435 440 445

Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile
450 455 460

Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala
465 470 475 480

Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg
485 490 495

Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala
500 505 510

Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val
515 520 525

Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val
530 535 540

Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg
545 550 555 560

Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr
565 570 575

Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu
580 585 590

048071

Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln
 595 600 605

Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala
 610 615 620

Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser
 625 630 635 640

Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn
 645 650 655

Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn
 660 665 670

Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser
 675 680 685

Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly
 690 695 700

Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala
 705 710 715 720

Ile Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala
 725 730 735

Pro Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp
 740 745 750

Leu Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr
 755 760 765

Lys Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile
 770 775 780

Asp Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile
 785 790 795 800

Leu Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg
 805 810 815

Glu Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro
 820 825 830

His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu
 835 840 845

048071

Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile
850 855 860

Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn
865 870 875 880

Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro
885 890 895

Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe
900 905 910

Ile Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val
915 920 925

Leu Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu
930 935 940

Leu Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe
945 950 955 960

Leu Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr
965 970 975

Asn Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys
980 985 990

Ile Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe
995 1000 1005

Asn Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys
1010 1015 1020

Asp Lys Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu
1025 1030 1035

Asp Ile Val Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile
1040 1045 1050

Glu Glu Arg Leu Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val
1055 1060 1065

Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu
1070 1075 1080

Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys Gln Ser Gly Lys

048071

1085						1090						1095			
Thr	Ile	Leu	Asp	Phe	Leu	Lys	Ser	Asp	Gly	Phe	Ala	Asn	Arg	Asn	
1100						1105					1110				
Phe	Met	Gln	Leu	Ile	His	Asp	Asp	Ser	Leu	Thr	Phe	Lys	Glu	Asp	
1115						1120					1125				
Ile	Gln	Lys	Ala	Gln	Val	Ser	Gly	Gln	Gly	Asp	Ser	Leu	His	Glu	
1130						1135					1140				
His	Ile	Ala	Asn	Leu	Ala	Gly	Ser	Pro	Ala	Ile	Lys	Lys	Gly	Ile	
1145						1150					1155				
Leu	Gln	Thr	Val	Lys	Val	Val	Asp	Glu	Leu	Val	Lys	Val	Met	Gly	
1160						1165					1170				
Arg	His	Lys	Pro	Glu	Asn	Ile	Val	Ile	Glu	Met	Ala	Arg	Glu	Asn	
1175						1180					1185				
Gln	Thr	Thr	Gln	Lys	Gly	Gln	Lys	Asn	Ser	Arg	Glu	Arg	Met	Lys	
1190						1195					1200				
Arg	Ile	Glu	Glu	Gly	Ile	Lys	Glu	Leu	Gly	Ser	Gln	Ile	Leu	Lys	
1205						1210					1215				
Glu	His	Pro	Val	Glu	Asn	Thr	Gln	Leu	Gln	Asn	Glu	Lys	Leu	Tyr	
1220						1225					1230				
Leu	Tyr	Tyr	Leu	Gln	Asn	Gly	Arg	Asp	Met	Tyr	Val	Asp	Gln	Glu	
1235						1240					1245				
Leu	Asp	Ile	Asn	Arg	Leu	Ser	Asp	Tyr	Asp	Val	Asp	Ala	Ile	Val	
1250						1255					1260				
Pro	Gln	Ser	Phe	Leu	Lys	Asp	Asp	Ser	Ile	Asp	Asn	Lys	Val	Leu	
1265						1270					1275				
Thr	Arg	Ser	Asp	Lys	Asn	Arg	Gly	Lys	Ser	Asp	Asn	Val	Pro	Ser	
1280						1285					1290				
Glu	Glu	Val	Val	Lys	Lys	Met	Lys	Asn	Tyr	Trp	Arg	Gln	Leu	Leu	
1295						1300					1305				
Asn	Ala	Lys	Leu	Ile	Thr	Gln	Arg	Lys	Phe	Asp	Asn	Leu	Thr	Lys	
1310						1315					1320				

048071

Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile
1325 1330 1335

Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val Ala
1340 1345 1350

Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp
1355 1360 1365

Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu
1370 1375 1380

Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu
1385 1390 1395

Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
1400 1405 1410

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu
1415 1420 1425

Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile
1430 1435 1440

Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe
1445 1450 1455

Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu
1460 1465 1470

Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly
1475 1480 1485

Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr
1490 1495 1500

Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys
1505 1510 1515

Thr Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro
1520 1525 1530

Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp
1535 1540 1545

Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser
1550 1555 1560

048071

Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu
 1565 1570 1575

Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser
 1580 1585 1590

Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr
 1595 1600 1605

Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser
 1610 1615 1620

Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala
 1625 1630 1635

Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr
 1640 1645 1650

Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly
 1655 1660 1665

Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His
 1670 1675 1680

Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser
 1685 1690 1695

Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser
 1700 1705 1710

Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu
 1715 1720 1725

Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala
 1730 1735 1740

Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr
 1745 1750 1755

Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile
 1760 1765 1770

Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly
 1775 1780 1785

Asp Ala Tyr Pro Tyr Asp Val Pro Asp Tyr Ala Ser Leu Gly Ser
 1790 1795 1800

048071

Gly Asp Gly Ile Gly Ser Gly Ser Asn Gly Ser Ser Leu Asp Ala
 1805 1810 1815

Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp
 1820 1825 1830

Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe
 1835 1840 1845

Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu
 1850 1855 1860

Asp Met Leu Gly Ser Pro Lys Lys Lys Arg Lys Val Gly Ser Gln
 1865 1870 1875

Tyr Leu Pro Asp Thr Asp Asp Arg His Arg Ile Glu Glu Lys Arg
 1880 1885 1890

Lys Arg Thr Tyr Glu Thr Phe Lys Ser Ile Met Lys Lys Ser Pro
 1895 1900 1905

Phe Ser Gly Pro Thr Asp Pro Arg Pro Pro Pro Arg Arg Ile Ala
 1910 1915 1920

Val Pro Ser Arg Ser Ser Ala Ser Val Pro Lys Pro Ala Pro Gln
 1925 1930 1935

Pro Tyr Pro Phe Thr Ser Ser Leu Ser Thr Ile Asn Tyr Asp Glu
 1940 1945 1950

Phe Pro Thr Met Val Phe Pro Ser Gly Gln Ile Ser Gln Ala Ser
 1955 1960 1965

Ala Leu Ala Pro Ala Pro Pro Gln Val Leu Pro Gln Ala Pro Ala
 1970 1975 1980

Pro Ala Pro Ala Pro Ala Met Val Ser Ala Leu Ala Gln Ala Pro
 1985 1990 1995

Ala Pro Val Pro Val Leu Ala Pro Gly Pro Pro Gln Ala Val Ala
 2000 2005 2010

Pro Pro Ala Pro Lys Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser
 2015 2020 2025

Glu Ala Leu Leu Gln Leu Gln Phe Asp Asp Glu Asp Leu Gly Ala

048071

2030 2035 2040

Leu Leu Gly Asn Ser Thr Asp Pro Ala Val Phe Thr Asp Leu Ala
 2045 2050 2055

Ser Val Asp Asn Ser Glu Phe Gln Gln Leu Leu Asn Gln Gly Ile
 2060 2065 2070

Pro Val Ala Pro His Thr Thr Glu Pro Met Leu Met Glu Tyr Pro
 2075 2080 2085

Glu Ala Ile Thr Arg Leu Val Thr Gly Ala Gln Arg Pro Pro Asp
 2090 2095 2100

Pro Ala Pro Ala Pro Leu Gly Ala Pro Gly Leu Pro Asn Gly Leu
 2105 2110 2115

Leu Ser Gly Asp Glu Asp Phe Ser Ser Ile Ala Asp Met Asp Phe
 2120 2125 2130

Ser Ala Leu Leu Ser Gln Ile Ser Ser Gly Ser Gly Ser Gly Ser
 2135 2140 2145

Arg Asp Ser Arg Glu Gly Met Phe Leu Pro Lys Pro Glu Ala Gly
 2150 2155 2160

Ser Ala Ile Ser Asp Val Phe Glu Gly Arg Glu Val Cys Gln Pro
 2165 2170 2175

Lys Arg Ile Arg Pro Phe His Pro Pro Gly Ser Pro Trp Ala Asn
 2180 2185 2190

Arg Pro Leu Pro Ala Ser Leu Ala Pro Thr Pro Thr Gly Pro Val
 2195 2200 2205

His Glu Pro Val Gly Ser Leu Thr Pro Ala Pro Val Pro Gln Pro
 2210 2215 2220

Leu Asp Pro Ala Pro Ala Val Thr Pro Glu Ala Ser His Leu Leu
 2225 2230 2235

Glu Asp Pro Asp Glu Glu Thr Ser Gln Ala Val Lys Ala Leu Arg
 2240 2245 2250

Glu Met Ala Asp Thr Val Ile Pro Gln Lys Glu Glu Ala Ala Ile
 2255 2260 2265

048071

Cys Gly Gln Met Asp Leu Ser His Pro Pro Pro Arg Gly His Leu
 2270 2275 2280

Asp Glu Leu Thr Thr Thr Leu Glu Ser Met Thr Glu Asp Leu Asn
 2285 2290 2295

Leu Asp Ser Pro Leu Thr Pro Glu Leu Asn Glu Ile Leu Asp Thr
 2300 2305 2310

Phe Leu Asn Asp Glu Cys Leu Leu His Ala Met His Ile Ser Thr
 2315 2320 2325

Gly Leu Ser Ile Phe Asp Thr Ser Leu Phe Leu Met Val Ser Lys
 2330 2335 2340

Gly Glu Glu Asp Asn Met Ala Ile Ile Lys Glu Phe Met Arg Phe
 2345 2350 2355

Lys Val His Met Glu Gly Ser Val Asn Gly His Glu Phe Glu Ile
 2360 2365 2370

Glu Gly Glu Gly Glu Gly Arg Pro Tyr Glu Gly Thr Gln Thr Ala
 2375 2380 2385

Lys Leu Lys Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp
 2390 2395 2400

Ile Leu Ser Pro Gln Phe Met Tyr Gly Ser Lys Ala Tyr Val Lys
 2405 2410 2415

His Pro Ala Asp Ile Pro Asp Tyr Leu Lys Leu Ser Phe Pro Glu
 2420 2425 2430

Gly Phe Lys Trp Glu Arg Val Met Asn Phe Glu Asp Gly Gly Val
 2435 2440 2445

Val Thr Val Thr Gln Asp Ser Ser Leu Gln Asp Gly Glu Phe Ile
 2450 2455 2460

Tyr Lys Val Lys Leu Arg Gly Thr Asn Phe Pro Ser Asp Gly Pro
 2465 2470 2475

Val Met Gln Lys Lys Thr Met Gly Trp Glu Ala Ser Ser Glu Arg
 2480 2485 2490

Met Tyr Pro Glu Asp Gly Ala Leu Lys Gly Glu Ile Lys Gln Arg
 2495 2500 2505

048071

Leu Lys Leu Lys Asp Gly Gly His Tyr Asp Ala Glu Val Lys Thr
 2510 2515 2520

Thr Tyr Lys Ala Lys Lys Pro Val Gln Leu Pro Gly Ala Tyr Asn
 2525 2530 2535

Val Asn Ile Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr Thr
 2540 2545 2550

Ile Val Glu Gln Tyr Glu Arg Ala Glu Gly Arg His Ser Thr Gly
 2555 2560 2565

Gly Met Asp Glu Leu Tyr Lys
 2570 2575

<210> 4

<211> 379

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 4

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

Asp Tyr Lys Asp Asp Asp Asp Ala Ser Ile Asp Met Glu Gly Ile Ser
 20 25 30

Ile Tyr Thr Ser Asp Asn Tyr Thr Glu Glu Met Gly Ser Gly Asp Tyr
 35 40 45

Asp Ser Met Lys Glu Pro Cys Phe Arg Glu Glu Asn Ala Asn Phe Asn
 50 55 60

Lys Ile Phe Leu Pro Thr Ile Tyr Ser Ile Ile Phe Leu Thr Gly Ile
 65 70 75 80

Val Gly Asn Gly Leu Val Ile Leu Val Met Gly Tyr Gln Lys Lys Leu
 85 90 95

Arg Ser Met Thr Asp Lys Tyr Arg Leu His Leu Ser Val Ala Asp Leu
 100 105 110

Leu Phe Val Ile Thr Leu Pro Phe Trp Ala Val Asp Ala Val Ala Asn
 115 120 125

048071

Trp Tyr Phe Gly Asn Phe Leu Cys Lys Ala Val His Val Ile Tyr Thr
 130 135 140

Val Asn Leu Tyr Ser Ser Val Leu Ile Leu Ala Phe Ile Ser Leu Asp
 145 150 155 160

Arg Tyr Leu Ala Ile Val His Ala Thr Asn Ser Gln Arg Pro Arg Lys
 165 170 175

Leu Leu Ala Glu Lys Val Val Tyr Val Gly Val Trp Ile Pro Ala Leu
 180 185 190

Leu Leu Thr Ile Pro Asp Phe Ile Phe Ala Asn Val Ser Glu Ala Asp
 195 200 205

Asp Arg Tyr Ile Cys Asp Arg Phe Tyr Pro Asn Asp Leu Trp Val Val
 210 215 220

Val Phe Gln Phe Gln His Ile Met Val Gly Leu Ile Leu Pro Gly Ile
 225 230 235 240

Val Ile Leu Ser Cys Tyr Cys Ile Ile Ile Ser Lys Leu Ser His Ser
 245 250 255

Lys Gly His Gln Lys Arg Lys Ala Leu Lys Thr Thr Val Ile Leu Ile
 260 265 270

Leu Ala Phe Phe Ala Cys Trp Leu Pro Tyr Tyr Ile Gly Ile Ser Ile
 275 280 285

Asp Ser Phe Ile Leu Leu Glu Ile Ile Lys Gln Gly Cys Glu Phe Glu
 290 295 300

Asn Thr Val His Lys Trp Ile Ser Ile Thr Glu Ala Leu Ala Phe Phe
 305 310 315 320

His Cys Cys Leu Asn Pro Ile Leu Tyr Ala Phe Leu Gly Ala Lys Phe
 325 330 335

Lys Thr Ser Ala Gln His Ala Leu Thr Ser Val Ser Arg Gly Ser Ser
 340 345 350

Leu Lys Ile Leu Ser Lys Gly Lys Arg Gly Gly His Ser Ser Val Ser
 355 360 365

Thr Glu Ser Glu Ser Ser Ser Phe His Ser Ser
 370 375

<210> 5
 <211> 27
 <212> белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: синтетическая конструкция

<400> 5
 Gly Arg Thr Pro Pro Ser Leu Gly Pro Gln Asp Glu Ser Cys Thr Thr
 1 5 10 15

Ala Ser Ser Ser Leu Ala Lys Asp Thr Ser Ser
 20 25

<210> 6
 <211> 7
 <212> белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетическая конструкция

<400> 6
 Glu Asn Leu Tyr Phe Gln Leu
 1 5

<210> 7
 <211> 1368
 <212> белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: синтетическая конструкция

<400> 7
 Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val
 1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
 20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
 35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
 50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys

048071

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
 325 330 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
 340 345 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
 355 360 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
 370 375 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
 385 390 395 400

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
 405 410 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
 420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
 435 440 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
 450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
 465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
 485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
 500 505 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
 515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
 530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
 545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
 565 570 575

048071

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
 580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
 595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
 610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
 625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
 645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
 660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
 675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
 690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
 705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
 725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
 740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
 755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
 770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
 785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
 805 810 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
 820 825 830

048071

Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys
 835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
 850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
 865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
 885 890 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
 900 905 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
 915 920 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
 930 935 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
 945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
 965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
 980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
 995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
 1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
 1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
 1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
 1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val

048071

1070						1075					1080			
Arg	Lys	Val	Leu	Ser	Met	Pro	Gln	Val	Asn	Ile	Val	Lys	Lys	Thr
1085						1090					1095			
Glu	Val	Gln	Thr	Gly	Gly	Phe	Ser	Lys	Glu	Ser	Ile	Leu	Pro	Lys
1100						1105					1110			
Arg	Asn	Ser	Asp	Lys	Leu	Ile	Ala	Arg	Lys	Lys	Asp	Trp	Asp	Pro
1115						1120					1125			
Lys	Lys	Tyr	Gly	Gly	Phe	Asp	Ser	Pro	Thr	Val	Ala	Tyr	Ser	Val
1130						1135					1140			
Leu	Val	Val	Ala	Lys	Val	Glu	Lys	Gly	Lys	Ser	Lys	Lys	Leu	Lys
1145						1150					1155			
Ser	Val	Lys	Glu	Leu	Leu	Gly	Ile	Thr	Ile	Met	Glu	Arg	Ser	Ser
1160						1165					1170			
Phe	Glu	Lys	Asn	Pro	Ile	Asp	Phe	Leu	Glu	Ala	Lys	Gly	Tyr	Lys
1175						1180					1185			
Glu	Val	Lys	Lys	Asp	Leu	Ile	Ile	Lys	Leu	Pro	Lys	Tyr	Ser	Leu
1190						1195					1200			
Phe	Glu	Leu	Glu	Asn	Gly	Arg	Lys	Arg	Met	Leu	Ala	Ser	Ala	Gly
1205						1210					1215			
Glu	Leu	Gln	Lys	Gly	Asn	Glu	Leu	Ala	Leu	Pro	Ser	Lys	Tyr	Val
1220						1225					1230			
Asn	Phe	Leu	Tyr	Leu	Ala	Ser	His	Tyr	Glu	Lys	Leu	Lys	Gly	Ser
1235						1240					1245			
Pro	Glu	Asp	Asn	Glu	Gln	Lys	Gln	Leu	Phe	Val	Glu	Gln	His	Lys
1250						1255					1260			
His	Tyr	Leu	Asp	Glu	Ile	Ile	Glu	Gln	Ile	Ser	Glu	Phe	Ser	Lys
1265						1270					1275			
Arg	Val	Ile	Leu	Ala	Asp	Ala	Asn	Leu	Asp	Lys	Val	Leu	Ser	Ala
1280						1285					1290			
Tyr	Asn	Lys	His	Arg	Asp	Lys	Pro	Ile	Arg	Glu	Gln	Ala	Glu	Asn
1295						1300					1305			

048071

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala
 1310 1315 1320

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser
 1325 1330 1335

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr
 1340 1345 1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
 1355 1360 1365

<210> 8

<211> 522

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 8

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
 1 5 10 15

Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe
 20 25 30

Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp
 35 40 45

Met Leu Gly Ser Pro Lys Lys Lys Arg Lys Val Gly Ser Gln Tyr Leu
 50 55 60

Pro Asp Thr Asp Asp Arg His Arg Ile Glu Glu Lys Arg Lys Arg Thr
 65 70 75 80

Tyr Glu Thr Phe Lys Ser Ile Met Lys Lys Ser Pro Phe Ser Gly Pro
 85 90 95

Thr Asp Pro Arg Pro Pro Pro Arg Arg Ile Ala Val Pro Ser Arg Ser
 100 105 110

Ser Ala Ser Val Pro Lys Pro Ala Pro Gln Pro Tyr Pro Phe Thr Ser
 115 120 125

Ser Leu Ser Thr Ile Asn Tyr Asp Glu Phe Pro Thr Met Val Phe Pro
 130 135 140

048071

Ser Gly Gln Ile Ser Gln Ala Ser Ala Leu Ala Pro Ala Pro Pro Gln
145 150 155 160

Val Leu Pro Gln Ala Pro Ala Pro Ala Pro Ala Pro Ala Met Val Ser
165 170 175

Ala Leu Ala Gln Ala Pro Ala Pro Val Pro Val Leu Ala Pro Gly Pro
180 185 190

Pro Gln Ala Val Ala Pro Pro Ala Pro Lys Pro Thr Gln Ala Gly Glu
195 200 205

Gly Thr Leu Ser Glu Ala Leu Leu Gln Leu Gln Phe Asp Asp Glu Asp
210 215 220

Leu Gly Ala Leu Leu Gly Asn Ser Thr Asp Pro Ala Val Phe Thr Asp
225 230 235 240

Leu Ala Ser Val Asp Asn Ser Glu Phe Gln Gln Leu Leu Asn Gln Gly
245 250 255

Ile Pro Val Ala Pro His Thr Thr Glu Pro Met Leu Met Glu Tyr Pro
260 265 270

Glu Ala Ile Thr Arg Leu Val Thr Gly Ala Gln Arg Pro Pro Asp Pro
275 280 285

Ala Pro Ala Pro Leu Gly Ala Pro Gly Leu Pro Asn Gly Leu Leu Ser
290 295 300

Gly Asp Glu Asp Phe Ser Ser Ile Ala Asp Met Asp Phe Ser Ala Leu
305 310 315 320

Leu Ser Gln Ile Ser Ser Gly Ser Gly Ser Gly Ser Arg Asp Ser Arg
325 330 335

Glu Gly Met Phe Leu Pro Lys Pro Glu Ala Gly Ser Ala Ile Ser Asp
340 345 350

Val Phe Glu Gly Arg Glu Val Cys Gln Pro Lys Arg Ile Arg Pro Phe
355 360 365

His Pro Pro Gly Ser Pro Trp Ala Asn Arg Pro Leu Pro Ala Ser Leu
370 375 380

Ala Pro Thr Pro Thr Gly Pro Val His Glu Pro Val Gly Ser Leu Thr
385 390 395 400

048071

Pro Ala Pro Val Pro Gln Pro Leu Asp Pro Ala Pro Ala Val Thr Pro
405 410 415

Glu Ala Ser His Leu Leu Glu Asp Pro Asp Glu Glu Thr Ser Gln Ala
420 425 430

Val Lys Ala Leu Arg Glu Met Ala Asp Thr Val Ile Pro Gln Lys Glu
435 440 445

Glu Ala Ala Ile Cys Gly Gln Met Asp Leu Ser His Pro Pro Pro Arg
450 455 460

Gly His Leu Asp Glu Leu Thr Thr Thr Leu Glu Ser Met Thr Glu Asp
465 470 475 480

Leu Asn Leu Asp Ser Pro Leu Thr Pro Glu Leu Asn Glu Ile Leu Asp
485 490 495

Thr Phe Leu Asn Asp Glu Cys Leu Leu His Ala Met His Ile Ser Thr
500 505 510

Gly Leu Ser Ile Phe Asp Thr Ser Leu Phe
515 520

<210> 9

<211> 7

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 9

Pro Lys Lys Lys Arg Lys Val

1 5

<210> 10

<211> 8

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 10

Tyr Pro Tyr Asp Val Pro Asp Tyr

1 5

<210> 11

048071

<211> 10

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 11

Gly Ile Gly Ser Gly Ser Asn Gly Ser Ser
1 5 10

<210> 12

<211> 2951

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 12

Met Asn Leu Gln Pro Ile Phe Trp Ile Gly Leu Ile Ser Ser Val Cys
1 5 10 15

Cys Val Phe Ala Gln Thr Asp Glu Asn Arg Cys Leu Lys Ala Asn Ala
20 25 30

Lys Ser Cys Gly Glu Cys Ile Gln Ala Gly Pro Asn Cys Gly Trp Cys
35 40 45

Thr Asn Ser Thr Phe Leu Gln Glu Gly Met Pro Thr Ser Ala Arg Cys
50 55 60

Asp Asp Leu Glu Ala Leu Lys Lys Lys Gly Cys Pro Pro Asp Asp Ile
65 70 75 80

Glu Asn Pro Arg Gly Ser Lys Asp Ile Lys Lys Asn Lys Asn Val Thr
85 90 95

Asn Arg Ser Lys Gly Thr Ala Glu Lys Leu Lys Pro Glu Asp Ile Thr
100 105 110

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

Gln Thr Phe Thr Leu Lys Phe Lys Arg Ala Glu Asp Tyr Pro Ile Asp
130 135 140

Leu Tyr Tyr Leu Met Asp Leu Ser Tyr Ser Met Lys Asp Asp Leu Glu
145 150 155 160

048071

Asn Val Lys Ser Leu Gly Thr Asp Leu Met Asn Glu Met Arg Arg Ile
 165 170 175

Thr Ser Asp Phe Arg Ile Gly Phe Gly Ser Phe Val Glu Lys Thr Val
 180 185 190

Met Pro Tyr Ile Ser Thr Thr Pro Ala Lys Leu Arg Asn Pro Cys Thr
 195 200 205

Ser Glu Gln Asn Cys Thr Ser Pro Phe Ser Tyr Lys Asn Val Leu Ser
 210 215 220

Leu Thr Asn Lys Gly Glu Val Phe Asn Glu Leu Val Gly Lys Gln Arg
 225 230 235 240

Ile Ser Gly Asn Leu Asp Ser Pro Glu Gly Gly Phe Asp Ala Ile Met
 245 250 255

Gln Val Ala Val Cys Gly Ser Leu Ile Gly Trp Arg Asn Val Thr Arg
 260 265 270

Leu Leu Val Phe Ser Thr Asp Ala Gly Phe His Phe Ala Gly Asp Gly
 275 280 285

Lys Leu Gly Gly Ile Val Leu Pro Asn Asp Gly Gln Cys His Leu Glu
 290 295 300

Asn Asn Met Tyr Thr Met Ser His Tyr Tyr Asp Tyr Pro Ser Ile Ala
 305 310 315 320

His Leu Val Gln Lys Leu Ser Glu Asn Asn Ile Gln Thr Ile Phe Ala
 325 330 335

Val Thr Glu Glu Phe Gln Pro Val Tyr Lys Glu Leu Lys Asn Leu Ile
 340 345 350

Pro Lys Ser Ala Val Gly Thr Leu Ser Ala Asn Ser Ser Asn Val Ile
 355 360 365

Gln Leu Ile Ile Asp Ala Tyr Asn Ser Leu Ser Ser Glu Val Ile Leu
 370 375 380

Glu Asn Gly Lys Leu Ser Glu Gly Val Thr Ile Ser Tyr Lys Ser Tyr
 385 390 395 400

Cys Lys Asn Gly Val Asn Gly Thr Gly Glu Asn Gly Arg Lys Cys Ser

048071

405 410 415
 Asn Ile Ser Ile Gly Asp Glu Val Gln Phe Glu Ile Ser Ile Thr Ser
 420 425 430
 Asn Lys Cys Pro Lys Lys Asp Ser Asp Ser Phe Lys Ile Arg Pro Leu
 435 440 445
 Gly Phe Thr Glu Glu Val Glu Val Ile Leu Gln Tyr Ile Cys Val Cys
 450 455 460
 Glu Cys Gln Ser Glu Gly Ile Pro Glu Ser Pro Lys Cys His Glu Gly
 465 470 475 480
 Asn Gly Thr Phe Glu Cys Gly Ala Cys Arg Cys Asn Glu Gly Arg Val
 485 490 495
 Gly Arg His Cys Glu Cys Ser Thr Asp Glu Val Asn Ser Glu Asp Met
 500 505 510
 Asp Ala Tyr Cys Arg Lys Glu Asn Ser Ser Glu Ile Cys Ser Asn Asn
 515 520 525
 Gly Glu Cys Val Cys Gly Gln Cys Val Cys Arg Lys Arg Asp Asn Thr
 530 535 540
 Asn Glu Ile Tyr Ser Gly Lys Phe Cys Glu Cys Asp Asn Phe Asn Cys
 545 550 555 560
 Asp Arg Ser Asn Gly Leu Ile Cys Gly Gly Asn Gly Val Cys Lys Cys
 565 570 575
 Arg Val Cys Glu Cys Asn Pro Asn Tyr Thr Gly Ser Ala Cys Asp Cys
 580 585 590
 Ser Leu Asp Thr Ser Thr Cys Glu Ala Ser Asn Gly Gln Ile Cys Asn
 595 600 605
 Gly Arg Gly Ile Cys Glu Cys Gly Val Cys Lys Cys Thr Asp Pro Lys
 610 615 620
 Phe Gln Gly Gln Thr Cys Glu Met Cys Gln Thr Cys Leu Gly Val Cys
 625 630 635 640
 Ala Glu His Lys Glu Cys Val Gln Cys Arg Ala Phe Asn Lys Gly Glu
 645 650 655

048071

Lys Lys Asp Thr Cys Thr Gln Glu Cys Ser Tyr Phe Asn Ile Thr Lys
660 665 670

Val Glu Ser Arg Asp Lys Leu Pro Gln Pro Val Gln Pro Asp Pro Val
675 680 685

Ser His Cys Lys Glu Lys Asp Val Asp Asp Cys Trp Phe Tyr Phe Thr
690 695 700

Tyr Ser Val Asn Gly Asn Asn Glu Val Met Val His Val Val Glu Asn
705 710 715 720

Pro Glu Cys Pro Thr Gly Pro Asp Ile Ile Pro Ile Val Ala Gly Val
725 730 735

Val Ala Gly Ile Val Leu Ile Gly Leu Ala Leu Leu Leu Ile Trp Lys
740 745 750

Leu Leu Met Ile Ile His Asp Arg Arg Glu Phe Ala Lys Phe Glu Lys
755 760 765

Glu Lys Met Asn Ala Lys Trp Asp Thr Gly Glu Asn Pro Ile Tyr Lys
770 775 780

Ser Ala Val Thr Thr Val Val Asn Pro Lys Tyr Glu Gly Lys Glu Asn
785 790 795 800

Leu Tyr Phe Gln Leu Ala Ser Met Asp Lys Lys Tyr Ser Ile Gly Leu
805 810 815

Ala Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr
820 825 830

Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His
835 840 845

Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu
850 855 860

Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr
865 870 875 880

Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu
885 890 895

Met Ala Lys Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe
900 905 910

048071

Leu Val Glu Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn
 915 920 925

Ile Val Asp Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His
 930 935 940

Leu Arg Lys Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu
 945 950 955 960

Ile Tyr Leu Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu
 965 970 975

Ile Glu Gly Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe
 980 985 990

Ile Gln Leu Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile
 995 1000 1005

Asn Ala Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu
 1010 1015 1020

Ser Lys Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly
 1025 1030 1035

Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu
 1040 1045 1050

Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp
 1055 1060 1065

Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp
 1070 1075 1080

Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu
 1085 1090 1095

Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu
 1100 1105 1110

Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met
 1115 1120 1125

Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
 1130 1135 1140

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe
 1145 1150 1155

048071

Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly
 1160 1165 1170

Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu
 1175 1180 1185

Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu
 1190 1195 1200

Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro
 1205 1210 1215

His Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln
 1220 1225 1230

Glu Asp Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu
 1235 1240 1245

Lys Ile Leu Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala
 1250 1255 1260

Arg Gly Asn Ser Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu
 1265 1270 1275

Thr Ile Thr Pro Trp Asn Phe Glu Glu Val Val Asp Lys Gly Ala
 1280 1285 1290

Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn Phe Asp Lys Asn
 1295 1300 1305

Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu Leu Tyr Glu
 1310 1315 1320

Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr Val Thr
 1325 1330 1335

Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys Lys
 1340 1345 1350

Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val
 1355 1360 1365

Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
 1370 1375 1380

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu

048071

1385						1390						1395			
Gly	Thr	Tyr	His	Asp	Leu	Leu	Lys	Ile	Ile	Lys	Asp	Lys	Asp	Phe	
1400						1405					1410				
Leu	Asp	Asn	Glu	Glu	Asn	Glu	Asp	Ile	Leu	Glu	Asp	Ile	Val	Leu	
1415						1420					1425				
Thr	Leu	Thr	Leu	Phe	Glu	Asp	Arg	Glu	Met	Ile	Glu	Glu	Arg	Leu	
1430						1435					1440				
Lys	Thr	Tyr	Ala	His	Leu	Phe	Asp	Asp	Lys	Val	Met	Lys	Gln	Leu	
1445						1450					1455				
Lys	Arg	Arg	Arg	Tyr	Thr	Gly	Trp	Gly	Arg	Leu	Ser	Arg	Lys	Leu	
1460						1465					1470				
Ile	Asn	Gly	Ile	Arg	Asp	Lys	Gln	Ser	Gly	Lys	Thr	Ile	Leu	Asp	
1475						1480					1485				
Phe	Leu	Lys	Ser	Asp	Gly	Phe	Ala	Asn	Arg	Asn	Phe	Met	Gln	Leu	
1490						1495					1500				
Ile	His	Asp	Asp	Ser	Leu	Thr	Phe	Lys	Glu	Asp	Ile	Gln	Lys	Ala	
1505						1510					1515				
Gln	Val	Ser	Gly	Gln	Gly	Asp	Ser	Leu	His	Glu	His	Ile	Ala	Asn	
1520						1525					1530				
Leu	Ala	Gly	Ser	Pro	Ala	Ile	Lys	Lys	Gly	Ile	Leu	Gln	Thr	Val	
1535						1540					1545				
Lys	Val	Val	Asp	Glu	Leu	Val	Lys	Val	Met	Gly	Arg	His	Lys	Pro	
1550						1555					1560				
Glu	Asn	Ile	Val	Ile	Glu	Met	Ala	Arg	Glu	Asn	Gln	Thr	Thr	Gln	
1565						1570					1575				
Lys	Gly	Gln	Lys	Asn	Ser	Arg	Glu	Arg	Met	Lys	Arg	Ile	Glu	Glu	
1580						1585					1590				
Gly	Ile	Lys	Glu	Leu	Gly	Ser	Gln	Ile	Leu	Lys	Glu	His	Pro	Val	
1595						1600					1605				
Glu	Asn	Thr	Gln	Leu	Gln	Asn	Glu	Lys	Leu	Tyr	Leu	Tyr	Tyr	Leu	
1610						1615					1620				

048071

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn
 1625 1630 1635

Arg Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe
 1640 1645 1650

Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp
 1655 1660 1665

Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val
 1670 1675 1680

Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu
 1685 1690 1695

Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly
 1700 1705 1710

Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu
 1715 1720 1725

Val Glu Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp
 1730 1735 1740

Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg
 1745 1750 1755

Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe
 1760 1765 1770

Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr
 1775 1780 1785

His His Ala His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala
 1790 1795 1800

Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly
 1805 1810 1815

Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu
 1820 1825 1830

Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn
 1835 1840 1845

Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu
 1850 1855 1860

048071

Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu
 1865 1870 1875

Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val
 1880 1885 1890

Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln
 1895 1900 1905

Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser
 1910 1915 1920

Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr
 1925 1930 1935

Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val
 1940 1945 1950

Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys
 1955 1960 1965

Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys
 1970 1975 1980

Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys
 1985 1990 1995

Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu
 2000 2005 2010

Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln
 2015 2020 2025

Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu
 2030 2035 2040

Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp
 2045 2050 2055

Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His Tyr Leu
 2060 2065 2070

Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile
 2075 2080 2085

Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys
 2090 2095 2100

048071

His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His
 2105 2110 2115

 Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr
 2120 2125 2130

 Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr Lys Glu
 2135 2140 2145

 Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr Gly Leu Tyr
 2150 2155 2160

 Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp Ala Tyr Pro
 2165 2170 2175

 Tyr Asp Val Pro Asp Tyr Ala Ser Leu Gly Ser Gly Asp Ala Leu
 2180 2185 2190

 Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp
 2195 2200 2205

 Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp
 2210 2215 2220

 Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp
 2225 2230 2235

 Met Leu Gly Ser Pro Lys Lys Lys Arg Lys Val Gly Ser Gln Tyr
 2240 2245 2250

 Leu Pro Asp Thr Asp Asp Arg His Arg Ile Glu Glu Lys Arg Lys
 2255 2260 2265

 Arg Thr Tyr Glu Thr Phe Lys Ser Ile Met Lys Lys Ser Pro Phe
 2270 2275 2280

 Ser Gly Pro Thr Asp Pro Arg Pro Pro Pro Arg Arg Ile Ala Val
 2285 2290 2295

 Pro Ser Arg Ser Ser Ala Ser Val Pro Lys Pro Ala Pro Gln Pro
 2300 2305 2310

 Tyr Pro Phe Thr Ser Ser Leu Ser Thr Ile Asn Tyr Asp Glu Phe
 2315 2320 2325

 Pro Thr Met Val Phe Pro Ser Gly Gln Ile Ser Gln Ala Ser Ala

048071

2330						2335					2340			
Leu	Ala	Pro	Ala	Pro	Pro	Gln	Val	Leu	Pro	Gln	Ala	Pro	Ala	Pro
2345						2350					2355			
Ala	Pro	Ala	Pro	Ala	Met	Val	Ser	Ala	Leu	Ala	Gln	Ala	Pro	Ala
2360						2365					2370			
Pro	Val	Pro	Val	Leu	Ala	Pro	Gly	Pro	Pro	Gln	Ala	Val	Ala	Pro
2375						2380					2385			
Pro	Ala	Pro	Lys	Pro	Thr	Gln	Ala	Gly	Glu	Gly	Thr	Leu	Ser	Glu
2390						2395					2400			
Ala	Leu	Leu	Gln	Leu	Gln	Phe	Asp	Asp	Glu	Asp	Leu	Gly	Ala	Leu
2405						2410					2415			
Leu	Gly	Asn	Ser	Thr	Asp	Pro	Ala	Val	Phe	Thr	Asp	Leu	Ala	Ser
2420						2425					2430			
Val	Asp	Asn	Ser	Glu	Phe	Gln	Gln	Leu	Leu	Asn	Gln	Gly	Ile	Pro
2435						2440					2445			
Val	Ala	Pro	His	Thr	Thr	Glu	Pro	Met	Leu	Met	Glu	Tyr	Pro	Glu
2450						2455					2460			
Ala	Ile	Thr	Arg	Leu	Val	Thr	Gly	Ala	Gln	Arg	Pro	Pro	Asp	Pro
2465						2470					2475			
Ala	Pro	Ala	Pro	Leu	Gly	Ala	Pro	Gly	Leu	Pro	Asn	Gly	Leu	Leu
2480						2485					2490			
Ser	Gly	Asp	Glu	Asp	Phe	Ser	Ser	Ile	Ala	Asp	Met	Asp	Phe	Ser
2495						2500					2505			
Ala	Leu	Leu	Ser	Gln	Ile	Ser	Ser	Gly	Ser	Gly	Ser	Gly	Ser	Arg
2510						2515					2520			
Asp	Ser	Arg	Glu	Gly	Met	Phe	Leu	Pro	Lys	Pro	Glu	Ala	Gly	Ser
2525						2530					2535			
Ala	Ile	Ser	Asp	Val	Phe	Glu	Gly	Arg	Glu	Val	Cys	Gln	Pro	Lys
2540						2545					2550			
Arg	Ile	Arg	Pro	Phe	His	Pro	Pro	Gly	Ser	Pro	Trp	Ala	Asn	Arg
2555						2560					2565			

048071

Pro Leu Pro Ala Ser Leu Ala Pro Thr Pro Thr Gly Pro Val His
 2570 2575 2580

Glu Pro Val Gly Ser Leu Thr Pro Ala Pro Val Pro Gln Pro Leu
 2585 2590 2595

Asp Pro Ala Pro Ala Val Thr Pro Glu Ala Ser His Leu Leu Glu
 2600 2605 2610

Asp Pro Asp Glu Glu Thr Ser Gln Ala Val Lys Ala Leu Arg Glu
 2615 2620 2625

Met Ala Asp Thr Val Ile Pro Gln Lys Glu Glu Ala Ala Ile Cys
 2630 2635 2640

Gly Gln Met Asp Leu Ser His Pro Pro Pro Arg Gly His Leu Asp
 2645 2650 2655

Glu Leu Thr Thr Thr Leu Glu Ser Met Thr Glu Asp Leu Asn Leu
 2660 2665 2670

Asp Ser Pro Leu Thr Pro Glu Leu Asn Glu Ile Leu Asp Thr Phe
 2675 2680 2685

Leu Asn Asp Glu Cys Leu Leu His Ala Met His Ile Ser Thr Gly
 2690 2695 2700

Leu Ser Ile Phe Asp Thr Ser Leu Phe Leu Met Val Ser Lys Gly
 2705 2710 2715

Glu Glu Asp Asn Met Ala Ile Ile Lys Glu Phe Met Arg Phe Lys
 2720 2725 2730

Val His Met Glu Gly Ser Val Asn Gly His Glu Phe Glu Ile Glu
 2735 2740 2745

Gly Glu Gly Glu Gly Arg Pro Tyr Glu Gly Thr Gln Thr Ala Lys
 2750 2755 2760

Leu Lys Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp Ile
 2765 2770 2775

Leu Ser Pro Gln Phe Met Tyr Gly Ser Lys Ala Tyr Val Lys His
 2780 2785 2790

Pro Ala Asp Ile Pro Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly
 2795 2800 2805

048071

Phe Lys Trp Glu Arg Val Met Asn Phe Glu Asp Gly Gly Val Val
 2810 2815 2820

Thr Val Thr Gln Asp Ser Ser Leu Gln Asp Gly Glu Phe Ile Tyr
 2825 2830 2835

Lys Val Lys Leu Arg Gly Thr Asn Phe Pro Ser Asp Gly Pro Val
 2840 2845 2850

Met Gln Lys Lys Thr Met Gly Trp Glu Ala Ser Ser Glu Arg Met
 2855 2860 2865

Tyr Pro Glu Asp Gly Ala Leu Lys Gly Glu Ile Lys Gln Arg Leu
 2870 2875 2880

Lys Leu Lys Asp Gly Gly His Tyr Asp Ala Glu Val Lys Thr Thr
 2885 2890 2895

Tyr Lys Ala Lys Lys Pro Val Gln Leu Pro Gly Ala Tyr Asn Val
 2900 2905 2910

Asn Ile Lys Leu Asp Ile Thr Ser His Asn Glu Asp Tyr Thr Ile
 2915 2920 2925

Val Glu Gln Tyr Glu Arg Ala Glu Gly Arg His Ser Thr Gly Gly
 2930 2935 2940

Met Asp Glu Leu Tyr Lys Leu Glu
 2945 2950

<210> 13

<211> 798

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 13

Met Asn Leu Gln Pro Ile Phe Trp Ile Gly Leu Ile Ser Ser Val Cys
 1 5 10 15

Cys Val Phe Ala Gln Thr Asp Glu Asn Arg Cys Leu Lys Ala Asn Ala
 20 25 30

Lys Ser Cys Gly Glu Cys Ile Gln Ala Gly Pro Asn Cys Gly Trp Cys
 35 40 45

048071

Thr Asn Ser Thr Phe Leu Gln Glu Gly Met Pro Thr Ser Ala Arg Cys
50 55 60

Asp Asp Leu Glu Ala Leu Lys Lys Lys Gly Cys Pro Pro Asp Asp Ile
65 70 75 80

Glu Asn Pro Arg Gly Ser Lys Asp Ile Lys Lys Asn Lys Asn Val Thr
85 90 95

Asn Arg Ser Lys Gly Thr Ala Glu Lys Leu Lys Pro Glu Asp Ile Thr
100 105 110

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

Gln Thr Phe Thr Leu Lys Phe Lys Arg Ala Glu Asp Tyr Pro Ile Asp
130 135 140

Leu Tyr Tyr Leu Met Asp Leu Ser Tyr Ser Met Lys Asp Asp Leu Glu
145 150 155 160

Asn Val Lys Ser Leu Gly Thr Asp Leu Met Asn Glu Met Arg Arg Ile
165 170 175

Thr Ser Asp Phe Arg Ile Gly Phe Gly Ser Phe Val Glu Lys Thr Val
180 185 190

Met Pro Tyr Ile Ser Thr Thr Pro Ala Lys Leu Arg Asn Pro Cys Thr
195 200 205

Ser Glu Gln Asn Cys Thr Ser Pro Phe Ser Tyr Lys Asn Val Leu Ser
210 215 220

Leu Thr Asn Lys Gly Glu Val Phe Asn Glu Leu Val Gly Lys Gln Arg
225 230 235 240

Ile Ser Gly Asn Leu Asp Ser Pro Glu Gly Gly Phe Asp Ala Ile Met
245 250 255

Gln Val Ala Val Cys Gly Ser Leu Ile Gly Trp Arg Asn Val Thr Arg
260 265 270

Leu Leu Val Phe Ser Thr Asp Ala Gly Phe His Phe Ala Gly Asp Gly
275 280 285

Lys Leu Gly Gly Ile Val Leu Pro Asn Asp Gly Gln Cys His Leu Glu
290 295 300

048071

Asn Asn Met Tyr Thr Met Ser His Tyr Tyr Asp Tyr Pro Ser Ile Ala
 305 310 315 320

His Leu Val Gln Lys Leu Ser Glu Asn Asn Ile Gln Thr Ile Phe Ala
 325 330 335

Val Thr Glu Glu Phe Gln Pro Val Tyr Lys Glu Leu Lys Asn Leu Ile
 340 345 350

Pro Lys Ser Ala Val Gly Thr Leu Ser Ala Asn Ser Ser Asn Val Ile
 355 360 365

Gln Leu Ile Ile Asp Ala Tyr Asn Ser Leu Ser Ser Glu Val Ile Leu
 370 375 380

Glu Asn Gly Lys Leu Ser Glu Gly Val Thr Ile Ser Tyr Lys Ser Tyr
 385 390 395 400

Cys Lys Asn Gly Val Asn Gly Thr Gly Glu Asn Gly Arg Lys Cys Ser
 405 410 415

Asn Ile Ser Ile Gly Asp Glu Val Gln Phe Glu Ile Ser Ile Thr Ser
 420 425 430

Asn Lys Cys Pro Lys Lys Asp Ser Asp Ser Phe Lys Ile Arg Pro Leu
 435 440 445

Gly Phe Thr Glu Glu Val Glu Val Ile Leu Gln Tyr Ile Cys Val Cys
 450 455 460

Glu Cys Gln Ser Glu Gly Ile Pro Glu Ser Pro Lys Cys His Glu Gly
 465 470 475 480

Asn Gly Thr Phe Glu Cys Gly Ala Cys Arg Cys Asn Glu Gly Arg Val
 485 490 495

Gly Arg His Cys Glu Cys Ser Thr Asp Glu Val Asn Ser Glu Asp Met
 500 505 510

Asp Ala Tyr Cys Arg Lys Glu Asn Ser Ser Glu Ile Cys Ser Asn Asn
 515 520 525

Gly Glu Cys Val Cys Gly Gln Cys Val Cys Arg Lys Arg Asp Asn Thr
 530 535 540

Asn Glu Ile Tyr Ser Gly Lys Phe Cys Glu Cys Asp Asn Phe Asn Cys

<210> 14

<211> 7

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 14

Glu Asn Leu Tyr Phe Gln Leu

1 5

<210> 15

<211> 1368

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 15

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile

35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu

50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys

65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser

85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys

100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr

115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp

130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His

048071

145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn
210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
290 295 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305 310 315 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
325 330 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
340 345 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
355 360 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
370 375 380

Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385 390 395 400

048071

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
435 440 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480

Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510

Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655

048071

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu
705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly
725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly
740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln
755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
770 775 780

Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg
820 825 830

Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys
835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys
885 890 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
900 905 910

048071

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
 915 920 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
 930 935 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser
 945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
 965 970 975

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
 980 985 990

Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
 995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
 1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
 1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
 1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
 1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
 1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
 1085 1090 1095

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys
 1100 1105 1110

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro
 1115 1120 1125

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val
 1130 1135 1140

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys

048071

1145						1150						1155			
Ser	Val	Lys	Glu	Leu	Leu	Gly	Ile	Thr	Ile	Met	Glu	Arg	Ser	Ser	
1160						1165					1170				
Phe	Glu	Lys	Asn	Pro	Ile	Asp	Phe	Leu	Glu	Ala	Lys	Gly	Tyr	Lys	
1175						1180					1185				
Glu	Val	Lys	Lys	Asp	Leu	Ile	Ile	Lys	Leu	Pro	Lys	Tyr	Ser	Leu	
1190						1195					1200				
Phe	Glu	Leu	Glu	Asn	Gly	Arg	Lys	Arg	Met	Leu	Ala	Ser	Ala	Gly	
1205						1210					1215				
Glu	Leu	Gln	Lys	Gly	Asn	Glu	Leu	Ala	Leu	Pro	Ser	Lys	Tyr	Val	
1220						1225					1230				
Asn	Phe	Leu	Tyr	Leu	Ala	Ser	His	Tyr	Glu	Lys	Leu	Lys	Gly	Ser	
1235						1240					1245				
Pro	Glu	Asp	Asn	Glu	Gln	Lys	Gln	Leu	Phe	Val	Glu	Gln	His	Lys	
1250						1255					1260				
His	Tyr	Leu	Asp	Glu	Ile	Ile	Glu	Gln	Ile	Ser	Glu	Phe	Ser	Lys	
1265						1270					1275				
Arg	Val	Ile	Leu	Ala	Asp	Ala	Asn	Leu	Asp	Lys	Val	Leu	Ser	Ala	
1280						1285					1290				
Tyr	Asn	Lys	His	Arg	Asp	Lys	Pro	Ile	Arg	Glu	Gln	Ala	Glu	Asn	
1295						1300					1305				
Ile	Ile	His	Leu	Phe	Thr	Leu	Thr	Asn	Leu	Gly	Ala	Pro	Ala	Ala	
1310						1315					1320				
Phe	Lys	Tyr	Phe	Asp	Thr	Thr	Ile	Asp	Arg	Lys	Arg	Tyr	Thr	Ser	
1325						1330					1335				
Thr	Lys	Glu	Val	Leu	Asp	Ala	Thr	Leu	Ile	His	Gln	Ser	Ile	Thr	
1340						1345					1350				
Gly	Leu	Tyr	Glu	Thr	Arg	Ile	Asp	Leu	Ser	Gln	Leu	Gly	Gly	Asp	
1355						1360					1365				

<210> 16

<211> 522

<212> белок

<213> Искусственная последовательность

048071

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 16

Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu
 1 5 10 15

Asp Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe
 20 25 30

Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp
 35 40 45

Met Leu Gly Ser Pro Lys Lys Lys Arg Lys Val Gly Ser Gln Tyr Leu
 50 55 60

Pro Asp Thr Asp Asp Arg His Arg Ile Glu Glu Lys Arg Lys Arg Thr
 65 70 75 80

Tyr Glu Thr Phe Lys Ser Ile Met Lys Lys Ser Pro Phe Ser Gly Pro
 85 90 95

Thr Asp Pro Arg Pro Pro Pro Arg Arg Ile Ala Val Pro Ser Arg Ser
 100 105 110

Ser Ala Ser Val Pro Lys Pro Ala Pro Gln Pro Tyr Pro Phe Thr Ser
 115 120 125

Ser Leu Ser Thr Ile Asn Tyr Asp Glu Phe Pro Thr Met Val Phe Pro
 130 135 140

Ser Gly Gln Ile Ser Gln Ala Ser Ala Leu Ala Pro Ala Pro Pro Gln
 145 150 155 160

Val Leu Pro Gln Ala Pro Ala Pro Ala Pro Ala Pro Ala Met Val Ser
 165 170 175

Ala Leu Ala Gln Ala Pro Ala Pro Val Pro Val Leu Ala Pro Gly Pro
 180 185 190

Pro Gln Ala Val Ala Pro Pro Ala Pro Lys Pro Thr Gln Ala Gly Glu
 195 200 205

Gly Thr Leu Ser Glu Ala Leu Leu Gln Leu Gln Phe Asp Asp Glu Asp
 210 215 220

048071

Leu Gly Ala Leu Leu Gly Asn Ser Thr Asp Pro Ala Val Phe Thr Asp
 225 230 235 240

Leu Ala Ser Val Asp Asn Ser Glu Phe Gln Gln Leu Leu Asn Gln Gly
 245 250 255

Ile Pro Val Ala Pro His Thr Thr Glu Pro Met Leu Met Glu Tyr Pro
 260 265 270

Glu Ala Ile Thr Arg Leu Val Thr Gly Ala Gln Arg Pro Pro Asp Pro
 275 280 285

Ala Pro Ala Pro Leu Gly Ala Pro Gly Leu Pro Asn Gly Leu Leu Ser
 290 295 300

Gly Asp Glu Asp Phe Ser Ser Ile Ala Asp Met Asp Phe Ser Ala Leu
 305 310 315 320

Leu Ser Gln Ile Ser Ser Gly Ser Gly Ser Gly Ser Arg Asp Ser Arg
 325 330 335

Glu Gly Met Phe Leu Pro Lys Pro Glu Ala Gly Ser Ala Ile Ser Asp
 340 345 350

Val Phe Glu Gly Arg Glu Val Cys Gln Pro Lys Arg Ile Arg Pro Phe
 355 360 365

His Pro Pro Gly Ser Pro Trp Ala Asn Arg Pro Leu Pro Ala Ser Leu
 370 375 380

Ala Pro Thr Pro Thr Gly Pro Val His Glu Pro Val Gly Ser Leu Thr
 385 390 395 400

Pro Ala Pro Val Pro Gln Pro Leu Asp Pro Ala Pro Ala Val Thr Pro
 405 410 415

Glu Ala Ser His Leu Leu Glu Asp Pro Asp Glu Glu Thr Ser Gln Ala
 420 425 430

Val Lys Ala Leu Arg Glu Met Ala Asp Thr Val Ile Pro Gln Lys Glu
 435 440 445

Glu Ala Ala Ile Cys Gly Gln Met Asp Leu Ser His Pro Pro Pro Arg
 450 455 460

Gly His Leu Asp Glu Leu Thr Thr Thr Leu Glu Ser Met Thr Glu Asp
 465 470 475 480

048071

Leu Asn Leu Asp Ser Pro Leu Thr Pro Glu Leu Asn Glu Ile Leu Asp
485 490 495

Thr Phe Leu Asn Asp Glu Cys Leu Leu His Ala Met His Ile Ser Thr
500 505 510

Gly Leu Ser Ile Phe Asp Thr Ser Leu Phe
515 520

<210> 17

<211> 7

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 17

Pro Lys Lys Lys Arg Lys Val
1 5

<210> 18

<211> 238

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 18

Met Val Ser Lys Gly Glu Glu Asp Asn Met Ala Ile Ile Lys Glu Phe
1 5 10 15

Met Arg Phe Lys Val His Met Glu Gly Ser Val Asn Gly His Glu Phe
20 25 30

Glu Ile Glu Gly Glu Gly Glu Gly Arg Pro Tyr Glu Gly Thr Gln Thr
35 40 45

Ala Lys Leu Lys Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp
50 55 60

Ile Leu Ser Pro Gln Phe Met Tyr Gly Ser Lys Ala Tyr Val Lys His
65 70 75 80

Pro Ala Asp Ile Pro Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Phe
85 90 95

048071

Lys Trp Glu Arg Val Met Asn Phe Glu Asp Gly Gly Val Val Thr Val
 100 105 110

Thr Gln Asp Ser Ser Leu Gln Asp Gly Glu Phe Ile Tyr Lys Val Lys
 115 120 125

Leu Arg Gly Thr Asn Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys
 130 135 140

Thr Met Gly Trp Glu Ala Ser Ser Glu Arg Met Tyr Pro Glu Asp Gly
 145 150 155 160

Ala Leu Lys Gly Glu Ile Lys Gln Arg Leu Lys Leu Lys Asp Gly Gly
 165 170 175

His Tyr Asp Ala Glu Val Lys Thr Thr Tyr Lys Ala Lys Lys Pro Val
 180 185 190

Gln Leu Pro Gly Ala Tyr Asn Val Asn Ile Lys Leu Asp Ile Thr Ser
 195 200 205

His Asn Glu Asp Tyr Thr Ile Val Glu Gln Tyr Glu Arg Ala Glu Gly
 210 215 220

Arg His Ser Thr Gly Gly Met Asp Glu Leu Tyr Lys Leu Glu
 225 230 235

<210> 19

<211> 9

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 19

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
 1 5

<210> 20

<211> 6

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 20

Gly Ser Gly Ser Gly Ser

1

5

<210> 21

<211> 3917

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 21

Met Pro Pro Leu Leu Ala Pro Leu Leu Cys Leu Ala Leu Leu Pro Ala
 1 5 10 15

Leu Ala Ala Arg Gly Pro Arg Cys Ser Gln Pro Gly Glu Thr Cys Leu
 20 25 30

Asn Gly Gly Lys Cys Glu Ala Ala Asn Gly Thr Glu Ala Cys Val Cys
 35 40 45

Gly Gly Ala Phe Val Gly Pro Arg Cys Gln Asp Pro Asn Pro Cys Leu
 50 55 60

Ser Thr Pro Cys Lys Asn Ala Gly Thr Cys His Val Val Asp Arg Arg
 65 70 75 80

Gly Val Ala Asp Tyr Ala Cys Ser Cys Ala Leu Gly Phe Ser Gly Pro
 85 90 95

Leu Cys Leu Thr Pro Leu Asp Asn Ala Cys Leu Thr Asn Pro Cys Arg
 100 105 110

Asn Gly Gly Thr Cys Asp Leu Leu Thr Leu Thr Glu Tyr Lys Cys Arg
 115 120 125

Cys Pro Pro Gly Trp Ser Gly Lys Ser Cys Gln Gln Ala Asp Pro Cys
 130 135 140

Ala Ser Asn Pro Cys Ala Asn Gly Gly Gln Cys Leu Pro Phe Glu Ala
 145 150 155 160

Ser Tyr Ile Cys His Cys Pro Pro Ser Phe His Gly Pro Thr Cys Arg
 165 170 175

Gln Asp Val Asn Glu Cys Gly Gln Lys Pro Gly Leu Cys Arg His Gly
 180 185 190

Gly Thr Cys His Asn Glu Val Gly Ser Tyr Arg Cys Val Cys Arg Ala

048071

195 200 205

Thr His Thr Gly Pro Asn Cys Glu Arg Pro Tyr Val Pro Cys Ser Pro
 210 215 220

Ser Pro Cys Gln Asn Gly Gly Thr Cys Arg Pro Thr Gly Asp Val Thr
 225 230 235 240

His Glu Cys Ala Cys Leu Pro Gly Phe Thr Gly Gln Asn Cys Glu Glu
 245 250 255

Asn Ile Asp Asp Cys Pro Gly Asn Asn Cys Lys Asn Gly Gly Ala Cys
 260 265 270

Val Asp Gly Val Asn Thr Tyr Asn Cys Arg Cys Pro Pro Glu Trp Thr
 275 280 285

Gly Gln Tyr Cys Thr Glu Asp Val Asp Glu Cys Gln Leu Met Pro Asn
 290 295 300

Ala Cys Gln Asn Gly Gly Thr Cys His Asn Thr His Gly Gly Tyr Asn
 305 310 315 320

Cys Val Cys Val Asn Gly Trp Thr Gly Glu Asp Cys Ser Glu Asn Ile
 325 330 335

Asp Asp Cys Ala Ser Ala Ala Cys Phe His Gly Ala Thr Cys His Asp
 340 345 350

Arg Val Ala Ser Phe Tyr Cys Glu Cys Pro His Gly Arg Thr Gly Leu
 355 360 365

Leu Cys His Leu Asn Asp Ala Cys Ile Ser Asn Pro Cys Asn Glu Gly
 370 375 380

Ser Asn Cys Asp Thr Asn Pro Val Asn Gly Lys Ala Ile Cys Thr Cys
 385 390 395 400

Pro Ser Gly Tyr Thr Gly Pro Ala Cys Ser Gln Asp Val Asp Glu Cys
 405 410 415

Ser Leu Gly Ala Asn Pro Cys Glu His Ala Gly Lys Cys Ile Asn Thr
 420 425 430

Leu Gly Ser Phe Glu Cys Gln Cys Leu Gln Gly Tyr Thr Gly Pro Arg
 435 440 445

048071

Cys Glu Ile Asp Val Asn Glu Cys Val Ser Asn Pro Cys Gln Asn Asp
450 455 460

Ala Thr Cys Leu Asp Gln Ile Gly Glu Phe Gln Cys Ile Cys Met Pro
465 470 475 480

Gly Tyr Glu Gly Val His Cys Glu Val Asn Thr Asp Glu Cys Ala Ser
485 490 495

Ser Pro Cys Leu His Asn Gly Arg Cys Leu Asp Lys Ile Asn Glu Phe
500 505 510 515

Gln Cys Glu Cys Pro Thr Gly Phe Thr Gly His Leu Cys Gln Tyr Asp
515 520 525

Val Asp Glu Cys Ala Ser Thr Pro Cys Lys Asn Gly Ala Lys Cys Leu
530 535 540

Asp Gly Pro Asn Thr Tyr Thr Cys Val Cys Thr Glu Gly Tyr Thr Gly
545 550 555 560

Thr His Cys Glu Val Asp Ile Asp Glu Cys Asp Pro Asp Pro Cys His
565 570 575

Tyr Gly Ser Cys Lys Asp Gly Val Ala Thr Phe Thr Cys Leu Cys Arg
580 585 590

Pro Gly Tyr Thr Gly His His Cys Glu Thr Asn Ile Asn Glu Cys Ser
595 600 605

Ser Gln Pro Cys Arg His Gly Gly Thr Cys Gln Asp Arg Asp Asn Ala
610 615 620

Tyr Leu Cys Phe Cys Leu Lys Gly Thr Thr Gly Pro Asn Cys Glu Ile
625 630 635 640

Asn Leu Asp Asp Cys Ala Ser Ser Pro Cys Asp Ser Gly Thr Cys Leu
645 650 655

Asp Lys Ile Asp Gly Tyr Glu Cys Ala Cys Glu Pro Gly Tyr Thr Gly
660 665 670

Ser Met Cys Asn Ile Asn Ile Asp Glu Cys Ala Gly Asn Pro Cys His
675 680 685

Asn Gly Gly Thr Cys Glu Asp Gly Ile Asn Gly Phe Thr Cys Arg Cys
690 695 700

048071

Pro Glu Gly Tyr His Asp Pro Thr Cys Leu Ser Glu Val Asn Glu Cys
705 710 715 720

Asn Ser Asn Pro Cys Val His Gly Ala Cys Arg Asp Ser Leu Asn Gly
725 730 735

Tyr Lys Cys Asp Cys Asp Pro Gly Trp Ser Gly Thr Asn Cys Asp Ile
740 745 750

Asn Asn Asn Glu Cys Glu Ser Asn Pro Cys Val Asn Gly Gly Thr Cys
755 760 765

Lys Asp Met Thr Ser Gly Tyr Val Cys Thr Cys Arg Glu Gly Phe Ser
770 775 780

Gly Pro Asn Cys Gln Thr Asn Ile Asn Glu Cys Ala Ser Asn Pro Cys
785 790 795 800

Leu Asn Gln Gly Thr Cys Ile Asp Asp Val Ala Gly Tyr Lys Cys Asn
805 810 815

Cys Leu Leu Pro Tyr Thr Gly Ala Thr Cys Glu Val Val Leu Ala Pro
820 825 830

Cys Ala Pro Ser Pro Cys Arg Asn Gly Gly Glu Cys Arg Gln Ser Glu
835 840 845

Asp Tyr Glu Ser Phe Ser Cys Val Cys Pro Thr Gly Trp Gln Ala Gly
850 855 860

Gln Thr Cys Glu Val Asp Ile Asn Glu Cys Val Leu Ser Pro Cys Arg
865 870 875 880

His Gly Ala Ser Cys Gln Asn Thr His Gly Gly Tyr Arg Cys His Cys
885 890 895

Gln Ala Gly Tyr Ser Gly Arg Asn Cys Glu Thr Asp Ile Asp Asp Cys
900 905 910

Arg Pro Asn Pro Cys His Asn Gly Gly Ser Cys Thr Asp Gly Ile Asn
915 920 925

Thr Ala Phe Cys Asp Cys Leu Pro Gly Phe Arg Gly Thr Phe Cys Glu
930 935 940

Glu Asp Ile Asn Glu Cys Ala Ser Asp Pro Cys Arg Asn Gly Ala Asn
945 950 955 960

048071

Cys Thr Asp Cys Val Asp Ser Tyr Thr Cys Thr Cys Pro Ala Gly Phe
 965 970 975

Ser Gly Ile His Cys Glu Asn Asn Thr Pro Asp Cys Thr Glu Ser Ser
 980 985 990

Cys Phe Asn Gly Gly Thr Cys Val Asp Gly Ile Asn Ser Phe Thr Cys
 995 1000 1005

Leu Cys Pro Pro Gly Phe Thr Gly Ser Tyr Cys Gln His Asp Val
 1010 1015 1020

Asn Glu Cys Asp Ser Gln Pro Cys Leu His Gly Gly Thr Cys Gln
 1025 1030 1035

Asp Gly Cys Gly Ser Tyr Arg Cys Thr Cys Pro Gln Gly Tyr Thr
 1040 1045 1050

Gly Pro Asn Cys Gln Asn Leu Val His Trp Cys Asp Ser Ser Pro
 1055 1060 1065

Cys Lys Asn Gly Gly Lys Cys Trp Gln Thr His Thr Gln Tyr Arg
 1070 1075 1080

Cys Glu Cys Pro Ser Gly Trp Thr Gly Leu Tyr Cys Asp Val Pro
 1085 1090 1095

Ser Val Ser Cys Glu Val Ala Ala Gln Arg Gln Gly Val Asp Val
 1100 1105 1110

Ala Arg Leu Cys Gln His Gly Gly Leu Cys Val Asp Ala Gly Asn
 1115 1120 1125

Thr His His Cys Arg Cys Gln Ala Gly Tyr Thr Gly Ser Tyr Cys
 1130 1135 1140

Glu Asp Leu Val Asp Glu Cys Ser Pro Ser Pro Cys Gln Asn Gly
 1145 1150 1155

Ala Thr Cys Thr Asp Tyr Leu Gly Gly Tyr Ser Cys Lys Cys Val
 1160 1165 1170

Ala Gly Tyr His Gly Val Asn Cys Ser Glu Glu Ile Asp Glu Cys
 1175 1180 1185

Leu Ser His Pro Cys Gln Asn Gly Gly Thr Cys Leu Asp Leu Pro

048071

1190						1195						1200			
Asn	Thr	Tyr	Lys	Cys	Ser	Cys	Pro	Arg	Gly	Thr	Gln	Gly	Val	His	
1205						1210					1215				
Cys	Glu	Ile	Asn	Val	Asp	Asp	Cys	Asn	Pro	Pro	Val	Asp	Pro	Val	
1220						1225					1230				
Ser	Arg	Ser	Pro	Lys	Cys	Phe	Asn	Asn	Gly	Thr	Cys	Val	Asp	Gln	
1235						1240					1245				
Val	Gly	Gly	Tyr	Ser	Cys	Thr	Cys	Pro	Pro	Gly	Phe	Val	Gly	Glu	
1250						1255					1260				
Arg	Cys	Glu	Gly	Asp	Val	Asn	Glu	Cys	Leu	Ser	Asn	Pro	Cys	Asp	
1265						1270					1275				
Ala	Arg	Gly	Thr	Gln	Asn	Cys	Val	Gln	Arg	Val	Asn	Asp	Phe	His	
1280						1285					1290				
Cys	Glu	Cys	Arg	Ala	Gly	His	Thr	Gly	Arg	Arg	Cys	Glu	Ser	Val	
1295						1300					1305				
Ile	Asn	Gly	Cys	Lys	Gly	Lys	Pro	Cys	Lys	Asn	Gly	Gly	Thr	Cys	
1310						1315					1320				
Ala	Val	Ala	Ser	Asn	Thr	Ala	Arg	Gly	Phe	Ile	Cys	Lys	Cys	Pro	
1325						1330					1335				
Ala	Gly	Phe	Glu	Gly	Ala	Thr	Cys	Glu	Asn	Asp	Ala	Arg	Thr	Cys	
1340						1345					1350				
Gly	Ser	Leu	Arg	Cys	Leu	Asn	Gly	Gly	Thr	Cys	Ile	Ser	Gly	Pro	
1355						1360					1365				
Arg	Ser	Pro	Thr	Cys	Leu	Cys	Leu	Gly	Pro	Phe	Thr	Gly	Pro	Glu	
1370						1375					1380				
Cys	Gln	Phe	Pro	Ala	Ser	Ser	Pro	Cys	Leu	Gly	Gly	Asn	Pro	Cys	
1385						1390					1395				
Tyr	Asn	Gln	Gly	Thr	Cys	Glu	Pro	Thr	Ser	Glu	Ser	Pro	Phe	Tyr	
1400						1405					1410				
Arg	Cys	Leu	Cys	Pro	Ala	Lys	Phe	Asn	Gly	Leu	Leu	Cys	His	Ile	
1415						1420					1425				

048071

Leu Asp Tyr Ser Phe Gly Gly Gly Ala Gly Arg Asp Ile Pro Pro
 1430 1435 1440

Pro Leu Ile Glu Glu Ala Cys Glu Leu Pro Glu Cys Gln Glu Asp
 1445 1450 1455

Ala Gly Asn Lys Val Cys Ser Leu Gln Cys Asn Asn His Ala Cys
 1460 1465 1470

Gly Trp Asp Gly Gly Asp Cys Ser Leu Asn Phe Asn Asp Pro Trp
 1475 1480 1485

Lys Asn Cys Thr Gln Ser Leu Gln Cys Trp Lys Tyr Phe Ser Asp
 1490 1495 1500

Gly His Cys Asp Ser Gln Cys Asn Ser Ala Gly Cys Leu Phe Asp
 1505 1510 1515

Gly Phe Asp Cys Gln Arg Ala Glu Gly Gln Cys Asn Pro Leu Tyr
 1520 1525 1530

Asp Gln Tyr Cys Lys Asp His Phe Ser Asp Gly His Cys Asp Gln
 1535 1540 1545

Gly Cys Asn Ser Ala Glu Cys Glu Trp Asp Gly Leu Asp Cys Ala
 1550 1555 1560

Glu His Val Pro Glu Arg Leu Ala Ala Gly Thr Leu Val Val Val
 1565 1570 1575

Val Leu Met Pro Pro Glu Gln Leu Arg Asn Ser Ser Phe His Phe
 1580 1585 1590

Leu Arg Glu Leu Ser Arg Val Leu His Thr Asn Val Val Phe Lys
 1595 1600 1605

Arg Asp Ala His Gly Gln Gln Met Ile Phe Pro Tyr Tyr Gly Arg
 1610 1615 1620

Glu Glu Glu Leu Arg Lys His Pro Ile Lys Arg Ala Ala Glu Gly
 1625 1630 1635

Trp Ala Ala Pro Asp Ala Leu Leu Gly Gln Val Lys Ala Ser Leu
 1640 1645 1650

Leu Pro Gly Gly Ser Glu Gly Gly Arg Arg Arg Arg Glu Leu Asp
 1655 1660 1665

048071

Pro Met Asp Val Arg Gly Ser Ile Val Tyr Leu Glu Ile Asp Asn
 1670 1675 1680

Arg Gln Cys Val Gln Ala Ser Ser Gln Cys Phe Gln Ser Ala Thr
 1685 1690 1695

Asp Val Ala Ala Phe Leu Gly Ala Leu Ala Ser Leu Gly Ser Leu
 1700 1705 1710

Asn Ile Pro Tyr Lys Ile Glu Ala Val Gln Ser Glu Thr Val Glu
 1715 1720 1725

Pro Pro Pro Pro Ala Gln Leu His Phe Met Tyr Val Ala Ala Ala
 1730 1735 1740

Ala Phe Val Leu Leu Phe Phe Val Gly Cys Gly Val Leu Leu Ser
 1745 1750 1755

Arg Lys Arg Arg Arg Ala Ser Met Asp Lys Lys Tyr Ser Ile Gly
 1760 1765 1770

Leu Ala Ile Gly Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp
 1775 1780 1785

Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys Val Leu Gly Asn Thr
 1790 1795 1800

Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly Ala Leu Leu Phe
 1805 1810 1815

Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys Arg Thr Ala
 1820 1825 1830

Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr Leu Gln
 1835 1840 1845

Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe Phe
 1850 1855 1860

His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His
 1865 1870 1875

Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
 1880 1885 1890

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val
 1895 1900 1905

048071

Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu
 1910 1915 1920

Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp
 1925 1930 1935

Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu
 1940 1945 1950

Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala
 1955 1960 1965

Ser Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys
 1970 1975 1980

Ser Arg Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys
 1985 1990 1995

Lys Asn Gly Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu
 2000 2005 2010

Thr Pro Asn Phe Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys
 2015 2020 2025

Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu
 2030 2035 2040

Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu Phe Leu Ala Ala
 2045 2050 2055

Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile Leu Arg Val
 2060 2065 2070

Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met Ile Lys
 2075 2080 2085

Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala Leu
 2090 2095 2100

Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp
 2105 2110 2115

Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
 2120 2125 2130

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met

048071

2135						2140						2145			
Asp	Gly	Thr	Glu	Glu	Leu	Leu	Val	Lys	Leu	Asn	Arg	Glu	Asp	Leu	
2150						2155					2160				
Leu	Arg	Lys	Gln	Arg	Thr	Phe	Asp	Asn	Gly	Ser	Ile	Pro	His	Gln	
2165						2170					2175				
Ile	His	Leu	Gly	Glu	Leu	His	Ala	Ile	Leu	Arg	Arg	Gln	Glu	Asp	
2180						2185					2190				
Phe	Tyr	Pro	Phe	Leu	Lys	Asp	Asn	Arg	Glu	Lys	Ile	Glu	Lys	Ile	
2195						2200					2205				
Leu	Thr	Phe	Arg	Ile	Pro	Tyr	Tyr	Val	Gly	Pro	Leu	Ala	Arg	Gly	
2210						2215					2220				
Asn	Ser	Arg	Phe	Ala	Trp	Met	Thr	Arg	Lys	Ser	Glu	Glu	Thr	Ile	
2225						2230					2235				
Thr	Pro	Trp	Asn	Phe	Glu	Glu	Val	Val	Asp	Lys	Gly	Ala	Ser	Ala	
2240						2245					2250				
Gln	Ser	Phe	Ile	Glu	Arg	Met	Thr	Asn	Phe	Asp	Lys	Asn	Leu	Pro	
2255						2260					2265				
Asn	Glu	Lys	Val	Leu	Pro	Lys	His	Ser	Leu	Leu	Tyr	Glu	Tyr	Phe	
2270						2275					2280				
Thr	Val	Tyr	Asn	Glu	Leu	Thr	Lys	Val	Lys	Tyr	Val	Thr	Glu	Gly	
2285						2290					2295				
Met	Arg	Lys	Pro	Ala	Phe	Leu	Ser	Gly	Glu	Gln	Lys	Lys	Ala	Ile	
2300						2305					2310				
Val	Asp	Leu	Leu	Phe	Lys	Thr	Asn	Arg	Lys	Val	Thr	Val	Lys	Gln	
2315						2320					2325				
Leu	Lys	Glu	Asp	Tyr	Phe	Lys	Lys	Ile	Glu	Cys	Phe	Asp	Ser	Val	
2330						2335					2340				
Glu	Ile	Ser	Gly	Val	Glu	Asp	Arg	Phe	Asn	Ala	Ser	Leu	Gly	Thr	
2345						2350					2355				
Tyr	His	Asp	Leu	Leu	Lys	Ile	Ile	Lys	Asp	Lys	Asp	Phe	Leu	Asp	
2360						2365					2370				

048071

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu
 2375 2380 2385
 Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr
 2390 2395 2400
 Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg
 2405 2410 2415
 Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn
 2420 2425 2430
 Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu
 2435 2440 2445
 Lys Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His
 2450 2455 2460
 Asp Asp Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val
 2465 2470 2475
 Ser Gly Gln Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala
 2480 2485 2490
 Gly Ser Pro Ala Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val
 2495 2500 2505
 Val Asp Glu Leu Val Lys Val Met Gly Arg His Lys Pro Glu Asn
 2510 2515 2520
 Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr Thr Gln Lys Gly
 2525 2530 2535
 Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu Glu Gly Ile
 2540 2545 2550
 Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val Glu Asn
 2555 2560 2565
 Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln Asn
 2570 2575 2580
 Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu
 2585 2590 2595
 Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys
 2600 2605 2610

048071

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn
 2615 2620 2625

Arg Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys
 2630 2635 2640

Met Lys Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr
 2645 2650 2655

Gln Arg Lys Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu
 2660 2665 2670

Ser Glu Leu Asp Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu
 2675 2680 2685

Thr Arg Gln Ile Thr Lys His Val Ala Gln Ile Leu Asp Ser Arg
 2690 2695 2700

Met Asn Thr Lys Tyr Asp Glu Asn Asp Lys Leu Ile Arg Glu Val
 2705 2710 2715

Lys Val Ile Thr Leu Lys Ser Lys Leu Val Ser Asp Phe Arg Lys
 2720 2725 2730

Asp Phe Gln Phe Tyr Lys Val Arg Glu Ile Asn Asn Tyr His His
 2735 2740 2745

Ala His Asp Ala Tyr Leu Asn Ala Val Val Gly Thr Ala Leu Ile
 2750 2755 2760

Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val Tyr Gly Asp Tyr
 2765 2770 2775

Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys Ser Glu Gln Glu
 2780 2785 2790

Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr Ser Asn Ile Met
 2795 2800 2805

Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn Gly Glu Ile Arg
 2810 2815 2820

Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr Gly Glu Ile Val
 2825 2830 2835

Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg Lys Val Leu Ser
 2840 2845 2850

048071

Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu Val Gln Thr Gly
 2855 2860 2865

Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg Asn Ser Asp Lys
 2870 2875 2880

Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys Lys Tyr Gly Gly
 2885 2890 2895

Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu Val Val Ala Lys
 2900 2905 2910

Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser Val Lys Glu Leu
 2915 2920 2925

Leu Gly Ile Thr Ile Met Glu Arg Ser Ser Phe Glu Lys Asn Pro
 2930 2935 2940

Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu Val Lys Lys Asp
 2945 2950 2955

Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe Glu Leu Glu Asn
 2960 2965 2970

Gly Arg Lys Arg Met Leu Ala Ser Ala Gly Glu Leu Gln Lys Gly
 2975 2980 2985

Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn Phe Leu Tyr Leu
 2990 2995 3000

Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro Glu Asp Asn Glu
 3005 3010 3015

Gln Lys Gln Leu Phe Val Glu Gln His Lys His Tyr Leu Asp Glu
 3020 3025 3030

Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg Val Ile Leu Ala
 3035 3040 3045

Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr Asn Lys His Arg
 3050 3055 3060

Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile Ile His Leu Phe
 3065 3070 3075

Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe Lys Tyr Phe Asp

048071

3080							3085								3090
Thr	Thr	Ile	Asp	Arg	Lys	Arg	Tyr	Thr	Ser	Thr	Lys	Glu	Val	Leu	
	3095					3100					3105				
Asp	Ala	Thr	Leu	Ile	His	Gln	Ser	Ile	Thr	Gly	Leu	Tyr	Glu	Thr	
	3110					3115					3120				
Arg	Ile	Asp	Leu	Ser	Gln	Leu	Gly	Gly	Asp	Ala	Tyr	Pro	Tyr	Asp	
	3125					3130					3135				
Val	Pro	Asp	Tyr	Ala	Ser	Leu	Gly	Ser	Gly	Asp	Gly	Ile	Gly	Ser	
	3140					3145					3150				
Gly	Ser	Asn	Gly	Ser	Ser	Leu	Asp	Ala	Leu	Asp	Asp	Phe	Asp	Leu	
	3155					3160					3165				
Asp	Met	Leu	Gly	Ser	Asp	Ala	Leu	Asp	Asp	Phe	Asp	Leu	Asp	Met	
	3170					3175					3180				
Leu	Gly	Ser	Asp	Ala	Leu	Asp	Asp	Phe	Asp	Leu	Asp	Met	Leu	Gly	
	3185					3190					3195				
Ser	Asp	Ala	Leu	Asp	Asp	Phe	Asp	Leu	Asp	Met	Leu	Gly	Ser	Gly	
	3200					3205					3210				
Gly	Ser	Gly	Ser	Gln	Tyr	Leu	Pro	Asp	Thr	Asp	Asp	Arg	His	Arg	
	3215					3220					3225				
Ile	Glu	Glu	Lys	Arg	Lys	Arg	Thr	Tyr	Glu	Thr	Phe	Lys	Ser	Ile	
	3230					3235					3240				
Met	Lys	Lys	Ser	Pro	Phe	Ser	Gly	Pro	Thr	Asp	Pro	Arg	Pro	Pro	
	3245					3250					3255				
Pro	Arg	Arg	Ile	Ala	Val	Pro	Ser	Arg	Ser	Ser	Ala	Ser	Val	Pro	
	3260					3265					3270				
Lys	Pro	Ala	Pro	Gln	Pro	Tyr	Pro	Phe	Thr	Ser	Ser	Leu	Ser	Thr	
	3275					3280					3285				
Ile	Asn	Tyr	Asp	Glu	Phe	Pro	Thr	Met	Val	Phe	Pro	Ser	Gly	Gln	
	3290					3295					3300				
Ile	Ser	Gln	Ala	Ser	Ala	Leu	Ala	Pro	Ala	Pro	Pro	Gln	Val	Leu	
	3305					3310					3315				

048071

Pro Gln Ala Pro Ala Pro Ala Pro Ala Pro Ala Met Val Ser Ala
 3320 3325 3330

Leu Ala Gln Ala Pro Ala Pro Val Pro Val Leu Ala Pro Gly Pro
 3335 3340 3345

Pro Gln Ala Val Ala Pro Pro Ala Pro Lys Pro Thr Gln Ala Gly
 3350 3355 3360

Glu Gly Thr Leu Ser Glu Ala Leu Leu Gln Leu Gln Phe Asp Asp
 3365 3370 3375

Glu Asp Leu Gly Ala Leu Leu Gly Asn Ser Thr Asp Pro Ala Val
 3380 3385 3390

Phe Thr Asp Leu Ala Ser Val Asp Asn Ser Glu Phe Gln Gln Leu
 3395 3400 3405

Leu Asn Gln Gly Ile Pro Val Ala Pro His Thr Thr Glu Pro Met
 3410 3415 3420

Leu Met Glu Tyr Pro Glu Ala Ile Thr Arg Leu Val Thr Gly Ala
 3425 3430 3435

Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro Leu Gly Ala Pro Gly
 3440 3445 3450

Leu Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp Phe Ser Ser Ile
 3455 3460 3465

Ala Asp Met Asp Phe Ser Ala Leu Leu Ser Gln Ile Ser Ser Gly
 3470 3475 3480

Ser Gly Ser Gly Ser Arg Asp Ser Arg Glu Gly Met Phe Leu Pro
 3485 3490 3495

Lys Pro Glu Ala Gly Ser Ala Ile Ser Asp Val Phe Glu Gly Arg
 3500 3505 3510

Glu Val Cys Gln Pro Lys Arg Ile Arg Pro Phe His Pro Pro Gly
 3515 3520 3525

Ser Pro Trp Ala Asn Arg Pro Leu Pro Ala Ser Leu Ala Pro Thr
 3530 3535 3540

Pro Thr Gly Pro Val His Glu Pro Val Gly Ser Leu Thr Pro Ala
 3545 3550 3555

048071

Pro Val Pro Gln Pro Leu Asp Pro Ala Pro Ala Val Thr Pro Glu
 3560 3565 3570

Ala Ser His Leu Leu Glu Asp Pro Asp Glu Glu Thr Ser Gln Ala
 3575 3580 3585

Val Lys Ala Leu Arg Glu Met Ala Asp Thr Val Ile Pro Gln Lys
 3590 3595 3600

Glu Glu Ala Ala Ile Cys Gly Gln Met Asp Leu Ser His Pro Pro
 3605 3610 3615

Pro Arg Gly His Leu Asp Glu Leu Thr Thr Thr Leu Glu Ser Met
 3620 3625 3630

Thr Glu Asp Leu Asn Leu Asp Ser Pro Leu Thr Pro Glu Leu Asn
 3635 3640 3645

Glu Ile Leu Asp Thr Phe Leu Asn Asp Glu Cys Leu Leu His Ala
 3650 3655 3660

Met His Ile Ser Thr Gly Leu Ser Ile Phe Asp Thr Ser Leu Phe
 3665 3670 3675

Leu Met Val Ser Lys Gly Glu Glu Asp Asn Met Ala Ile Ile Lys
 3680 3685 3690

Glu Phe Met Arg Phe Lys Val His Met Glu Gly Ser Val Asn Gly
 3695 3700 3705

His Glu Phe Glu Ile Glu Gly Glu Gly Glu Gly Arg Pro Tyr Glu
 3710 3715 3720

Gly Thr Gln Thr Ala Lys Leu Lys Val Thr Lys Gly Gly Pro Leu
 3725 3730 3735

Pro Phe Ala Trp Asp Ile Leu Ser Pro Gln Phe Met Tyr Gly Ser
 3740 3745 3750

Lys Ala Tyr Val Lys His Pro Ala Asp Ile Pro Asp Tyr Leu Lys
 3755 3760 3765

Leu Ser Phe Pro Glu Gly Phe Lys Trp Glu Arg Val Met Asn Phe
 3770 3775 3780

Glu Asp Gly Gly Val Val Thr Val Thr Gln Asp Ser Ser Leu Gln
 3785 3790 3795

048071

Asp Gly Glu Phe Ile Tyr Lys Val Lys Leu Arg Gly Thr Asn Phe
3800 3805 3810

Pro Ser Asp Gly Pro Val Met Gln Lys Lys Thr Met Gly Trp Glu
3815 3820 3825

Ala Ser Ser Glu Arg Met Tyr Pro Glu Asp Gly Ala Leu Lys Gly
3830 3835 3840

Glu Ile Lys Gln Arg Leu Lys Leu Lys Asp Gly Gly His Tyr Asp
3845 3850 3855

Ala Glu Val Lys Thr Thr Tyr Lys Ala Lys Lys Pro Val Gln Leu
3860 3865 3870

Pro Gly Ala Tyr Asn Val Asn Ile Lys Leu Asp Ile Thr Ser His
3875 3880 3885

Asn Glu Asp Tyr Thr Ile Val Glu Gln Tyr Glu Arg Ala Glu Gly
3890 3895 3900

Arg His Ser Thr Gly Gly Met Asp Glu Leu Tyr Lys Leu Glu
3905 3910 3915

<210> 22

<211> 1734

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 22

Met Pro Pro Leu Leu Ala Pro Leu Leu Cys Leu Ala Leu Leu Pro Ala
1 5 10 15

Leu Ala Ala Arg Gly Pro Arg Cys Ser Gln Pro Gly Glu Thr Cys Leu
20 25 30

Asn Gly Gly Lys Cys Glu Ala Ala Asn Gly Thr Glu Ala Cys Val Cys
35 40 45

Gly Gly Ala Phe Val Gly Pro Arg Cys Gln Asp Pro Asn Pro Cys Leu
50 55 60

Ser Thr Pro Cys Lys Asn Ala Gly Thr Cys His Val Val Asp Arg Arg
65 70 75 80

048071

Gly Val Ala Asp Tyr Ala Cys Ser Cys Ala Leu Gly Phe Ser Gly Pro
85 90 95

Leu Cys Leu Thr Pro Leu Asp Asn Ala Cys Leu Thr Asn Pro Cys Arg
100 105 110

Asn Gly Gly Thr Cys Asp Leu Leu Thr Leu Thr Glu Tyr Lys Cys Arg
115 120 125

Cys Pro Pro Gly Trp Ser Gly Lys Ser Cys Gln Gln Ala Asp Pro Cys
130 135 140

Ala Ser Asn Pro Cys Ala Asn Gly Gly Gln Cys Leu Pro Phe Glu Ala
145 150 155 160

Ser Tyr Ile Cys His Cys Pro Pro Ser Phe His Gly Pro Thr Cys Arg
165 170 175

Gln Asp Val Asn Glu Cys Gly Gln Lys Pro Gly Leu Cys Arg His Gly
180 185 190

Gly Thr Cys His Asn Glu Val Gly Ser Tyr Arg Cys Val Cys Arg Ala
195 200 205

Thr His Thr Gly Pro Asn Cys Glu Arg Pro Tyr Val Pro Cys Ser Pro
210 215 220

Ser Pro Cys Gln Asn Gly Gly Thr Cys Arg Pro Thr Gly Asp Val Thr
225 230 235 240

His Glu Cys Ala Cys Leu Pro Gly Phe Thr Gly Gln Asn Cys Glu Glu
245 250 255

Asn Ile Asp Asp Cys Pro Gly Asn Asn Cys Lys Asn Gly Gly Ala Cys
260 265 270

Val Asp Gly Val Asn Thr Tyr Asn Cys Arg Cys Pro Pro Glu Trp Thr
275 280 285

Gly Gln Tyr Cys Thr Glu Asp Val Asp Glu Cys Gln Leu Met Pro Asn
290 295 300

Ala Cys Gln Asn Gly Gly Thr Cys His Asn Thr His Gly Gly Tyr Asn
305 310 315 320

Cys Val Cys Val Asn Gly Trp Thr Gly Glu Asp Cys Ser Glu Asn Ile
325 330 335

048071

Asp Asp Cys Ala Ser Ala Ala Cys Phe His Gly Ala Thr Cys His Asp
340 345 350

Arg Val Ala Ser Phe Tyr Cys Glu Cys Pro His Gly Arg Thr Gly Leu
355 360 365

Leu Cys His Leu Asn Asp Ala Cys Ile Ser Asn Pro Cys Asn Glu Gly
370 375 380

Ser Asn Cys Asp Thr Asn Pro Val Asn Gly Lys Ala Ile Cys Thr Cys
385 390 395 400

Pro Ser Gly Tyr Thr Gly Pro Ala Cys Ser Gln Asp Val Asp Glu Cys
405 410 415

Ser Leu Gly Ala Asn Pro Cys Glu His Ala Gly Lys Cys Ile Asn Thr
420 425 430

Leu Gly Ser Phe Glu Cys Gln Cys Leu Gln Gly Tyr Thr Gly Pro Arg
435 440 445

Cys Glu Ile Asp Val Asn Glu Cys Val Ser Asn Pro Cys Gln Asn Asp
450 455 460

Ala Thr Cys Leu Asp Gln Ile Gly Glu Phe Gln Cys Ile Cys Met Pro
465 470 475 480

Gly Tyr Glu Gly Val His Cys Glu Val Asn Thr Asp Glu Cys Ala Ser
485 490 495

Ser Pro Cys Leu His Asn Gly Arg Cys Leu Asp Lys Ile Asn Glu Phe
500 505 510

Gln Cys Glu Cys Pro Thr Gly Phe Thr Gly His Leu Cys Gln Tyr Asp
515 520 525

Val Asp Glu Cys Ala Ser Thr Pro Cys Lys Asn Gly Ala Lys Cys Leu
530 535 540

Asp Gly Pro Asn Thr Tyr Thr Cys Val Cys Thr Glu Gly Tyr Thr Gly
545 550 555 560

Thr His Cys Glu Val Asp Ile Asp Glu Cys Asp Pro Asp Pro Cys His
565 570 575

Tyr Gly Ser Cys Lys Asp Gly Val Ala Thr Phe Thr Cys Leu Cys Arg

048071

580 585 590
 Pro Gly Tyr Thr Gly His His Cys Glu Thr Asn Ile Asn Glu Cys Ser
 595 600 605
 Ser Gln Pro Cys Arg His Gly Gly Thr Cys Gln Asp Arg Asp Asn Ala
 610 615 620
 Tyr Leu Cys Phe Cys Leu Lys Gly Thr Thr Gly Pro Asn Cys Glu Ile
 625 630 635 640
 Asn Leu Asp Asp Cys Ala Ser Ser Pro Cys Asp Ser Gly Thr Cys Leu
 645 650 655
 Asp Lys Ile Asp Gly Tyr Glu Cys Ala Cys Glu Pro Gly Tyr Thr Gly
 660 665 670
 Ser Met Cys Asn Ile Asn Ile Asp Glu Cys Ala Gly Asn Pro Cys His
 675 680 685
 Asn Gly Gly Thr Cys Glu Asp Gly Ile Asn Gly Phe Thr Cys Arg Cys
 690 695 700
 Pro Glu Gly Tyr His Asp Pro Thr Cys Leu Ser Glu Val Asn Glu Cys
 705 710 715 720
 Asn Ser Asn Pro Cys Val His Gly Ala Cys Arg Asp Ser Leu Asn Gly
 725 730 735
 Tyr Lys Cys Asp Cys Asp Pro Gly Trp Ser Gly Thr Asn Cys Asp Ile
 740 745 750
 Asn Asn Asn Glu Cys Glu Ser Asn Pro Cys Val Asn Gly Gly Thr Cys
 755 760 765
 Lys Asp Met Thr Ser Gly Tyr Val Cys Thr Cys Arg Glu Gly Phe Ser
 770 775 780
 Gly Pro Asn Cys Gln Thr Asn Ile Asn Glu Cys Ala Ser Asn Pro Cys
 785 790 795 800
 Leu Asn Gln Gly Thr Cys Ile Asp Asp Val Ala Gly Tyr Lys Cys Asn
 805 810 815
 Cys Leu Leu Pro Tyr Thr Gly Ala Thr Cys Glu Val Val Leu Ala Pro
 820 825 830

048071

Cys Ala Pro Ser Pro Cys Arg Asn Gly Gly Glu Cys Arg Gln Ser Glu
835 840 845

Asp Tyr Glu Ser Phe Ser Cys Val Cys Pro Thr Gly Trp Gln Ala Gly
850 855 860

Gln Thr Cys Glu Val Asp Ile Asn Glu Cys Val Leu Ser Pro Cys Arg
865 870 875 880

His Gly Ala Ser Cys Gln Asn Thr His Gly Gly Tyr Arg Cys His Cys
885 890 895

Gln Ala Gly Tyr Ser Gly Arg Asn Cys Glu Thr Asp Ile Asp Asp Cys
900 905 910

Arg Pro Asn Pro Cys His Asn Gly Gly Ser Cys Thr Asp Gly Ile Asn
915 920 925

Thr Ala Phe Cys Asp Cys Leu Pro Gly Phe Arg Gly Thr Phe Cys Glu
930 935 940

Glu Asp Ile Asn Glu Cys Ala Ser Asp Pro Cys Arg Asn Gly Ala Asn
945 950 955 960 965

Cys Thr Asp Cys Val Asp Ser Tyr Thr Cys Thr Cys Pro Ala Gly Phe
965 970 975

Ser Gly Ile His Cys Glu Asn Asn Thr Pro Asp Cys Thr Glu Ser Ser
980 985 990

Cys Phe Asn Gly Gly Thr Cys Val Asp Gly Ile Asn Ser Phe Thr Cys
995 1000 1005

Leu Cys Pro Pro Gly Phe Thr Gly Ser Tyr Cys Gln His Asp Val
1010 1015 1020

Asn Glu Cys Asp Ser Gln Pro Cys Leu His Gly Gly Thr Cys Gln
1025 1030 1035

Asp Gly Cys Gly Ser Tyr Arg Cys Thr Cys Pro Gln Gly Tyr Thr
1040 1045 1050

Gly Pro Asn Cys Gln Asn Leu Val His Trp Cys Asp Ser Ser Pro
1055 1060 1065

Cys Lys Asn Gly Gly Lys Cys Trp Gln Thr His Thr Gln Tyr Arg
1070 1075 1080

048071

Cys Glu Cys Pro Ser Gly Trp Thr Gly Leu Tyr Cys Asp Val Pro
 1085 1090 1095

Ser Val Ser Cys Glu Val Ala Ala Gln Arg Gln Gly Val Asp Val
 1100 1105 1110

Ala Arg Leu Cys Gln His Gly Gly Leu Cys Val Asp Ala Gly Asn
 1115 1120 1125

Thr His His Cys Arg Cys Gln Ala Gly Tyr Thr Gly Ser Tyr Cys
 1130 1135 1140

Glu Asp Leu Val Asp Glu Cys Ser Pro Ser Pro Cys Gln Asn Gly
 1145 1150 1155

Ala Thr Cys Thr Asp Tyr Leu Gly Gly Tyr Ser Cys Lys Cys Val
 1160 1165 1170

Ala Gly Tyr His Gly Val Asn Cys Ser Glu Glu Ile Asp Glu Cys
 1175 1180 1185

Leu Ser His Pro Cys Gln Asn Gly Gly Thr Cys Leu Asp Leu Pro
 1190 1195 1200

Asn Thr Tyr Lys Cys Ser Cys Pro Arg Gly Thr Gln Gly Val His
 1205 1210 1215

Cys Glu Ile Asn Val Asp Asp Cys Asn Pro Pro Val Asp Pro Val
 1220 1225 1230

Ser Arg Ser Pro Lys Cys Phe Asn Asn Gly Thr Cys Val Asp Gln
 1235 1240 1245

Val Gly Gly Tyr Ser Cys Thr Cys Pro Pro Gly Phe Val Gly Glu
 1250 1255 1260

Arg Cys Glu Gly Asp Val Asn Glu Cys Leu Ser Asn Pro Cys Asp
 1265 1270 1275

Ala Arg Gly Thr Gln Asn Cys Val Gln Arg Val Asn Asp Phe His
 1280 1285 1290

Cys Glu Cys Arg Ala Gly His Thr Gly Arg Arg Cys Glu Ser Val
 1295 1300 1305

Ile Asn Gly Cys Lys Gly Lys Pro Cys Lys Asn Gly Gly Thr Cys
 1310 1315 1320

048071

Ala Val Ala Ser Asn Thr Ala Arg Gly Phe Ile Cys Lys Cys Pro
1325 1330 1335

Ala Gly Phe Glu Gly Ala Thr Cys Glu Asn Asp Ala Arg Thr Cys
1340 1345 1350

Gly Ser Leu Arg Cys Leu Asn Gly Gly Thr Cys Ile Ser Gly Pro
1355 1360 1365

Arg Ser Pro Thr Cys Leu Cys Leu Gly Pro Phe Thr Gly Pro Glu
1370 1375 1380

Cys Gln Phe Pro Ala Ser Ser Pro Cys Leu Gly Gly Asn Pro Cys
1385 1390 1395

Tyr Asn Gln Gly Thr Cys Glu Pro Thr Ser Glu Ser Pro Phe Tyr
1400 1405 1410

Arg Cys Leu Cys Pro Ala Lys Phe Asn Gly Leu Leu Cys His Ile
1415 1420 1425

Leu Asp Tyr Ser Phe Gly Gly Gly Ala Gly Arg Asp Ile Pro Pro
1430 1435 1440

Pro Leu Ile Glu Glu Ala Cys Glu Leu Pro Glu Cys Gln Glu Asp
1445 1450 1455

Ala Gly Asn Lys Val Cys Ser Leu Gln Cys Asn Asn His Ala Cys
1460 1465 1470

Gly Trp Asp Gly Gly Asp Cys Ser Leu Asn Phe Asn Asp Pro Trp
1475 1480 1485

Lys Asn Cys Thr Gln Ser Leu Gln Cys Trp Lys Tyr Phe Ser Asp
1490 1495 1500

Gly His Cys Asp Ser Gln Cys Asn Ser Ala Gly Cys Leu Phe Asp
1505 1510 1515

Gly Phe Asp Cys Gln Arg Ala Glu Gly Gln Cys Asn Pro Leu Tyr
1520 1525 1530

Asp Gln Tyr Cys Lys Asp His Phe Ser Asp Gly His Cys Asp Gln
1535 1540 1545

Gly Cys Asn Ser Ala Glu Cys Glu Trp Asp Gly Leu Asp Cys Ala

048071

1550 1555 1560

Glu His Val Pro Glu Arg Leu Ala Ala Gly Thr Leu Val Val Val
1565 1570 1575

Val Leu Met Pro Pro Glu Gln Leu Arg Asn Ser Ser Phe His Phe
1580 1585 1590

Leu Arg Glu Leu Ser Arg Val Leu His Thr Asn Val Val Phe Lys
1595 1600 1605

Arg Asp Ala His Gly Gln Gln Met Ile Phe Pro Tyr Tyr Gly Arg
1610 1615 1620

Glu Glu Glu Leu Arg Lys His Pro Ile Lys Arg Ala Ala Glu Gly
1625 1630 1635

Trp Ala Ala Pro Asp Ala Leu Leu Gly Gln Val Lys Ala Ser Leu
1640 1645 1650

Leu Pro Gly Gly Ser Glu Gly Gly Arg Arg Arg Arg Glu Leu Asp
1655 1660 1665

Pro Met Asp Val Arg Gly Ser Ile Val Tyr Leu Glu Ile Asp Asn
1670 1675 1680

Arg Gln Cys Val Gln Ala Ser Ser Gln Cys Phe Gln Ser Ala Thr
1685 1690 1695

Asp Val Ala Ala Phe Leu Gly Ala Leu Ala Ser Leu Gly Ser Leu
1700 1705 1710

Asn Ile Pro Tyr Lys Ile Glu Ala Val Gln Ser Glu Thr Val Glu
1715 1720 1725

Pro Pro Pro Pro Ala Gln
1730

<210> 23

<211> 20

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 23

Leu His Phe Met Tyr Val Ala Ala Ala Ala Phe Val Leu Leu Phe Phe

048071

1 5 10 15

Val Gly Cys Gly
20

<210> 24

<211> 9

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 24

Val Leu Leu Ser Arg Lys Arg Arg Arg
1 5

<210> 25

<211> 1367

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 25

Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val Gly
1 5 10 15

Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys
20 25 30

Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly
35 40 45

Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys
50 55 60

Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr
65 70 75 80

Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe
85 90 95

Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His
100 105 110

Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His
115 120 125

048071

Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser
 130 135 140

Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met
 145 150 155 160

Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp
 165 170 175

Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn
 180 185 190

Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys
 195 200 205

Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu
 210 215 220

Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu
 225 230 235 240

Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp
 245 250 255

Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp
 260 265 270

Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu
 275 280 285

Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile
 290 295 300

Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met
 305 310 315 320

Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala
 325 330 335

Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp
 340 345 350

Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln
 355 360 365

Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly
 370 375 380

048071

Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys
385 390 395 400

Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly
405 410 415

Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu
420 425 430

Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro
435 440 445

Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met
450 455 460

Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val
465 470 475 480

Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn
485 490 495

Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu
500 505 510

Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr
515 520 525

Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys
530 535 540

Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val
545 550 555 560

Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser
565 570 575

Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr
580 585 590

Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn
595 600 605

Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu
610 615 620

Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His

048071

Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe
885 890 895

Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys
900 905 910

Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys
915 920 925

His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu
930 935 940

Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys
945 950 955 960

Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu
965 970 975

Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val Val
980 985 990

Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val
995 1000 1005

Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys
1010 1015 1020

Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr
1025 1030 1035

Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn
1040 1045 1050

Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr
1055 1060 1065

Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg
1070 1075 1080

Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu
1085 1090 1095

Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg
1100 1105 1110

Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys
1115 1120 1125

048071

Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu
 1130 1135 1140

 Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser
 1145 1150 1155

 Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser Phe
 1160 1165 1170

 Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu
 1175 1180 1185

 Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe
 1190 1195 1200

 Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly Glu
 1205 1210 1215

 Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn
 1220 1225 1230

 Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro
 1235 1240 1245

 Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His
 1250 1255 1260

 Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg
 1265 1270 1275

 Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr
 1280 1285 1290

 Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile
 1295 1300 1305

 Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe
 1310 1315 1320

 Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr
 1325 1330 1335

 Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr Gly
 1340 1345 1350

 Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
 1355 1360 1365

048071

<210> 26

<211> 530

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 26

Asp Gly Ile Gly Ser Gly Ser Asn Gly Ser Ser Leu Asp Ala Leu Asp
1 5 10 15

Asp Phe Asp Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp
20 25 30

Leu Asp Met Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met
35 40 45

Leu Gly Ser Asp Ala Leu Asp Asp Phe Asp Leu Asp Met Leu Gly Ser
50 55 60

Gly Gly Ser Gly Ser Gln Tyr Leu Pro Asp Thr Asp Asp Arg His Arg
65 70 75 80

Ile Glu Glu Lys Arg Lys Arg Thr Tyr Glu Thr Phe Lys Ser Ile Met
85 90 95

Lys Lys Ser Pro Phe Ser Gly Pro Thr Asp Pro Arg Pro Pro Pro Arg
100 105 110

Arg Ile Ala Val Pro Ser Arg Ser Ser Ala Ser Val Pro Lys Pro Ala
115 120 125

Pro Gln Pro Tyr Pro Phe Thr Ser Ser Leu Ser Thr Ile Asn Tyr Asp
130 135 140

Glu Phe Pro Thr Met Val Phe Pro Ser Gly Gln Ile Ser Gln Ala Ser
145 150 155 160

Ala Leu Ala Pro Ala Pro Pro Gln Val Leu Pro Gln Ala Pro Ala Pro
165 170 175

Ala Pro Ala Pro Ala Met Val Ser Ala Leu Ala Gln Ala Pro Ala Pro
180 185 190

Val Pro Val Leu Ala Pro Gly Pro Pro Gln Ala Val Ala Pro Pro Ala
195 200 205

048071

Pro Lys Pro Thr Gln Ala Gly Glu Gly Thr Leu Ser Glu Ala Leu Leu
 210 215 220

Gln Leu Gln Phe Asp Asp Glu Asp Leu Gly Ala Leu Leu Gly Asn Ser
 225 230 235 240

Thr Asp Pro Ala Val Phe Thr Asp Leu Ala Ser Val Asp Asn Ser Glu
 245 250 255

Phe Gln Gln Leu Leu Asn Gln Gly Ile Pro Val Ala Pro His Thr Thr
 260 265 270

Glu Pro Met Leu Met Glu Tyr Pro Glu Ala Ile Thr Arg Leu Val Thr
 275 280 285

Gly Ala Gln Arg Pro Pro Asp Pro Ala Pro Ala Pro Leu Gly Ala Pro
 290 295 300

Gly Leu Pro Asn Gly Leu Leu Ser Gly Asp Glu Asp Phe Ser Ser Ile
 305 310 315 320

Ala Asp Met Asp Phe Ser Ala Leu Leu Ser Gln Ile Ser Ser Gly Ser
 325 330 335

Gly Ser Gly Ser Arg Asp Ser Arg Glu Gly Met Phe Leu Pro Lys Pro
 340 345 350

Glu Ala Gly Ser Ala Ile Ser Asp Val Phe Glu Gly Arg Glu Val Cys
 355 360 365

Gln Pro Lys Arg Ile Arg Pro Phe His Pro Pro Gly Ser Pro Trp Ala
 370 375 380

Asn Arg Pro Leu Pro Ala Ser Leu Ala Pro Thr Pro Thr Gly Pro Val
 385 390 395 400

His Glu Pro Val Gly Ser Leu Thr Pro Ala Pro Val Pro Gln Pro Leu
 405 410 415

Asp Pro Ala Pro Ala Val Thr Pro Glu Ala Ser His Leu Leu Glu Asp
 420 425 430

Pro Asp Glu Glu Thr Ser Gln Ala Val Lys Ala Leu Arg Glu Met Ala
 435 440 445

Asp Thr Val Ile Pro Gln Lys Glu Glu Ala Ala Ile Cys Gly Gln Met

048071

450 455 460

Asp Leu Ser His Pro Pro Pro Arg Gly His Leu Asp Glu Leu Thr Thr
 465 470 475 480

Thr Leu Glu Ser Met Thr Glu Asp Leu Asn Leu Asp Ser Pro Leu Thr
 485 490 495

Pro Glu Leu Asn Glu Ile Leu Asp Thr Phe Leu Asn Asp Glu Cys Leu
 500 505 510

Leu His Ala Met His Ile Ser Thr Gly Leu Ser Ile Phe Asp Thr Ser
 515 520 525

Leu Phe
 530

<210> 27

<211> 238

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 27

Met Val Ser Lys Gly Glu Glu Asp Asn Met Ala Ile Ile Lys Glu Phe
 1 5 10 15

Met Arg Phe Lys Val His Met Glu Gly Ser Val Asn Gly His Glu Phe
 20 25 30

Glu Ile Glu Gly Glu Gly Glu Gly Arg Pro Tyr Glu Gly Thr Gln Thr
 35 40 45

Ala Lys Leu Lys Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp
 50 55 60

Ile Leu Ser Pro Gln Phe Met Tyr Gly Ser Lys Ala Tyr Val Lys His
 65 70 75 80

Pro Ala Asp Ile Pro Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Phe
 85 90 95

Lys Trp Glu Arg Val Met Asn Phe Glu Asp Gly Gly Val Val Thr Val
 100 105 110

Thr Gln Asp Ser Ser Leu Gln Asp Gly Glu Phe Ile Tyr Lys Val Lys

048071

115 120 125

Leu Arg Gly Thr Asn Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys
 130 135 140

Thr Met Gly Trp Glu Ala Ser Ser Glu Arg Met Tyr Pro Glu Asp Gly
 145 150 155 160

Ala Leu Lys Gly Glu Ile Lys Gln Arg Leu Lys Leu Lys Asp Gly Gly
 165 170 175

His Tyr Asp Ala Glu Val Lys Thr Thr Tyr Lys Ala Lys Lys Pro Val
 180 185 190

Gln Leu Pro Gly Ala Tyr Asn Val Asn Ile Lys Leu Asp Ile Thr Ser
 195 200 205

His Asn Glu Asp Tyr Thr Ile Val Glu Gln Tyr Glu Arg Ala Glu Gly
 210 215 220

Arg His Ser Thr Gly Gly Met Asp Glu Leu Tyr Lys Leu Glu
 225 230 235

<210> 28

<211> 9

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 28

Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
 1 5

<210> 29

<211> 11

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 29

Asp Gly Ile Gly Ser Gly Ser Asn Gly Ser Ser
 1 5 10

<210> 30

<211> 12

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 30

Arg Ser Gln Gln Glu Ala Ala Ala Lys Lys Phe Phe
1 5 10

<210> 31

<211> 24

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 31

Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu
1 5 10 15

Ser Arg Lys Leu Ile Asn Gly Ile
 20

<210> 32

<211> 24

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: синтетическая конструкция

<400> 32

Val Met Ala Gln Leu Lys Ala Ala Ala Tyr Thr Gly Trp Gly Arg Leu
1 5 10 15

Ser Ala Ala Leu Ile Asn Gly Ile
 20

<210> 33

<211> 113

<212> ДНК

<213> Streptococcus pyogenes

<220>

<221> модифицированное основание

<222> (1)..(20)

<223> а, с, т, г, неизвестный или другой

<400> 33

nnnnnnnnnn nnnnnnnnnn gtttaagagc tatgctggaa acagcatagc aagtttaaat 60

aaggctagtc cgttatcaac ttgaaaaagt ggcaccgagt cggtgctttt ttt 113

<210> 34
 <211> 104
 <212> ДНК
 <213> Staphylococcus aureus

<220>
 <221> модифицированное основание
 <222> (1)..(23)
 <223> а, с, т, г, неизвестный или другой

<400> 34
 nnnnnnnnnn nnnnnnnnnn nnnngtttttag tactctggaa acagaatcta ctaaaacaag 60
 gcaaaatgcc gtgtttatct cgtcaacttg ttggcgagat tttt 104

<210> 35
 <211> 19
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетическая конструкция

<400> 35
 gtactccgac ctctagtgt 19

<210> 36
 <211> 23
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетическая конструкция

<400> 36
 ggtgcscttc cgccatttt ccc 23

<210> 37
 <211> 20
 <212> ДНК
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетическая конструкция

<400> 37
 гассaggatg ggcассассс 20

<210> 38
 <211> 20
 <212> ДНК
 <213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая конструкция

<400> 38

ggctggcgag cgcggcctta

20

<210> 39

<211> 326

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая конструкция

<400> 39

Ile Leu Asp Tyr Ser Phe Thr Gly Gly Ala Gly Arg Asp Ile Pro Pro
1 5 10 15Pro Gln Ile Glu Glu Ala Cys Glu Leu Pro Glu Cys Gln Val Asp Ala
20 25 30Gly Asn Lys Val Cys Asn Leu Gln Cys Asn Asn His Ala Cys Gly Trp
35 40 45Asp Gly Gly Asp Cys Ser Leu Asn Phe Asn Asp Pro Trp Lys Asn Cys
50 55 60Thr Gln Ser Leu Gln Cys Trp Lys Tyr Phe Ser Asp Gly His Cys Asp
65 70 75 80Ser Gln Cys Asn Ser Ala Gly Cys Leu Phe Asp Gly Phe Asp Cys Gln
85 90 95Leu Thr Glu Gly Gln Cys Asn Pro Leu Tyr Asp Gln Tyr Cys Lys Asp
100 105 110His Phe Ser Asp Gly His Cys Asp Gln Gly Cys Asn Ser Ala Glu Cys
115 120 125Glu Trp Asp Gly Leu Asp Cys Ala Glu His Val Pro Glu Arg Leu Ala
130 135 140Ala Gly Thr Leu Val Leu Val Val Leu Leu Pro Pro Asp Gln Leu Arg
145 150 155 160Asn Asn Ser Phe His Phe Leu Arg Glu Leu Ser His Val Leu His Thr
165 170 175

Asn Val Val Phe Lys Arg Asp Ala Gln Gly Gln Gln Met Ile Phe Pro

048071

	180		185		190													
Tyr	Tyr	Gly	His	Glu	Glu	Glu	Leu	Arg	Lys	His	Pro	Ile	Lys	Arg	Ser			
		195					200					205						
Thr	Val	Gly	Trp	Ala	Thr	Ser	Ser	Leu	Leu	Pro	Gly	Thr	Ser	Gly	Gly			
	210					215					220							
Arg	Gln	Arg	Arg	Glu	Leu	Asp	Pro	Met	Asp	Ile	Arg	Gly	Ser	Ile	Val			
225					230					235					240			
Tyr	Leu	Glu	Ile	Asp	Asn	Arg	Gln	Cys	Val	Gln	Ser	Ser	Ser	Gln	Cys			
				245					250					255				
Phe	Gln	Ser	Ala	Thr	Asp	Val	Ala	Ala	Phe	Leu	Gly	Ala	Leu	Ala	Ser			
			260					265					270					
Leu	Gly	Ser	Leu	Asn	Ile	Pro	Tyr	Lys	Ile	Glu	Ala	Val	Lys	Ser	Glu			
		275					280					285						
Pro	Val	Glu	Pro	Pro	Leu	Pro	Ser	Gln	Leu	His	Leu	Met	Tyr	Val	Ala			
	290					295					300							
Ala	Ala	Ala	Phe	Val	Leu	Leu	Phe	Phe	Val	Gly	Cys	Gly	Val	Leu	Leu			
305					310					315					320			
Ser	Arg	Lys	Arg	Arg	Arg													
					325													

<210> 40

<211> 7

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая конструкция

<400> 40

Pro Lys Lys Lys Arg Lys Val

1

5

<210> 41

<211> 16

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая конструкция

<400> 41

048071

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys
1 5 10 15

<210> 42

<211> 9

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая конструкция

<400> 42

Pro Ala Ala Lys Arg Val Lys Leu Asp
1 5

<210> 43

<211> 11

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая конструкция

<400> 43

Arg Gln Arg Arg Asn Glu Leu Lys Arg Ser Pro
1 5 10

<210> 44

<211> 38

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая конструкция

<400> 44

Asn Gln Ser Ser Asn Phe Gly Pro Met Lys Gly Gly Asn Phe Gly Gly
1 5 10 15

Arg Ser Ser Gly Pro Tyr Gly Gly Gly Gly Gln Tyr Phe Ala Lys Pro
20 25 30

Arg Asn Gln Gly Gly Tyr
35

<210> 45

<211> 42

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая конструкция

048071

<400> 45

Arg Met Arg Ile Glx Phe Lys Asn Lys Gly Lys Asp Thr Ala Glu Leu
1 5 10 15

Arg Arg Arg Arg Val Glu Val Ser Val Glu Leu Arg Lys Ala Lys Lys
20 25 30

Asp Glu Gln Ile Leu Lys Arg Arg Asn Val
35 40

<210> 46

<211> 8

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая
конструкция

<400> 46

Val Ser Arg Lys Arg Pro Arg Pro
1 5

<210> 47

<211> 8

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая
конструкция

<400> 47

Pro Pro Lys Lys Ala Arg Glu Asp
1 5

<210> 48

<211> 8

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая
конструкция

<400> 48

Pro Gln Pro Lys Lys Lys Pro Leu
1 5

<210> 49

<211> 12

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая
конструкция

<400> 49
Ser Ala Leu Ile Lys Lys Lys Lys Lys Met Ala Pro
1 5 10

<210> 50
<211> 5
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
конструкция

<400> 50
Asp Arg Leu Arg Arg
1 5

<210> 51
<211> 7
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
конструкция

<400> 51
Pro Lys Gln Lys Lys Arg Lys
1 5

<210> 52
<211> 10
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
конструкция

<400> 52
Arg Lys Leu Lys Lys Lys Ile Lys Lys Leu
1 5 10

<210> 53
<211> 10
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
конструкция

<400> 53
Arg Glu Lys Lys Lys Phe Leu Lys Arg Arg
1 5 10

<210> 54

<211> 20
 <212> белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетическая конструкция

<400> 54
 Lys Arg Lys Gly Asp Glu Val Asp Gly Val Asp Glu Val Ala Lys Lys
 1 5 10 15

Lys Ser Lys Lys
 20

<210> 55
 <211> 17
 <212> белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетическая конструкция

<400> 55
 Arg Lys Cys Leu Gln Ala Gly Met Asn Leu Glu Ala Arg Lys Thr Lys
 1 5 10 15

Lys

<210> 56
 <211> 236
 <212> белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетическая конструкция

<400> 56
 Met Val Ser Lys Gly Glu Glu Asp Asn Met Ala Ile Ile Lys Glu Phe
 1 5 10 15

Met Arg Phe Lys Val His Met Glu Gly Ser Val Asn Gly His Glu Phe
 20 25 30

Glu Ile Glu Gly Glu Gly Glu Gly Arg Pro Tyr Glu Gly Thr Gln Thr
 35 40 45

Ala Lys Leu Lys Val Thr Lys Gly Gly Pro Leu Pro Phe Ala Trp Asp
 50 55 60

Ile Leu Ser Pro Gln Phe Met Tyr Gly Ser Lys Ala Tyr Val Lys His
 65 70 75 80

048071

Pro Ala Asp Ile Pro Asp Tyr Leu Lys Leu Ser Phe Pro Glu Gly Phe
85 90 95

Lys Trp Glu Arg Val Met Asn Phe Glu Asp Gly Gly Val Val Thr Val
100 105 110

Thr Gln Asp Ser Ser Leu Gln Asp Gly Glu Phe Ile Tyr Lys Val Lys
115 120 125

Leu Arg Gly Thr Asn Phe Pro Ser Asp Gly Pro Val Met Gln Lys Lys
130 135 140

Thr Met Gly Trp Glu Ala Ser Ser Glu Arg Met Tyr Pro Glu Asp Gly
145 150 155 160

Ala Leu Lys Gly Glu Ile Lys Gln Arg Leu Lys Leu Lys Asp Gly Gly
165 170 175

His Tyr Asp Ala Glu Val Lys Thr Thr Tyr Lys Ala Lys Lys Pro Val
180 185 190

Gln Leu Pro Gly Ala Tyr Asn Val Asn Ile Lys Leu Asp Ile Thr Ser
195 200 205

His Asn Glu Asp Tyr Thr Ile Val Glu Gln Tyr Glu Arg Ala Glu Gly
210 215 220

Arg His Ser Thr Gly Gly Met Asp Glu Leu Tyr Lys
225 230 235

<210> 57

<211> 78

<212> белок

<213> Homo sapiens

<400> 57

Gln Asp Val Asp Glu Cys Ser Leu Gly Ala Asn Pro Cys Glu His Ala
1 5 10 15

Gly Lys Cys Ile Asn Thr Leu Gly Ser Phe Glu Cys Gln Cys Leu Gln
20 25 30

Gly Tyr Thr Gly Pro Arg Cys Glu Ile Asp Val Asn Glu Cys Val Ser
35 40 45

Asn Pro Cys Gln Asn Asp Ala Thr Cys Leu Asp Gln Ile Gly Glu Phe
50 55 60

048071

Gln Cys Met Cys Met Pro Gly Tyr Glu Gly Val His Cys Glu
65 70 75

<210> 58

<211> 78

<212> белок

<213> Xenopus sp.

<400> 58

Asn Asp Val Asp Glu Cys Ser Leu Gly Ala Asn Pro Cys Glu His Gly
1 5 10 15

Gly Arg Cys Thr Asn Thr Leu Gly Ser Phe Gln Cys Asn Cys Pro Gln
20 25 30

Gly Tyr Ala Gly Pro Arg Cys Glu Ile Asp Val Asn Glu Cys Leu Ser
35 40 45

Asn Pro Cys Gln Asn Asp Ser Thr Cys Leu Asp Gln Ile Gly Glu Phe
50 55 60

Gln Cys Ile Cys Met Pro Gly Tyr Glu Gly Leu Tyr Cys Glu
65 70 75

<210> 59

<211> 78

<212> белок

<213> Danio rerio

<400> 59

Gln Asp Ile Asp Glu Cys Ser Leu Gly Ala Asn Pro Cys Glu His Gly
1 5 10 15

Gly Arg Cys Leu Asn Thr Lys Gly Ser Phe Gln Cys Lys Cys Leu Gln
20 25 30

Gly Tyr Glu Gly Pro Arg Cys Glu Met Asp Val Asn Glu Cys Lys Ser
35 40 45

Asn Pro Cys Gln Asn Asp Ala Thr Cys Leu Asp Gln Ile Gly Gly Phe
50 55 60

His Cys Ile Cys Met Pro Gly Tyr Glu Gly Val Phe Cys Gln
65 70 75

<210> 60

<211> 77

<212> белок

<213> Drosophila sp.

<400> 60

048071

Glu Asp Ile Asp Glu Cys Asp Gln Gly Ser Pro Cys Glu His Asn Gly
 1 5 10 15

Ile Cys Val Asn Thr Pro Gly Ser Tyr Arg Cys Asn Cys Ser Gln Gly
 20 25 30

Phe Thr Gly Pro Arg Cys Glu Thr Asn Ile Asn Glu Cys Glu Ser His
 35 40 45

Pro Cys Gln Asn Glu Gly Ser Cys Leu Asp Asp Pro Gly Thr Phe Arg
 50 55 60

Cys Val Cys Met Pro Gly Phe Thr Gly Thr Gln Cys Glu
 65 70 75

<210> 61

<211> 5

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая
 распознавательная последовательность каспазы-1

<220>

<221> модифицированный остаток

<222> (1)..(1)

<223> Phe, Leu, Trp или Tyr

<220>

<221> модифицированный остаток

<222> (2)..(2)

<223> Любая аминокислота

<220>

<221> модифицированный остаток

<222> (3)..(3)

<223> Ala, His или Thr

<220>

<221> модифицированный остаток

<222> (5)..(5)

<223> Любая аминокислота за исключением Asp, Glu, Lys, Pro, Gln или Arg

<400> 61

Xaa Xaa Xaa Asp Xaa

1 5

<210> 62

<211> 5

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая
 распознавательная последовательность каспазы-10

<220>

<221> модифицированный остаток
<222> (5)..(5)
<223> Любая аминокислота

<400> 62
Ile Glu Ala Asp Xaa
1 5

<210> 63
<211> 5
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность каспазы-2
<220>
<221> модифицированный остаток
<222> (5)..(5)
<223> Любая аминокислота за исключением Asp, Glu, Lys, Pro, Gln или Arg

<400> 63
Asp Val Ala Asp Xaa
1 5

<210> 64
<211> 5
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность каспазы-2

<220>
<221> модифицированный остаток
<222> (5)..(5)
<223> Любая аминокислота за исключением Asp, Glu, Lys, Pro, Gln или Arg

<400> 64
Asp Glu His Asp Xaa
1 5

<210> 65
<211> 5
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность каспазы-3

<220>
<221> модифицированный остаток
<222> (5)..(5)
<223> Любая аминокислота за исключением Asp, Glu, Lys, Pro, Gln или Arg

<400> 65
Asp Met Gln Asp Xaa
1 5

<210> 66
 <211> 5
 <212> белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетическая
 распознавательная последовательность каспазы-3

<220>
 <221> модифицированный остаток
 <222> (5)..(5)
 <223> Любая аминокислота за исключением Asp, Glu, Lys, Pro, Gln или Arg

<400> 66
 Asp Glu Val Asp Xaa
 1 5

<210> 67
 <211> 5
 <212> белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетическая
 распознавательная последовательность каспазы-4

<220>
 <221> модифицированный остаток
 <222> (5)..(5)
 <223> Любая аминокислота за исключением Asp, Glu, Lys, Pro, Gln или Arg

<400> 67
 Leu Glu Val Asp Xaa
 1 5

<210> 68
 <211> 5
 <212> белок
 <213> Искусственная последовательность

<220>
 <223> Описание искусственной последовательности: Синтетическая
 распознавательная последовательность каспазы-4

<220>
 <221> модифицированный остаток
 <222> (1)..(1)
 <223> Leu или Trp

<220>
 <221> модифицированный остаток
 <222> (5)..(5)
 <223> Любая аминокислота за исключением Asp, Glu, Lys, Pro, Gln или Arg

<400> 68
 Xaa Glu His Asp Xaa
 1 5

<210> 69
<211> 5
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность каспазы-5

<220>
<221> модифицированный остаток
<222> (1)..(1)
<223> Leu или Trp

<220>
<221> модифицированный остаток
<222> (5)..(5)
<223> Любая аминокислота

<400> 69
Xaa Glu His Asp Xaa
1 5

<210> 70
<211> 5
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность каспазы-6

<220>
<221> модифицированный остаток
<222> (3)..(3)
<223> His или Ile

<220>
<221> модифицированный остаток
<222> (5)..(5)
<223> Любая аминокислота за исключением Asp, Glu, Lys, Pro, Gln или Arg

<400> 70
Val Glu Xaa Asp Xaa
1 5

<210> 71
<211> 5
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность каспазы-7

<220>
<221> модифицированный остаток
<222> (5)..(5)
<223> Любая аминокислота за исключением Asp, Glu, Lys, Pro, Gln или Arg

<400> 71

Asp Glu Val Asp Xaa

1 5

<210> 72

<211> 5

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность каспазы-8

<220>

<221> модифицированный остаток

<222> (1)..(1)

<223> Ile или Leu

<220>

<221> модифицированный остаток

<222> (5)..(5)

<223> Любая аминокислота за исключением Asp, Glu, Lys, Pro, Gln или Arg

<400> 72

Xaa Glu Thr Asp Xaa

1 5

<210> 73

<211> 5

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность каспазы-9

<220>

<221> модифицированный остаток

<222> (5)..(5)

<223> Любая аминокислота

<400> 73

Leu Glu His Asp Xaa

1 5

<210> 74

<211> 6

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность энтерокиназы

<220>

<221> модифицированный остаток

<222> (1)..(4)

<223> Asp или Glu

<220>

<221> модифицированный остаток
<222> (6)..(6)
<223> Любая аминокислота

<400> 74
Xaa Xaa Xaa Xaa Lys Xaa
1 5

<210> 75
<211> 5
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность фактора Ха

<220>
<221> модифицированный остаток
<222> (1)..(1)
<223> Ala, Phe, Gly, Ile, Leu, Thr, Val или Met

<220>
<221> модифицированный остаток
<222> (2)..(2)
<223> Asp или Glu

<220>
<221> модифицированный остаток
<222> (5)..(5)
<223> Любая аминокислота

<400> 75
Xaa Xaa Gly Arg Xaa
1 5

<210> 76
<211> 5
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность транзима В

<220>
<221> модифицированный остаток
<222> (5)..(5)
<223> Любая аминокислота

<400> 76
Ile Glu Pro Asp Xaa
1 5

<210> 77
<211> 8
<212> белок
<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность протеазы HRV3C

<400> 77

Leu Glu Val Leu Phe Gln Gly Pro

1 5

<210> 78

<211> 5

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность пепсина А

<220>

<221> модифицированный остаток

<222> (1)..(1)

<223> Любая аминокислота за исключением His, Lys или Arg

<220>

<221> модифицированный остаток

<222> (2)..(2)

<223> Любая аминокислота за исключением Pro

<220>

<221> модифицированный остаток

<222> (3)..(3)

<223> Любая аминокислота за исключением Arg

<220>

<221> модифицированный остаток

<222> (4)..(4)

<223> Phe, Leu, Trp или Tyr

<220>

<221> модифицированный остаток

<222> (5)..(5)

<223> Любая аминокислота за исключением Pro

<400> 78

Xaa Xaa Xaa Xaa Xaa

1 5

<210> 79

<211> 5

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность пепсина А

<220>

<221> модифицированный остаток

<222> (1)..(1)

<223> Любая аминокислота за исключением His, Lys или Arg

<220>

<221> модифицированный остаток

<222> (2)..(2)

<223> Любая аминокислота за исключением Pro

<220>

<221> модифицированный остаток

<222> (3) .. (3)

<223> Phe, Leu, Trp или Tyr

<220>

<221> модифицированный остаток

<222> (4) .. (4)

<223> Любая аминокислота

<220>

<221> модифицированный остаток

<222> (5) .. (5)

<223> Любая аминокислота за исключением Pro

<400> 79

Хаа Хаа Хаа Хаа Хаа

1

5

<210> 80

<211> 5

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая распознавательная последовательность пепсина А (низкая специфичность)

<220>

<221> модифицированный остаток

<222> (1) .. (1)

<223> Любая аминокислота за исключением His, Lys или Arg

<220>

<221> модифицированный остаток

<222> (2) .. (2)

<223> Любая аминокислота за исключением Pro

<220>

<221> модифицированный остаток

<222> (3) .. (3)

<223> Любая аминокислота за исключением Arg

<220>

<221> модифицированный остаток

<222> (4) .. (4)

<223> Phe или Leu

<220>

<221> модифицированный остаток

<222> (5) .. (5)

<223> Любая аминокислота за исключением Pro

<400> 80

Хаа Хаа Хаа Хаа Хаа

1

5

<210> 81

<211> 5
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая распознавательная последовательность пепсина А (низкая специфичность)

<220>
<221> модифицированный остаток
<222> (1)..(1)
<223> Любая аминокислота за исключением His, Lys или Arg

<220>
<221> модифицированный остаток
<222> (2)..(2)
<223> Любая аминокислота за исключением Pro

<220>
<221> модифицированный остаток
<222> (3)..(3)
<223> Phe или Leu

<220>
<221> модифицированный остаток
<222> (4)..(4)
<223> Любая аминокислота

<220>
<221> модифицированный остаток
<222> (5)..(5)
<223> Любая аминокислота за исключением Pro

<400> 81
Хаа Хаа Хаа Хаа Хаа
1 5

<210> 82
<211> 7
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая распознавательная последовательность протеазы TEV

<220>
<221> модифицированный остаток
<222> (2)..(3)
<223> Любая аминокислота

<220>
<221> модифицированный остаток
<222> (5)..(5)
<223> Любая аминокислота

<220>
<221> модифицированный остаток
<222> (7)..(7)
<223> Gly или Ser

<400> 82
Glu Xaa Xaa Tyr Xaa Gln Xaa
1 5

<210> 83
<211> 6
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность тромбина

<220>
<221> модифицированный остаток
<222> (1)..(2)
<223> Любая аминокислота

<220>
<221> модифицированный остаток
<222> (6)..(6)
<223> Любая аминокислота

<400> 83
Xaa Xaa Gly Arg Gly Xaa
1 5

<210> 84
<211> 6
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетическая
распознавательная последовательность тромбина

<220>
<221> модифицированный остаток
<222> (1)..(2)
<223> Ala, Phe, Gly, Ile, Leu, Thr, Val или Trp

<220>
<221> модифицированный остаток
<222> (5)..(6)
<223> Любая аминокислота за исключением Asp или Glu

<400> 84
Xaa Xaa Pro Arg Xaa Xaa
1 5

<210> 85
<211> 16
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический
пептид пенетратин

048071

<400> 85
Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys
1 5 10 15

<210> 86
<211> 5
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический пептид

<220>
<221> модифицированный остаток
<222> (4) .. (4)
<223> любая аминокислота
<400> 86
Met Gly Cys Xaa Cys
1 5

<210> 87
<211> 50
<212> белок
<213> Искусственная последовательность

<220>
<223> Описание искусственной последовательности: Синтетический полипептид

<220>
<221> прочий признак
<222> (1) .. (50)
<223> Эта последовательность может охватывать 3-50 остатков

<400> 87
Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg
1 5 10 15

Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg
20 25 30

Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg
35 40 45

Arg Arg
50

<210> 88
<211> 9
<212> белок
<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<400> 88

Arg Arg Arg Arg Arg Arg Arg Arg Arg
1 5

<210> 89

<211> 9

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<400> 89

Glu Glu Glu Glu Glu Glu Glu Glu Glu
1 5

<210> 90

<211> 8

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетический пептид

<400> 90

Arg Arg Arg Arg Arg Arg Arg Arg Arg
1 5

<210> 91

<211> 4

<212> белок

<213> Неизвестно

<220>

<223> Описание неизвестного:
Бета-интегриновый мотив

<400> 91

His Asp Arg Lys
1

<210> 92

<211> 8

<212> белок

<213> Искусственная последовательность

<220>

<223> Описание искусственной последовательности: Синтетическая метка 8xHis

<400> 92

His His His His His His His His
1 5

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Система для регуляции экспрессии целевого полинуклеотида в клетке, которая включает:
- (a) химерный рецепторный полипептид, который подвергается рецепторной модификации при связывании антигена, причем эта рецепторная модификация включает конформационное изменение или химическую модификацию;
 - (b) химерный адаптерный полипептид, который связывается с рецептором в ответ на модификацию рецептора;
 - (c) полипептид, модулирующий ген, (GMP), содержащий исполнительный элемент, соединенный с сайтом распознавания при расщеплении, причем при расщеплении сайта распознавания при расщеплении исполнительный элемент способен образовывать комплекс с целевым полинуклеотидом, чтобы регулировать генную экспрессию и/или активность этого целевого полинуклеотида; и
 - (d) расщепляющий элемент, который расщепляет сайт распознавания при расщеплении, когда находится вблизи сайта распознавания при расщеплении;
- где в ответ на связывание антигеном химерный рецепторный полипептид модифицируется, и химерный адаптерный полипептид привлекается к этому химерному рецепторному полипептиду, и причем:
- (i) GMP входит в состав внутриклеточной области химерного рецепторного полипептида, а расщепляющий элемент входит в состав химерного адаптерного полипептида, где связывание указанного химерного адаптерного химерного полипептида с указанным химерным рецепторным полипептидом приводит расщепляющий элемент в близость с сайтом распознавания при расщеплении, и где расщепляющий элемент расщепляет сайт распознавания при расщеплении, чтобы высвободить исполнительный элемент из GMP; или
 - (ii) GMP входит в состав химерного адаптерного полипептида, а расщепляющий элемент входит в состав внутриклеточной области химерного рецепторного полипептида, где связывание указанного химерного адаптерного химерного полипептида с указанным химерным рецепторным полипептидом приводит расщепляющий элемент в близость с сайтом распознавания при расщеплении и где расщепляющий элемент расщепляет сайт распознавания при расщеплении, чтобы высвободить исполнительный элемент из GMP; или
 - (iii) расщепляющий элемент образует комплекс со вторым адаптерным полипептидом, который связывает химерный рецепторный полипептид в ответ на модификацию рецептора, а GMP входит в состав химерного адаптерного полипептида, где связывание указанного химерного адаптерного полипептида, а также второго адаптерного полипептида с указанным химерным рецепторным полипептидом приводит расщепляющий элемент в близость с сайтом распознавания при расщеплении и где расщепляющий элемент расщепляет сайт распознавания при расщеплении, чтобы высвободить исполнительный элемент из GMP.
2. Система по п.1, где рецептор не содержит SEQ ID NO: 39.
3. Система по п.1, где целевой полинуклеотид представлен геномной ДНК.
4. Система по п.1, где целевой полинуклеотид представлен РНК.
5. Система по п.1, где модификация представляет собой фосфорилирование.
6. Система по п.1, где исполнительный элемент представлен CRISPR-ассоциированным (Cas) белком, а система дополнительно содержит нуклеиновую кислоту, которая может гибридизоваться с другой нуклеиновой кислотой, предпочтительно направляющую РНК, и которая действует, образуя комплекс с белком Cas.
7. Система по п.1, где (i) исполнительный элемент представлен белком, связывающим РНК (RBP), необязательно в комплексе с направляющей РНК, и (ii) система дополнительно содержит белок Cas, который способен образовывать комплекс с направляющей РНК.
8. Система по п.6 или 7, где белок Cas практически лишен активности расщепления ДНК.
9. Система по п.1, где (i) GMP входит в состав химерного адаптерного полипептида, (ii) при расщеплении сайта распознавания при расщеплении химерный адаптерный полипептид эффективно высвобождается из рецептора, и (iii) система содержит дополнительный химерный адаптерный полипептид, содержащий GMP, который связывается с модифицированным рецептором.
10. Система по п.1, где модификация рецептора включает модификации по нескольким сайтам модификации, причем каждый сайт модификации эффективно связывает адаптерный полипептид.
11. Система по п.1, где сайт распознавания при расщеплении включает полипептидную последовательность, а расщепляющий элемент обладает протеазной активностью.
12. Система по п.1, где сайт распознавания при расщеплении содержит дисульфидную связь, а расщепляющий элемент обладает оксидоредуктазной активностью.
13. Система по п.1, где сайт распознавания при расщеплении содержит первую часть последовательности интеина, которая реагирует со второй частью последовательности интеина с высвобождением исполнительного элемента.

14. Система по п.1, где рецептор является трансмембранным рецептором.
15. Система по п.1, где рецептор является ядерным рецептором.
16. Система по п.1, где исполнительный элемент регулирует экспрессию целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида.
17. Система по п.1, где исполнительный элемент содержит активатор для эффективного повышения экспрессии целевого полинуклеотида.
18. Система по п.1, где исполнительный элемент соединяется по меньшей мере с одним сигналом ядерной локализации (NLS).
19. Система по п.1, где химерный рецепторный полипептид связан по меньшей мере с одной нацеливающей последовательностью, которая направляет транспорт рецептора в определенный район клетки.
20. Система по п.19, где нацеливающая последовательность направляет транспорт рецептора в ядро, цитоплазму, митохондрии, эндоплазматический ретикулум (ER), хлоропласт, апопласт, пероксисому или плазматическую мембрану.
21. Система по п.19, где нацеливающая последовательность включает сигнал ядерного экспорта (NES).
22. Система по п.19, где нацеливающая последовательность включает пептид, нацеливающий на плазматическую мембрану.
23. Система по п.1, где химерный адаптерный полипептид связан по меньшей мере с одной нацеливающей последовательностью, которая направляет транспорт адаптера в определенный район клетки.
24. Система по п.23, где нацеливающая последовательность направляет транспорт химерного адаптерного полипептида в ядро, цитоплазму, митохондрии, эндоплазматический ретикулум (ER), хлоропласт, апопласт, пероксисому или плазматическую мембрану.
25. Система по п.23, где нацеливающая последовательность включает сигнал ядерного экспорта (NES).
26. Система по п.23, где нацеливающая последовательность включает пептид, нацеливающий на плазматическую мембрану.
27. Система по п.1, где рецептор связан с доменом фолдинга полипептидов.
28. Система по п.1, где химерный адаптерный полипептид связан с доменом фолдинга полипептидов.
29. Способ регулирования экспрессии целевого полинуклеотида в клетках, который включает:
- контактирование химерного рецепторного полипептида и антигена, причем (i) этот рецептор подвергается рецепторной модификации при связывании антигена, и (ii) рецепторная модификация включает конформационное изменение или химическую модификацию;
 - связывание химерного адаптерного полипептида с химерным рецепторным полипептидом в ответ на рецепторную модификацию с образованием комплекса между геномодулирующим полипептидом (GMP) и расщепляющим элементом, причем GMP содержит исполнительный элемент, соединенный с сайтом распознавания при расщеплении; и
 - расщепление сайта распознавания при расщеплении расщепляющим элементом, причем при расщеплении сайта распознавания при расщеплении исполнительный элемент активируется, чтобы образовать комплекс с целевым полинуклеотидом, тем самым регулируя экспрессию целевого полинуклеотида в клетке;
- где целевой полинуклеотид связан с генетическим заболеванием или медицинским состоянием; и где:
- GMP входит в состав внутриклеточной области химерного рецепторного полипептида, а расщепляющий элемент входит в состав химерного адаптерного полипептида; (ii) расщепляющий элемент входит в состав химерного адаптерного полипептида, а GMP входит в состав внутриклеточной области химерного рецептора; или (iii) расщепляющий элемент образует комплекс со вторым адаптерным полипептидом, который связывается с рецептором в ответ на модификацию рецептора, а GMP входит в состав химерного адаптерного полипептида.
30. Способ по п.29, где рецептор не содержит SEQ ID NO: 39.
31. Способ по п.29, где целевой полинуклеотид представлен геномной ДНК.
32. Способ по п.29, где целевой полинуклеотид представлен РНК.
33. Способ по п.29, где модификация представляет собой фосфорилирование.
34. Способ по п.29, где исполнительный элемент представлен белком Cas, который образует комплекс с направляющей РНК.
35. Способ по п.29, где исполнительный элемент представлен белком, связывающим РНК, (RBP) в комплексе с направляющей РНК, которая образует комплекс с белком Cas.
36. Способ по п.34 или 35, где белок Cas практически лишен активности расщепления ДНК.
37. Способ по п.29, где (i) GMP входит в состав химерного адаптерного полипептида, (ii) химерный адаптерный полипептид высвобождается из рецептора после расщепления сайта распознавания при расщеплении, и (iii) дополнительный химерный адаптерный полипептид, содержащий GMP, связывается с

модифицированным рецептором.

38. Способ по п.29, где рецепторная модификация включает модификацию по нескольким сайтам модификации, причем каждый сайт модификации эффективно связывает адаптерный полипептид.

39. Способ по п.29, где сайт распознавания при расщеплении включает полипептидную последовательность, а расщепляющий элемент обладает протеазной активностью.

40. Способ по п.29, где сайт распознавания при расщеплении содержит дисульфидную связь, а расщепляющий элемент обладает оксидоредуктазной активностью.

41. Способ по п.29, где сайт распознавания при расщеплении содержит первую часть последовательности интеина, которая реагирует со второй частью последовательности интеина с высвобождением исполнительного элемента.

42. Способ по п.29, где рецептор является трансмембранным рецептором.

43. Способ по п.29, где рецептор является ядерным рецептором.

44. Способ по п.29, где исполнительный элемент регулирует экспрессию целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида.

45. Способ по п.29, где исполнительный элемент содержит активатор для эффективного повышения экспрессии целевого полинуклеотида.

46. Химерный внутриклеточный рецептор, который включает:

(a) взаимодействующий с антигеном домен, который специфически связывает антиген; и

(b) исполнительный элемент, соединенный со взаимодействующим с антигеном доменом;

причем:

(i) химерный внутриклеточный рецептор подвергается модификации в ответ на связывание антигена; (ii) весь рецепторный полипептид транслоцируется в ядро клетки в ответ на указанную модификацию; и (iii) исполнительный элемент образует комплекс с целевым полинуклеотидом в ядре этой клетки.

47. Химерный внутриклеточный рецептор по п.46, где исполнительный элемент представлен белком Cas, который образует комплекс с направляющей РНК.

48. Химерный внутриклеточный рецептор по п.47, где белок Cas практически лишен активности расщепления ДНК.

49. Химерный внутриклеточный рецептор по п.46, где исполнительный элемент регулирует экспрессию целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида.

50. Химерный внутриклеточный рецептор по п.46, где исполнительный элемент содержит активатор для эффективного повышения экспрессии целевого полинуклеотида.

51. Химерный внутриклеточный рецептор по п.46, где антиген представлен гормоном.

52. Химерный внутриклеточный рецептор по п.46, где исполнительный элемент соединяется по меньшей мере с одним сигналом ядерной локализации (NLS).

53. Химерный внутриклеточный рецептор по п.46, где рецептор связан по меньшей мере с одной нацеливающей последовательностью, которая направляет транспорт рецептора в определенный район клетки.

54. Химерный внутриклеточный рецептор по п.53, где нацеливающая последовательность направляет транспорт рецептора в ядро, цитоплазму, митохондрию, эндоплазматический ретикулум (ER), хлоропласт, апопласт или пероксисому.

55. Химерный внутриклеточный рецептор по п.53, где нацеливающая последовательность включает сигнал ядерного экспорта (NES).

56. Химерный внутриклеточный рецептор по п.53, где нацеливающая последовательность включает пептид, нацеливающий на плазматическую мембрану.

57. Химерный внутриклеточный рецептор по п.46, где рецептор связан с доменом фолдинга полипептидов.

58. Способ регулирования экспрессии целевого полинуклеотида в клетке, содержащей ядро, который включает:

(a) контактирование химерного внутриклеточного рецептора и антигена, причем

(i) рецептор содержит взаимодействующий с антигеном домен и исполнительный элемент, и (ii) рецептор подвергается модификации при экспозиции с антигеном;

(b) транслокацию модифицированного рецептора в ядро; и

(c) образование комплекса между исполнительным элементом и целевым полинуклеотидом,

где целевой полинуклеотид связан с генетическим заболеванием или медицинским состоянием.

59. Способ по п.58, где исполнительный элемент представлен белком Cas, который образует комплекс с направляющей РНК.

60. Способ по п.58, где белок Cas практически лишен активности расщепления ДНК.

61. Способ по п.58, где исполнительный элемент регулирует экспрессию целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов

для эффективного подавления или усиления экспрессии целевого полинуклеотида.

62. Способ по п.58, где исполнительный элемент содержит активатор для эффективного повышения экспрессии целевого полинуклеотида.

63. Способ по п.58, где антиген представлен гормоном.

64. Химерный рецепторный полипептид, который включает:

(a) взаимодействующий с антигеном домен; и

(b) полипептид, модулирующий ген, (GMP), содержащий исполнительный элемент, соединенный с сайтом распознавания при расщеплении;

причем (i) в ответ на связывание антигеном взаимодействующего с антигеном домена рецепторный полипептид подвергается рецепторной модификации, чтобы привлечь, по меньшей мере, расщепляющий элемент к этому рецепторному полипептиду; (ii) сайт распознавания при расщеплении расщепляется указанным расщепляющим элементом в ответ на указанную рецепторную модификацию; (iii) исполнительный элемент образует комплекс с целевым полинуклеотидом после отщепления от рецепторного полипептида по сайту распознавания при расщеплении; и (iv) при условии, что указанный рецепторный полипептид не содержит SEQ ID NO: 39.

65. Химерный рецепторный полипептид по п.64, где сайт распознавания при расщеплении фланкирован взаимодействующим с антигеном доменом и исполнительным элементом.

66. Химерный рецепторный полипептид по п.64, где взаимодействующий с антигеном домен входит в состав внеклеточной области химерного рецепторного полипептида, а GMP входит в состав внутриклеточной области химерного рецепторного полипептида.

67. Химерный рецепторный полипептид по п.64, где исполнительный элемент транслоцируется в ядро клетки после расщепления последовательности распознавания при расщеплении.

68. Химерный рецепторный полипептид по п.64, где химерный рецепторный полипептид представлен ядерным рецептором, который транслоцируется в ядро клетки в ответ на связывание антигена.

69. Химерный рецепторный полипептид по п.64, где исполнительный элемент представлен белком Cas, который образует комплекс с направляющей РНК.

70. Химерный рецепторный полипептид по п.64, где исполнительный элемент представлен белком, связывающим РНК, (RBP), необязательно в комплексе с направляющей РНК, которая способна образовывать комплекс с белком Cas.

71. Химерный рецепторный полипептид по п.69 или 70, где белок Cas практически лишен активности расщепления ДНК.

72. Химерный рецепторный полипептид по п.64, где сайт распознавания при расщеплении включает полипептидную последовательность, которая является распознаваемой протеазой последовательностью.

73. Химерный рецепторный полипептид по п.64, где сайт распознавания при расщеплении содержит первую часть последовательности интеина, которая реагирует со второй частью последовательности интеина с высвобождением исполнительного элемента.

74. Химерный рецепторный полипептид по п.64, где сайт распознавания при расщеплении содержит дисульфидную связь.

75. Химерный рецепторный полипептид по п.64, где рецептор является трансмембранным рецептором.

76. Химерный рецепторный полипептид по п.64, где рецептор является ядерным рецептором.

77. Химерный рецепторный полипептид по п.64, где исполнительный элемент регулирует экспрессию целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида.

78. Химерный рецепторный полипептид по п.64, где исполнительный элемент содержит активатор для эффективного повышения экспрессии целевого полинуклеотида.

79. Химерный рецепторный полипептид по п.64, где исполнительный элемент соединяется по меньшей мере с одним сигналом ядерной локализации (NLS).

80. Химерный рецепторный полипептид по п.64, где рецептор связан по меньшей мере с одной нацеливающей последовательностью, которая направляет транспорт рецептора в определенный район клетки.

81. Химерный рецепторный полипептид по п.80, где нацеливающая последовательность направляет транспорт рецептора в ядро, цитоплазму, митохондрию, эндо-плазматический ретикулум (ER), хлоропласт, апопласт, пероксисому или плазматическую мембрану.

82. Химерный рецепторный полипептид по п.80, где нацеливающая последовательность включает сигнал ядерного экспорта (NES).

83. Химерный рецепторный полипептид по п.80, где нацеливающая последовательность включает пептид, нацеливающий на плазматическую мембрану.

84. Химерный рецепторный полипептид по п.64, где рецептор связан с доменом фолдинга полипептидов.

85. Химерный адаптерный полипептид, который включает:

(a) участок связывания рецептора, который связывает рецептор, подвергшийся модификации при связывании рецептора с антигеном; и

(b) полипептид, модулирующий ген, (GMP), соединенный с участком рецепторного связывания, причем GMP содержит исполнительный элемент, соединенный с сайтом распознавания при расщеплении, где сайт распознавания при расщеплении может расщепляться расщепляющим элементом, функционально соединенным с указанным рецептором;

при этом:

(i) в ответ на связывание антигена с рецептором адаптерный полипептид привлекается к рецептору, и сайт распознавания при расщеплении расщепляется расщепляющим элементом, чтобы высвободить исполнительный элемент из GMP; и (ii) исполнительный элемент образует комплекс с целевым полинуклеотидом.

86. Химерный адаптерный полипептид по п.85, где исполнительный элемент способен транслоцироваться в ядро клетки после расщепления последовательности распознавания при расщеплении.

87. Химерный адаптерный полипептид по п.85, где исполнительный элемент представлен белком Cas, который образует комплекс с направляющей РНК.

88. Химерный адаптерный полипептид по п.85, где исполнительный элемент представлен белком, связывающим РНК, (RBP), необязательно в комплексе с направляющей РНК, которая способна образовывать комплекс с белком Cas.

89. Химерный адаптерный полипептид по п.87 или 88, где белок Cas практически лишен активности расщепления ДНК.

90. Химерный адаптерный полипептид по п.85, где сайт распознавания при расщеплении включает полипептидную последовательность, которая является распознаваемой протеазой последовательностью.

91. Химерный адаптерный полипептид по п.85, где сайт распознавания при расщеплении содержит первую часть последовательности интeина, которая реагирует со второй частью последовательности интeина с высвобождением исполнительного элемента.

92. Химерный адаптерный полипептид по п.85, где сайт распознавания при расщеплении содержит дисульфидную связь.

93. Химерный адаптерный полипептид по п.85, где исполнительный элемент регулирует экспрессию целевого полинуклеотида путем физического ограждения целевого полинуклеотида или привлечения дополнительных факторов для эффективного подавления или усиления экспрессии целевого полинуклеотида.

94. Химерный адаптерный полипептид по п.85, где исполнительный элемент содержит активатор для эффективного повышения экспрессии целевого полинуклеотида.

95. Химерный адаптерный полипептид по п.85, где исполнительный элемент соединяется по меньшей мере с одним сигналом ядерной локализации (NLS).

96. Химерный адаптерный полипептид по п.85, где адаптерный полипептид связан по меньшей мере с одной нацеливающей последовательностью, которая направляет транспорт адаптера в определенный район клетки.

97. Химерный адаптерный полипептид по п.96, где нацеливающая последовательность направляет транспорт адаптера в ядро, цитоплазму, митохондрии, эндоплазматический ретикулум (ER), хлоропласт, апопласт, пероксисому или плазматическую мембрану.

98. Химерный адаптерный полипептид по п.96, где нацеливающая последовательность содержит сигнал ядерного экспорта (NES).

99. Химерный адаптерный полипептид по п.96, где нацеливающая последовательность включает пептид, нацеливающий на плазматическую мембрану.

100. Система для регуляции экспрессии целевого полинуклеотида в клетке, которая включает:

(a) химерный рецепторный полипептид, который подвергается модификации при связывании антигена, причем модификация рецептора включает конформационное изменение или химическую модификацию;

(b) химерный адаптерный полипептид, который связывается с рецептором в ответ на модификацию рецептора;

(c) исполнительный элемент, соединенный с доменом расщепления пептидов, причем при расщеплении домена расщепления пептидов исполнительный элемент активируется и образует комплекс с целевым полинуклеотидом; и

(d) расщепляющий элемент, который расщепляет домен расщепления пептидов, когда находится вблизи домена расщепления пептидов;

причем:

(i) расщепляющий элемент входит в состав внутриклеточной части рецептора, а исполнительный элемент, связанный с доменом расщепления пептидов, входит в состав химерного адаптерного полипептида; (ii) расщепляющий элемент образует комплекс со вторым адаптерным полипептидом, который связывается с рецептором в ответ на модификацию рецептора, а исполнительный элемент, связанный с до-

меном расщепления пептидов, входит в состав химерного адаптерного полипептида; или же (iii) расщепляющий элемент входит в состав адаптерного полипептида, а исполнительный элемент, связанный с доменом расщепления пептидов, входит в состав внутриклеточной части рецептора.

101. Система для регуляции экспрессии целевого полинуклеотида в клетке, которая включает:

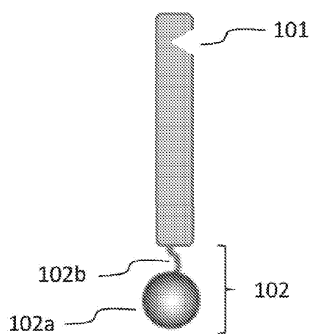
(a) химерный рецепторный полипептид, который подвергается модификации при связывании антигена, причем модификация рецептора включает конформационное изменение или химическую модификацию;

(b) химерный адаптерный полипептид, который связывается с рецептором в ответ на модификацию рецептора;

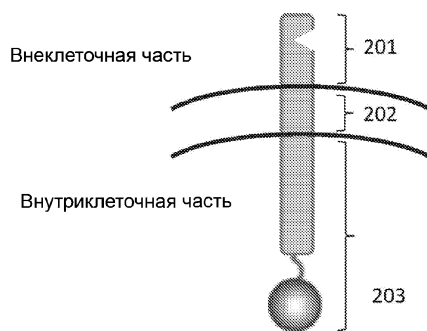
(c) исполнительный элемент, соединенный с доменом расщепления пептидов, причем при расщеплении домена расщепления пептидов исполнительный элемент активируется и образует комплекс с целевым полинуклеотидом; и

(d) рекомбинантный протеазный домен, который расщепляет домен расщепления пептидов, когда находится вблизи домена расщепления пептидов;

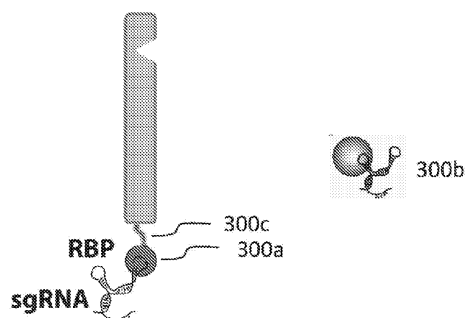
причем: (i) рекомбинантный протеазный домен входит в состав внутриклеточной части рецептора, а исполнительный элемент, связанный с доменом расщепления пептидов, входит в состав химерного адаптерного полипептида; (ii) рекомбинантный протеазный домен образует комплекс со вторым адаптерным полипептидом, который связывается с рецептором в ответ на модификацию рецептора, а исполнительный элемент, связанный с доменом расщепления пептидов, входит в состав химерного адаптерного полипептида; или же (iii) рекомбинантный протеазный домен входит в состав адаптерного полипептида, а исполнительный элемент, связанный с доменом расщепления пептидов, входит в состав внутриклеточной части рецептора.



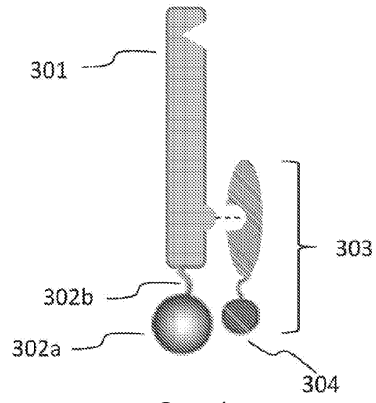
Фиг. 1



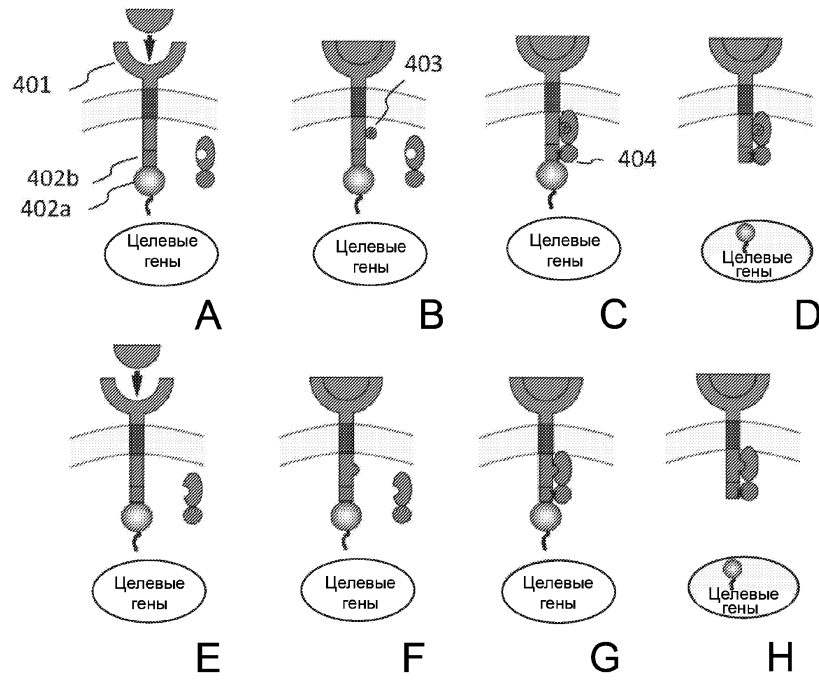
Фиг. 2



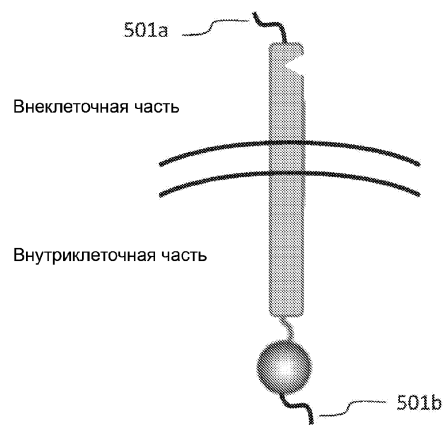
Фиг. 3



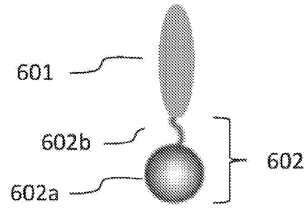
Фиг. 4



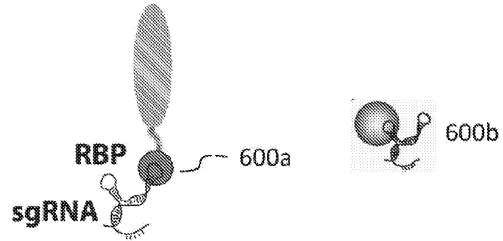
Фиг. 4А-Н



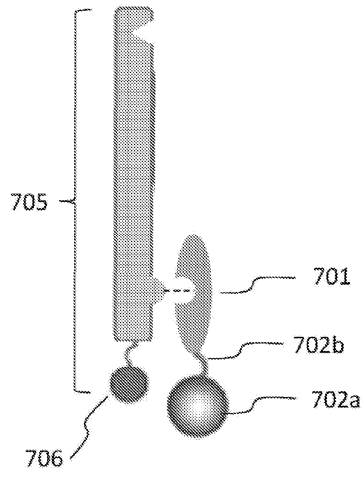
Фиг. 5



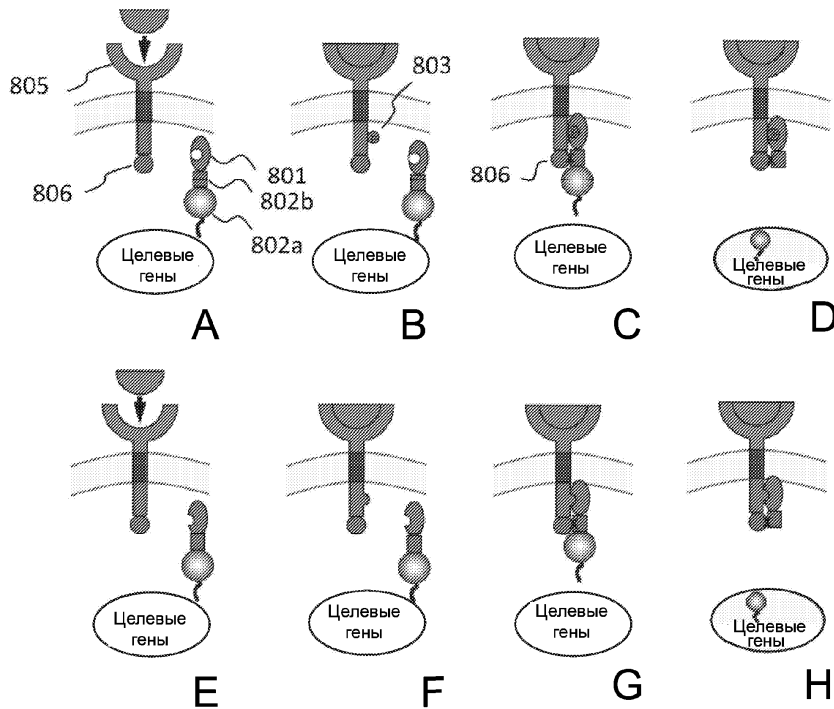
Фиг. 6А



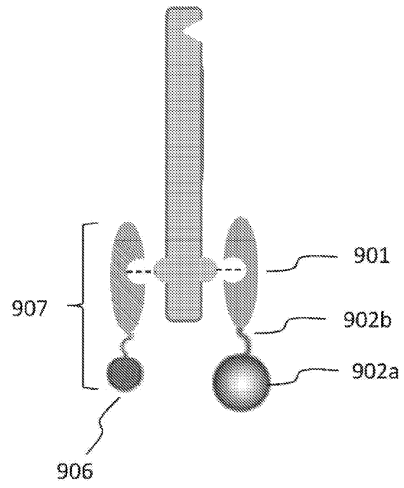
Фиг. 6В



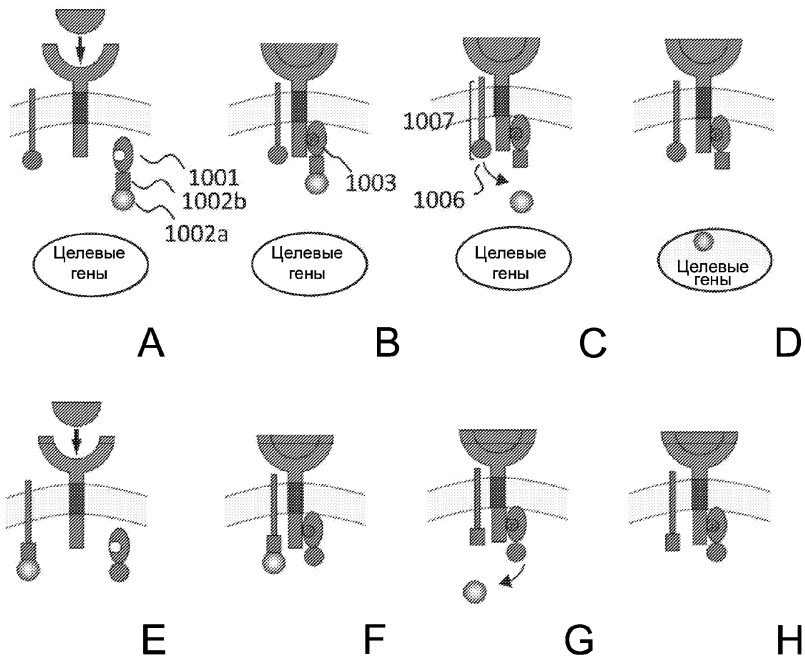
Фиг. 7



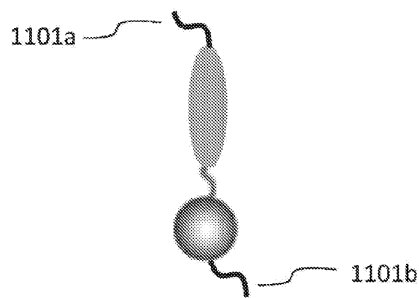
Фиг. 8А-Н



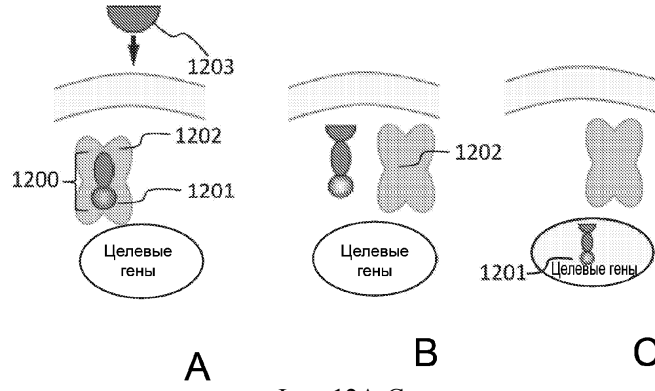
Фиг. 9



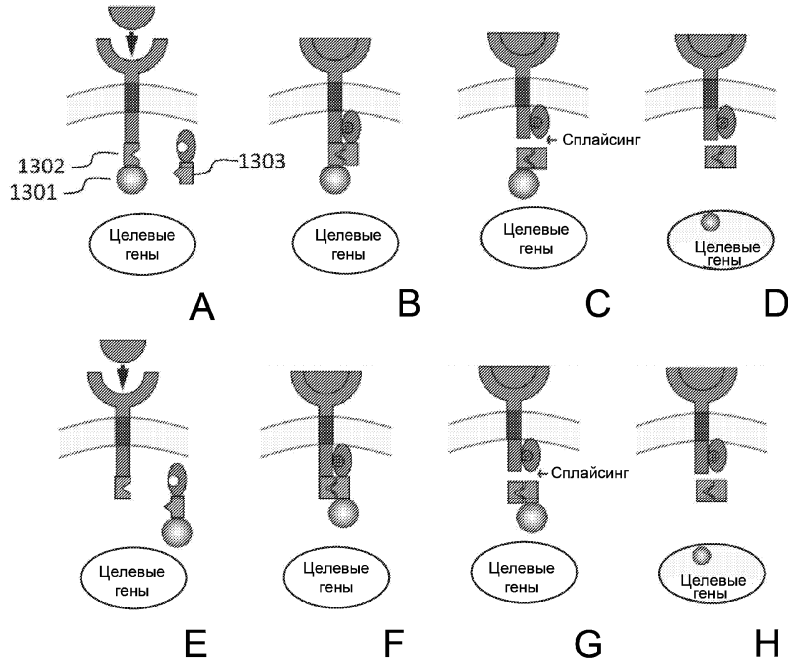
Фиг. 10А-Н



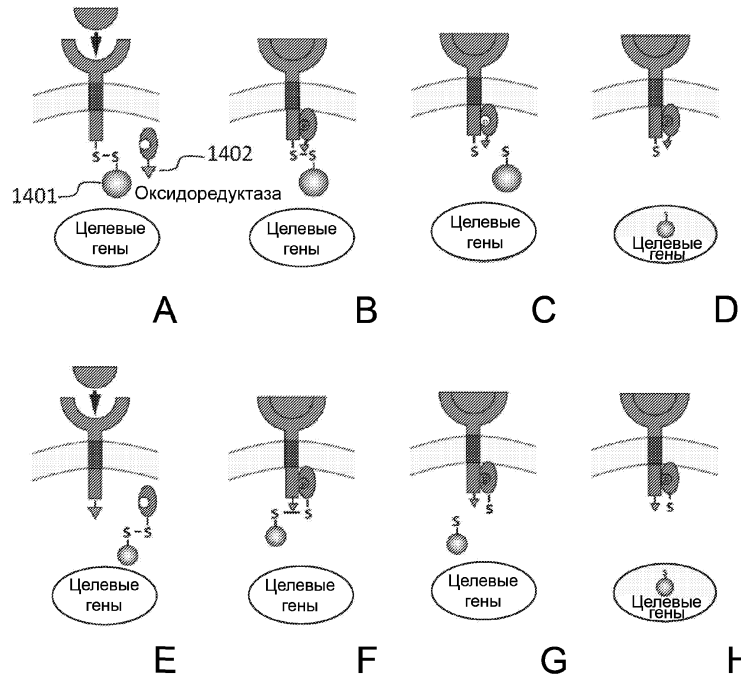
Фиг. 11



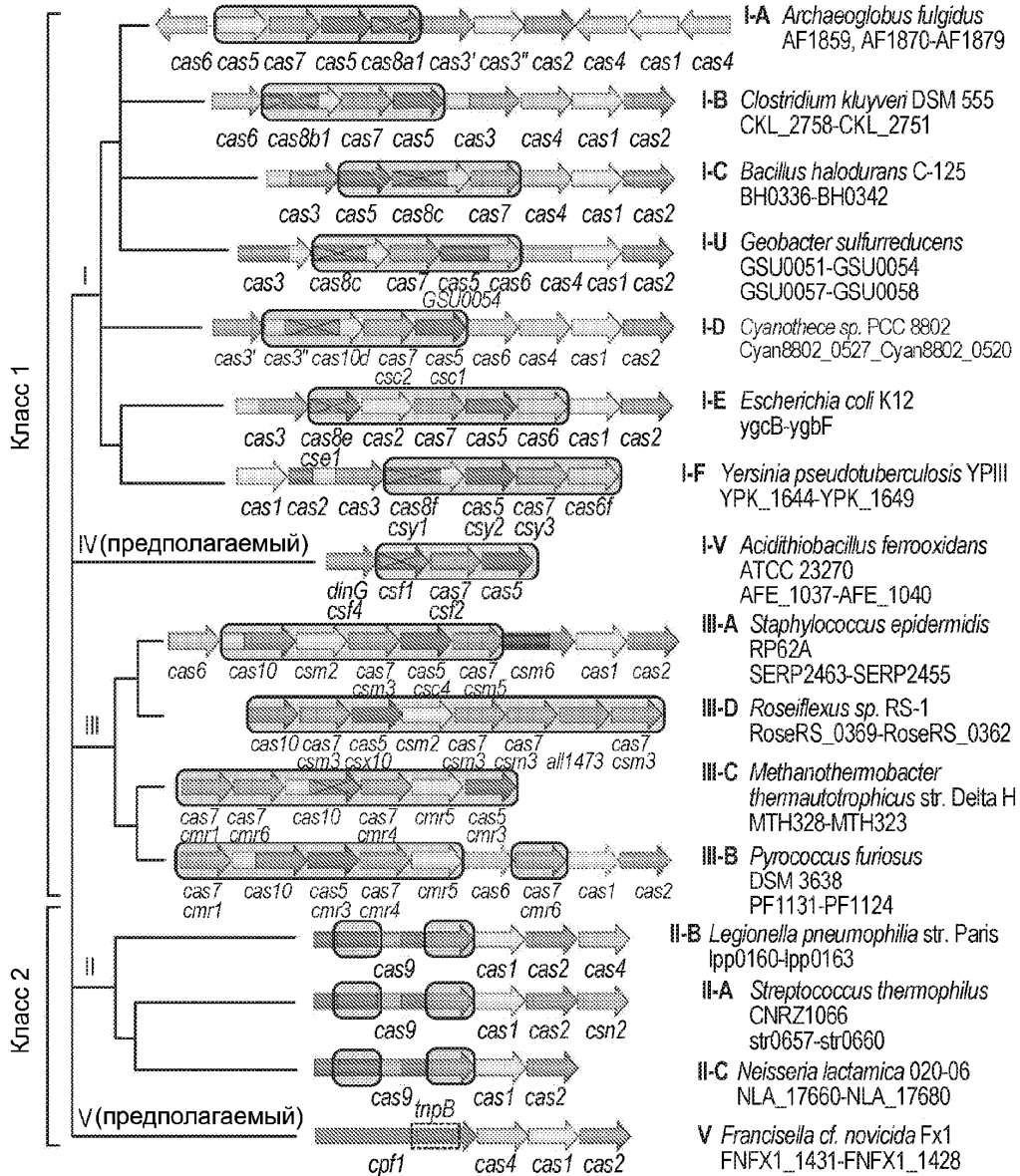
Фиг. 12А-С



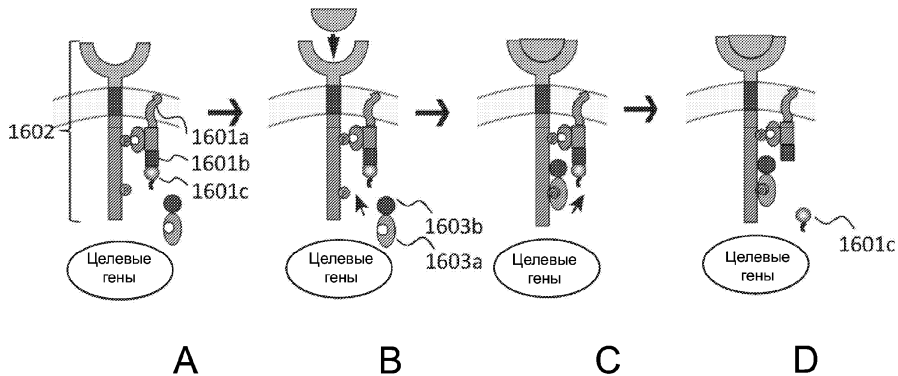
Фиг. 13А-Н



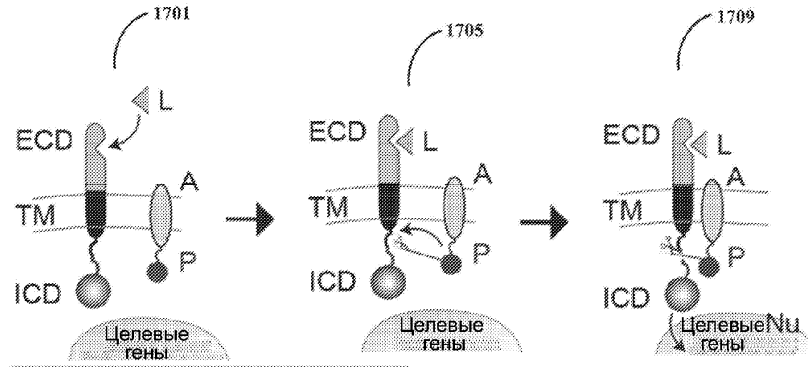
Фиг. 14А-Н



Фиг. 15

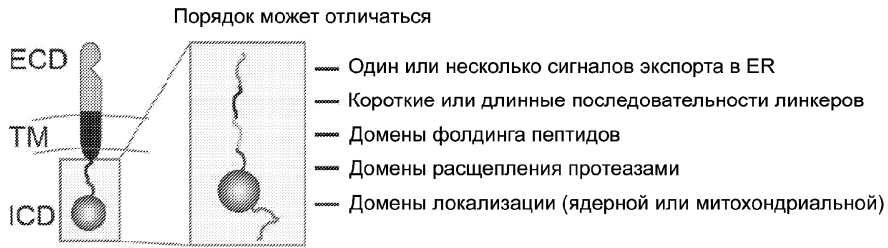


Фиг. 16A-D

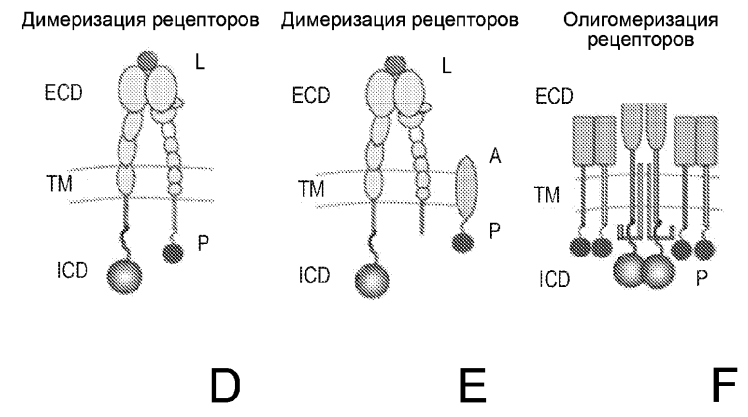
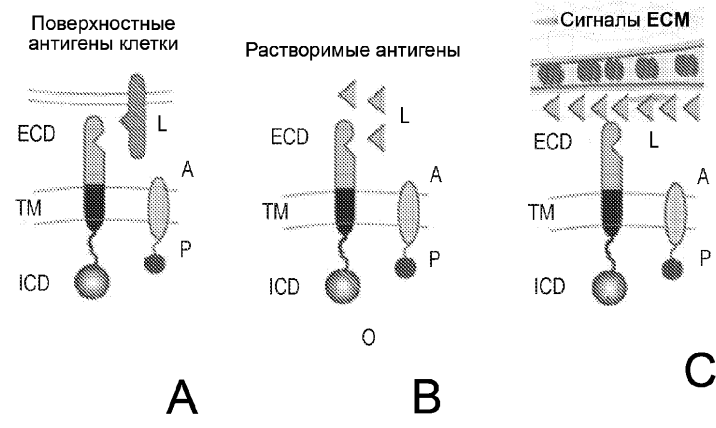


Фиг. 17

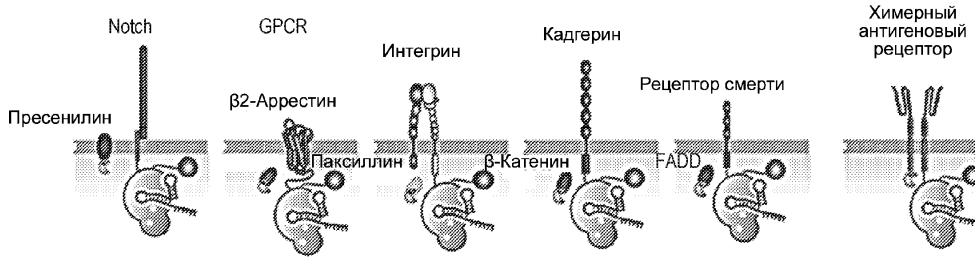
Варианты конструкции линкеров



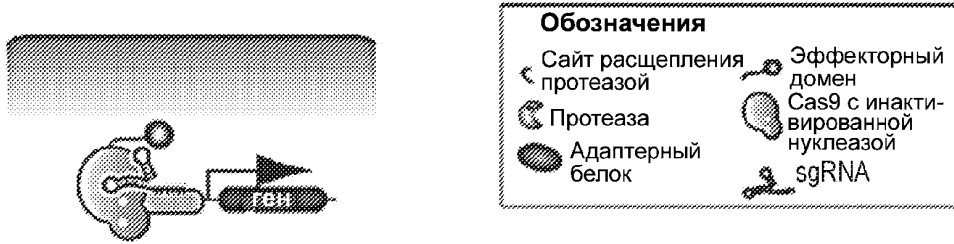
Фиг. 18



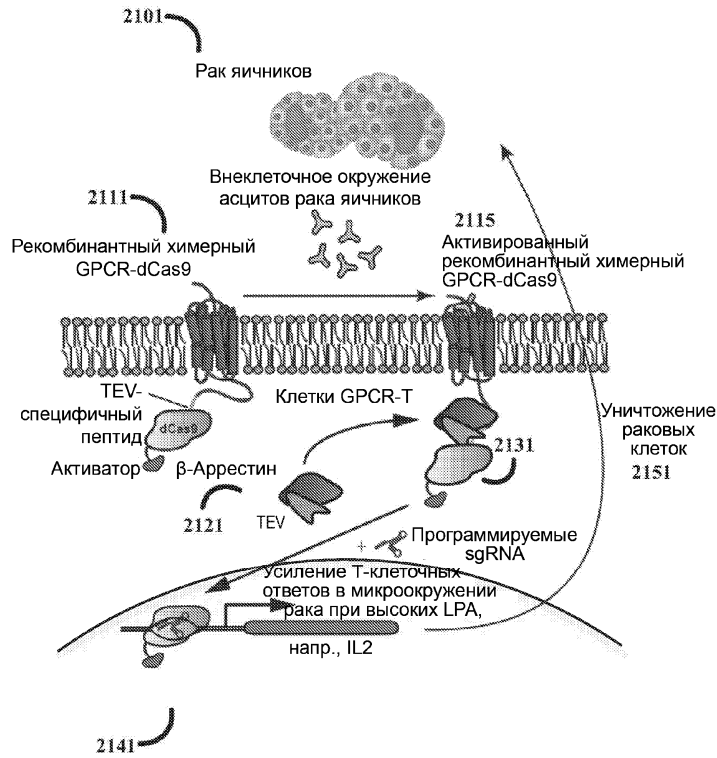
Фиг. 19A-F



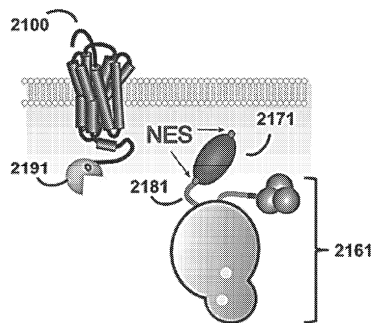
Фиг. 20



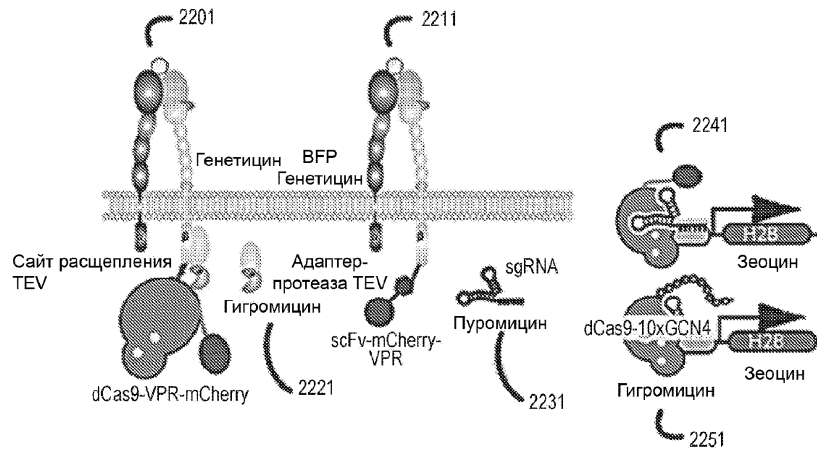
Фиг. 20В



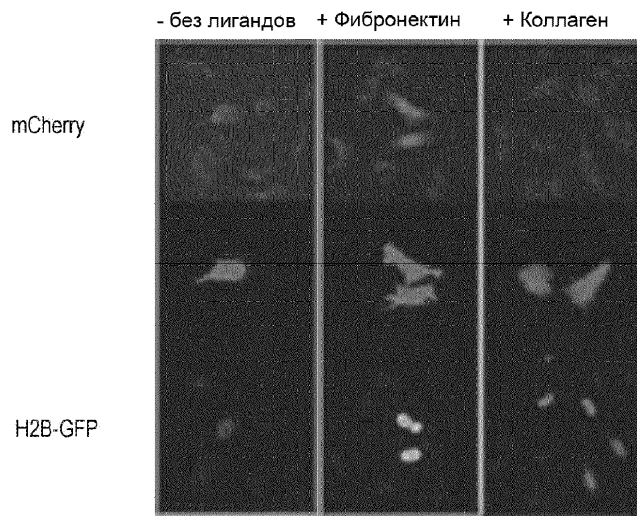
Фиг. 21А



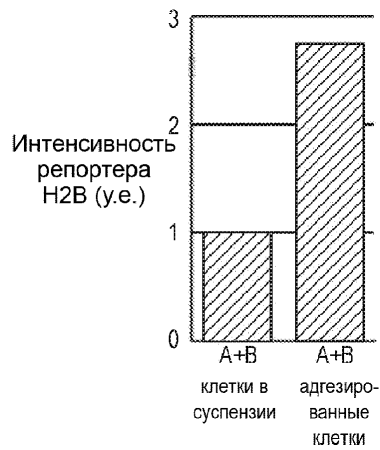
Фиг. 21В



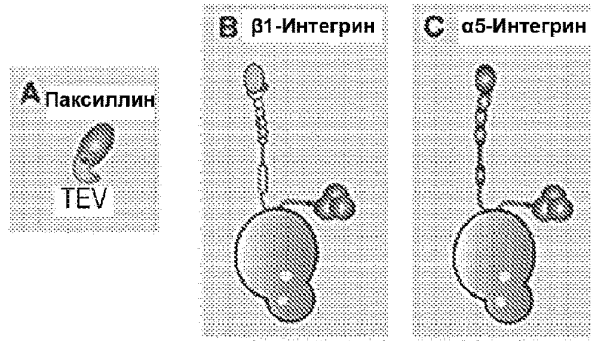
Фиг. 22А



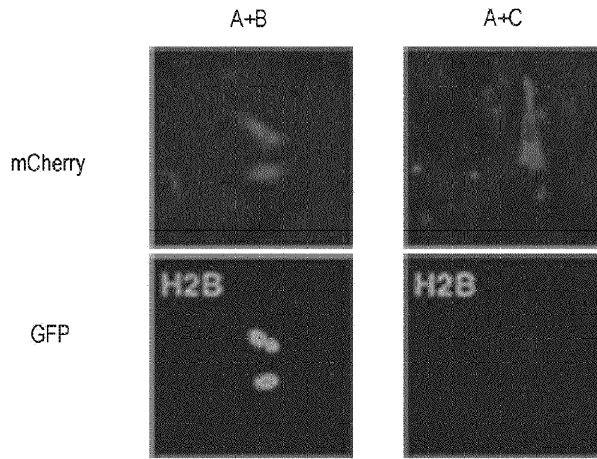
Фиг. 22В



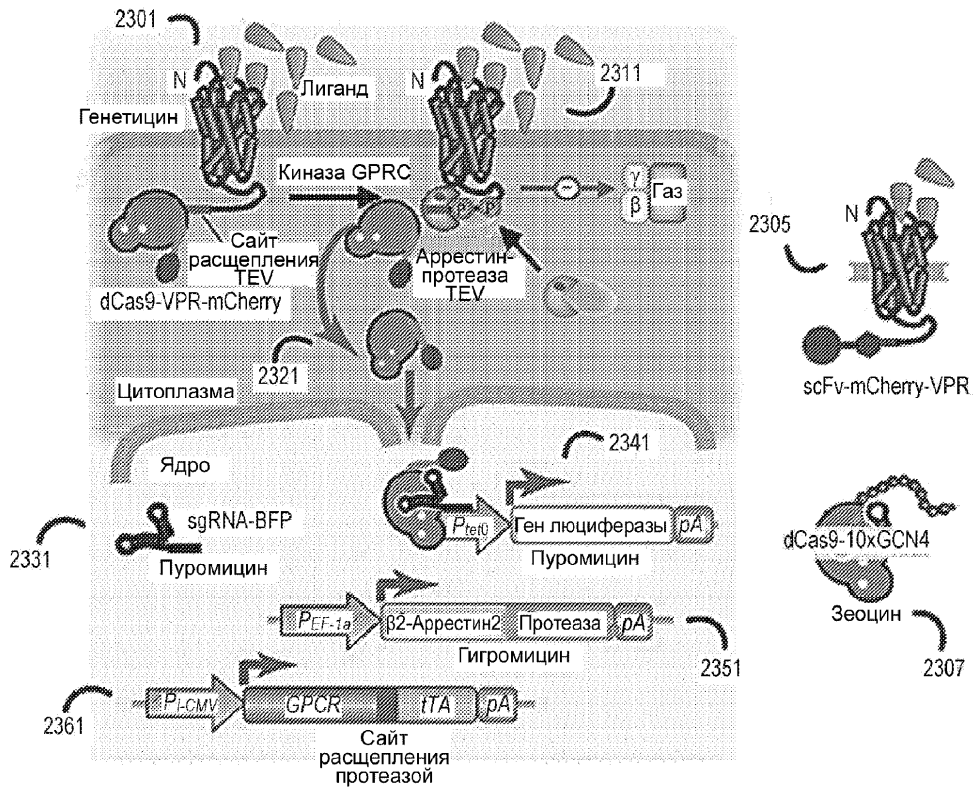
Фиг. 22С



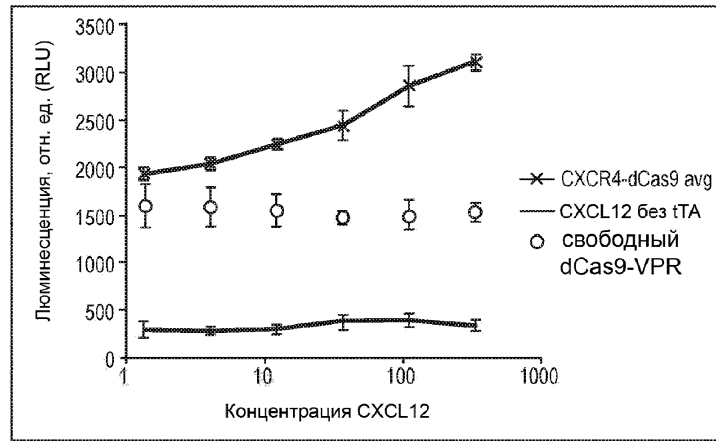
Фиг. 22D



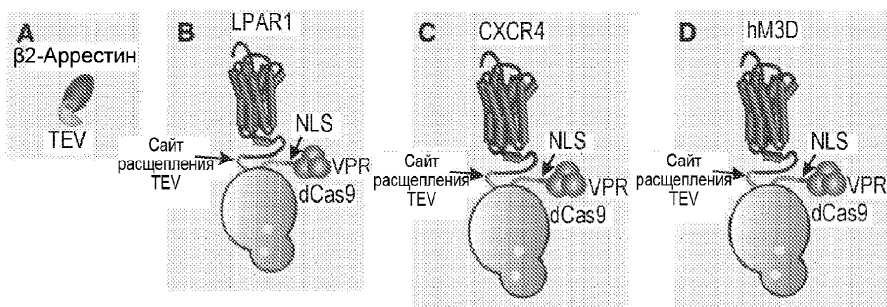
Фиг. 22E



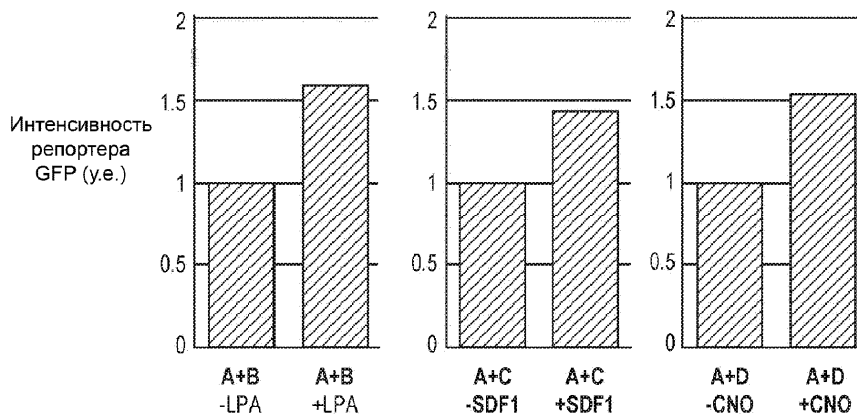
Фиг. 23A



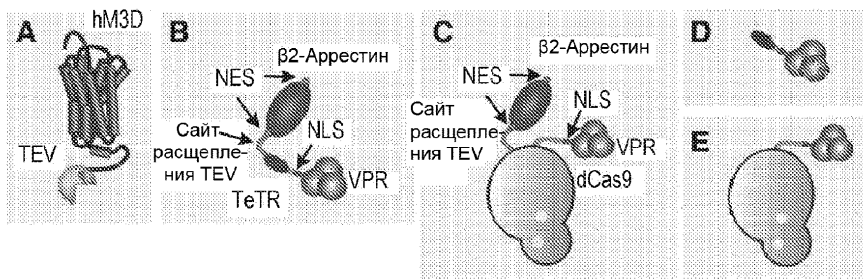
Фиг. 23В



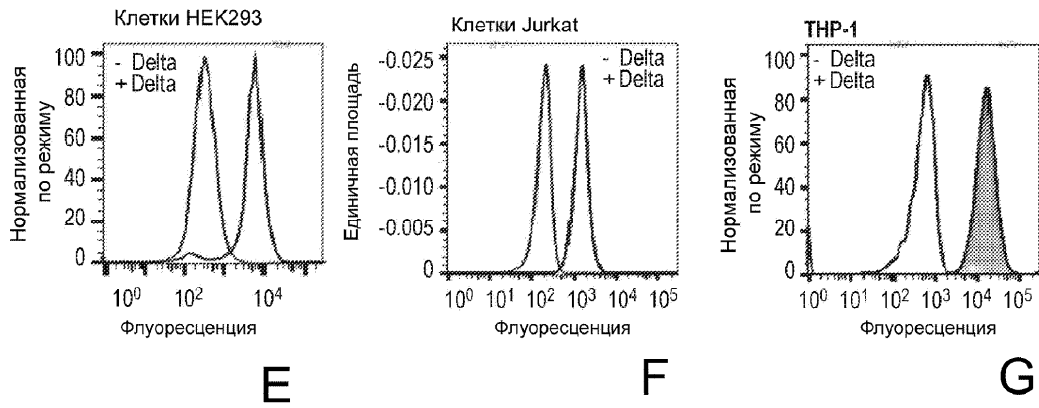
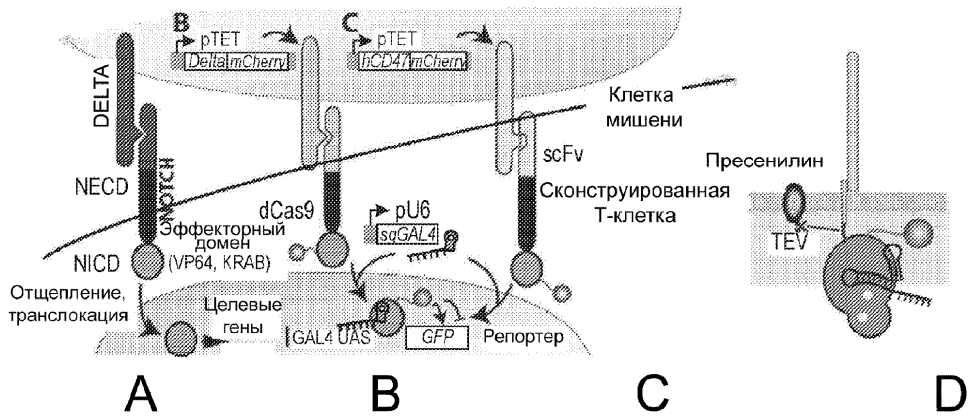
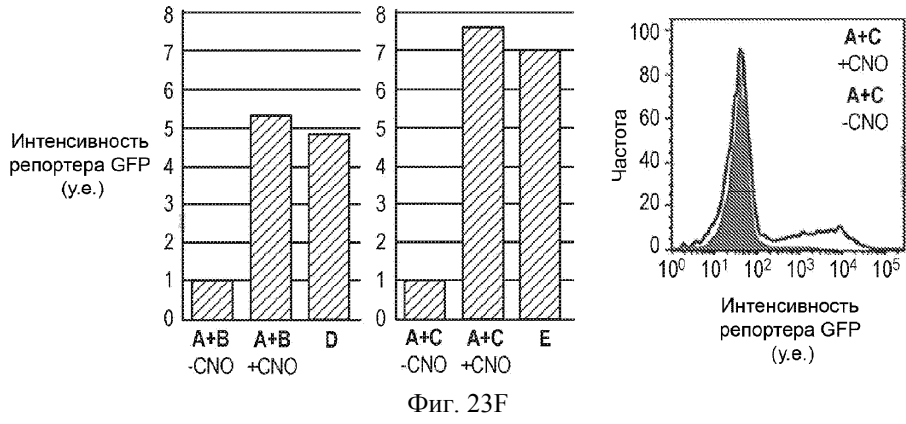
Фиг. 23С



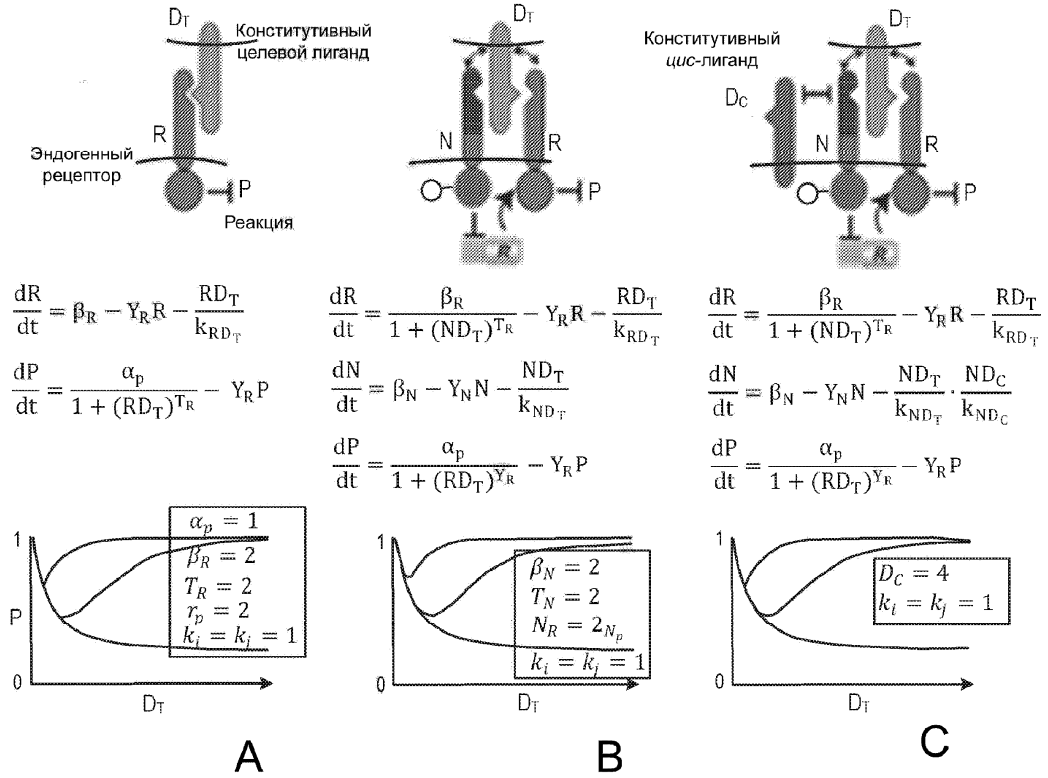
Фиг. 23D



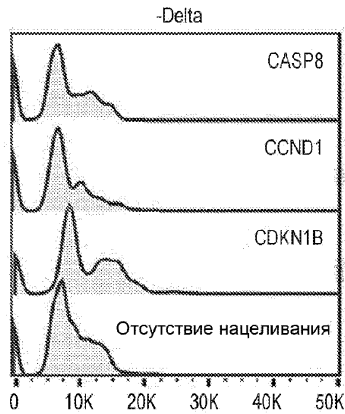
Фиг. 23Е



Фиг. 24А-Г

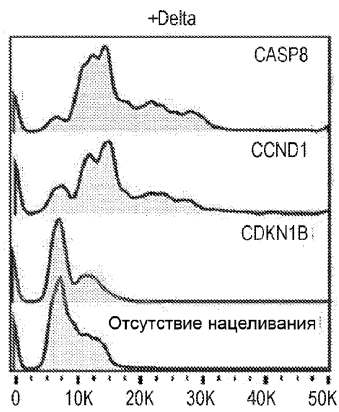


Фиг. 25



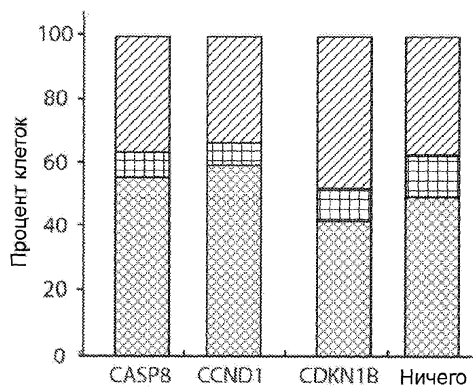
Фиолетовая FL1

Фиг. 26А

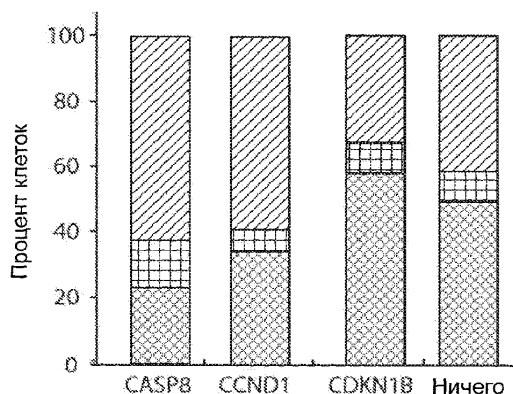


Фиолетовая FL1

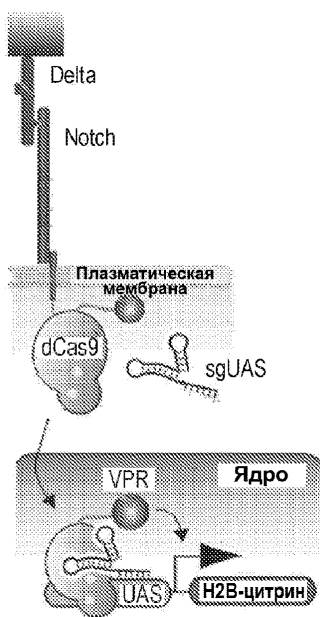
Фиг. 26В



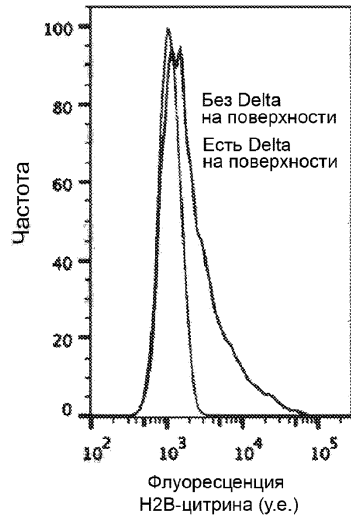
Фиг. 26С



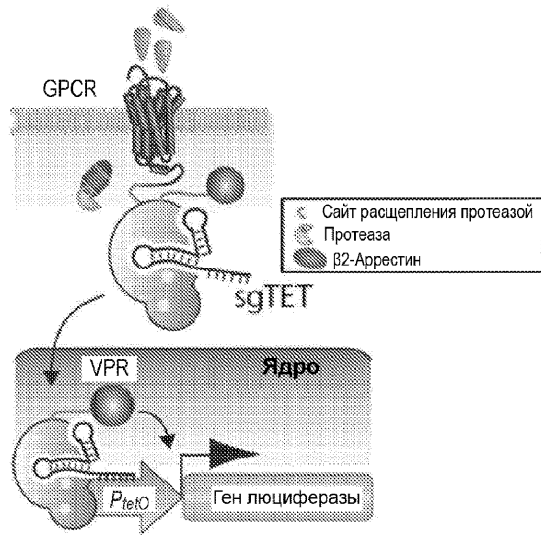
Фиг. 26D



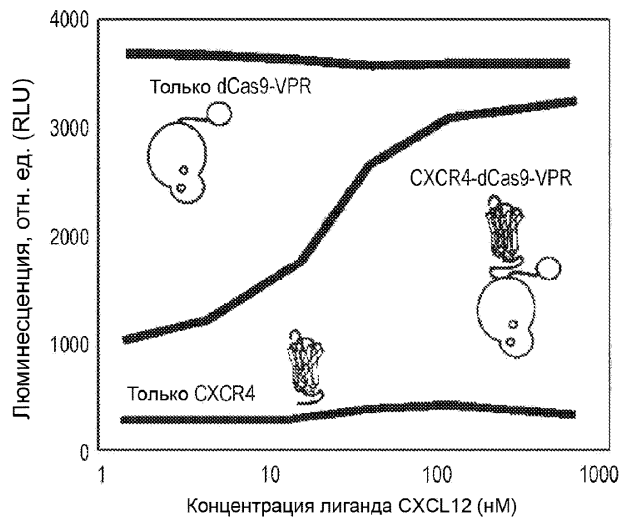
Фиг. 27А



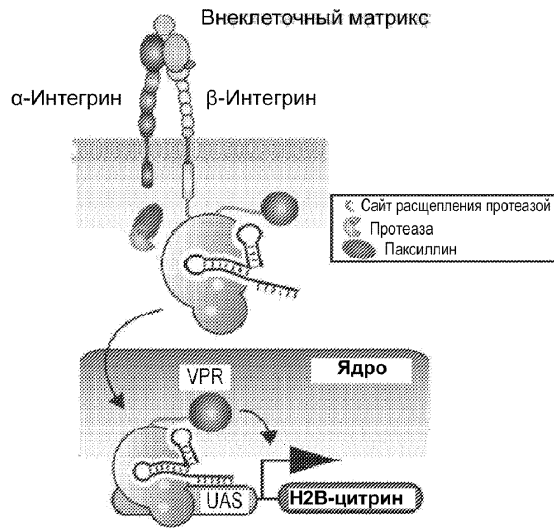
Фиг. 27В



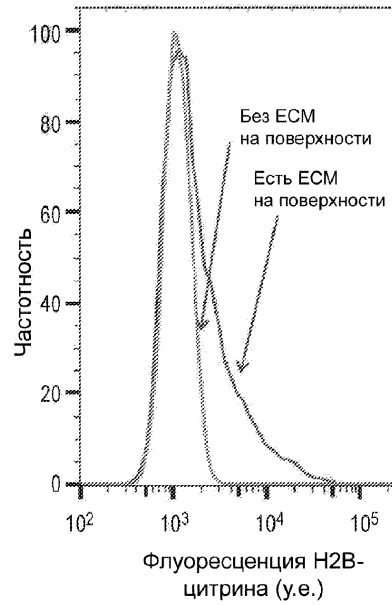
Фиг. 28А



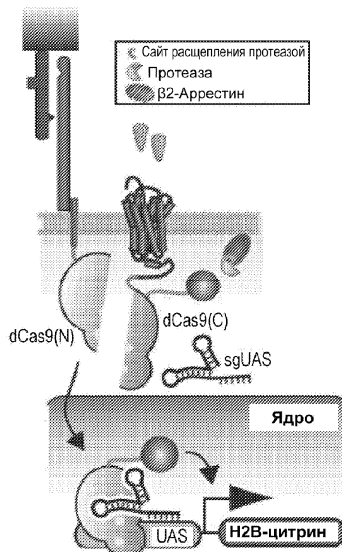
Фиг. 28В



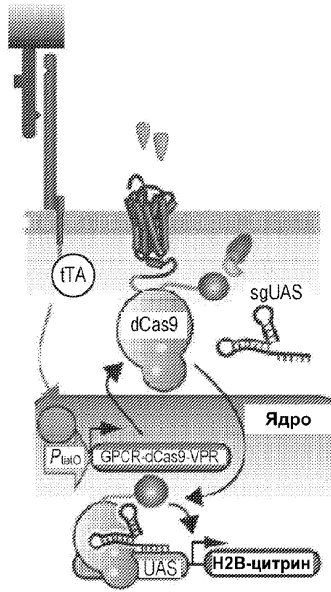
Фиг. 29А



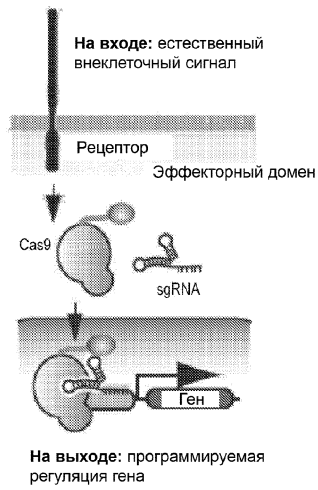
Фиг. 29В



Фиг. 30А

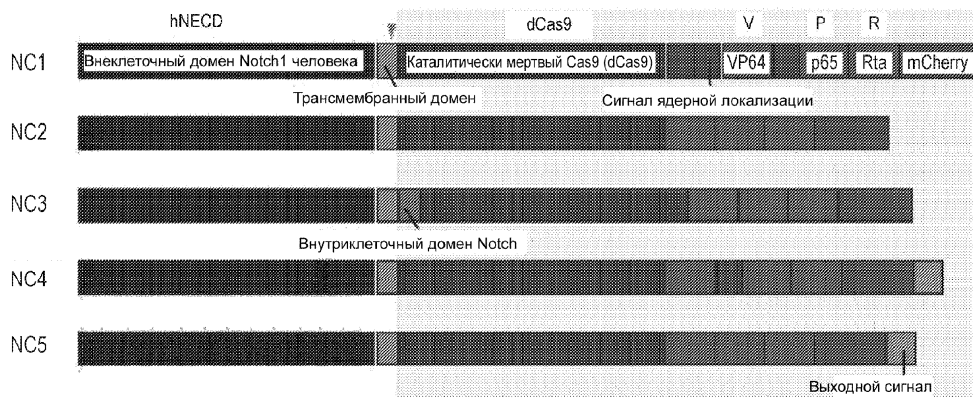


Фиг. 30В

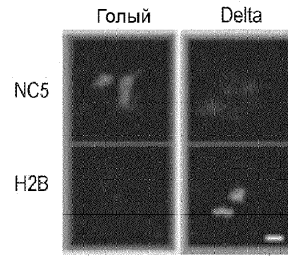
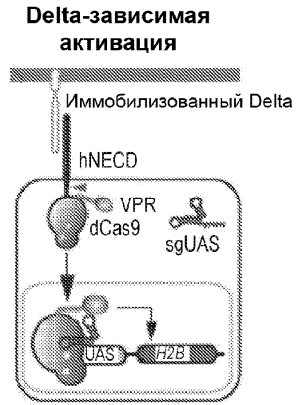


Фиг. 31А

Разработка химерных рецепторов Notch

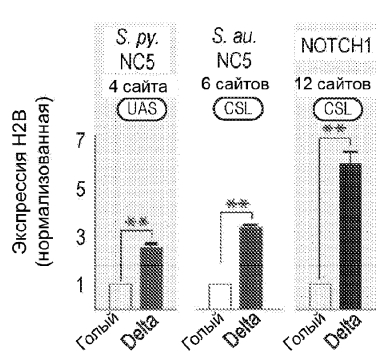


Фиг. 31В

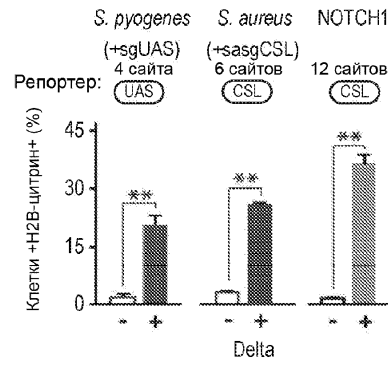


D

C



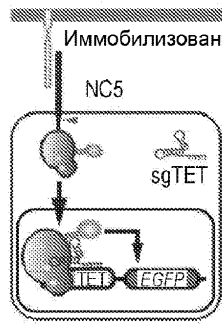
E



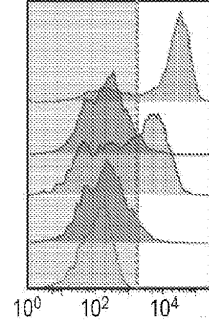
F

Фиг. 31С-F

Проверка репортера EGFP



Выключено Включено

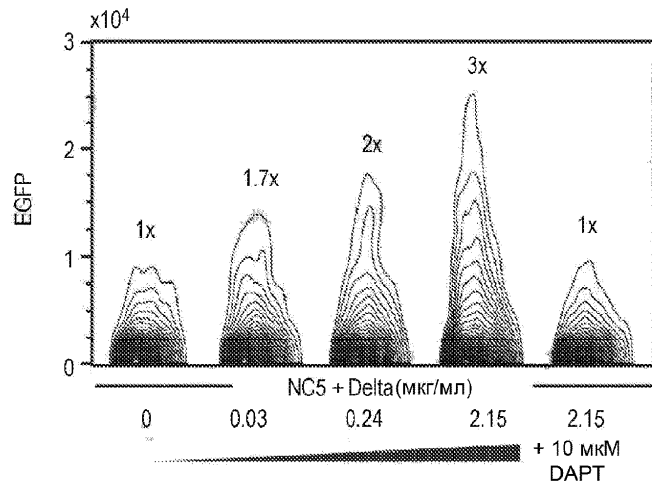


ВКЛЮЧЕНО %	Условия
77	dCas9-VPR
4.0	NC5 +Delta + 10 мкМ ингибитора DAPT
42	NC5 +Delta
5.1	NC5 без Delta
0.3	без конструкции

Активация репортера EGFP

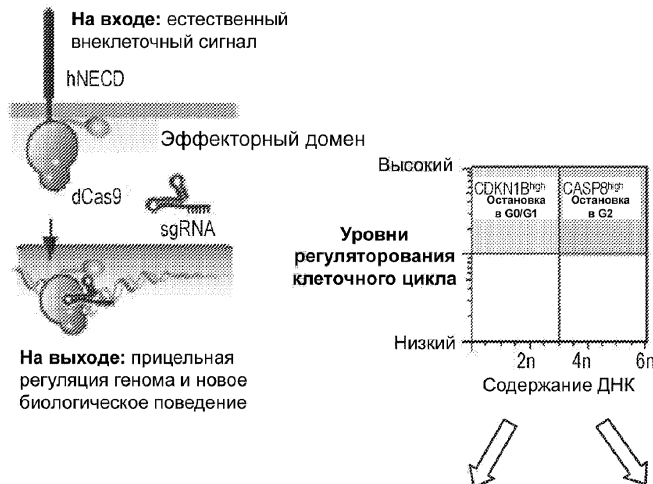
G

H



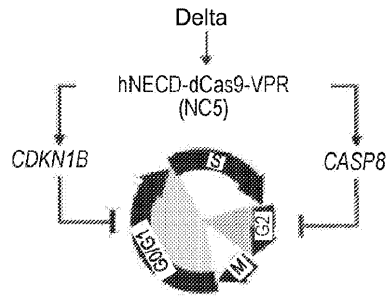
I

Фиг. 31G-I

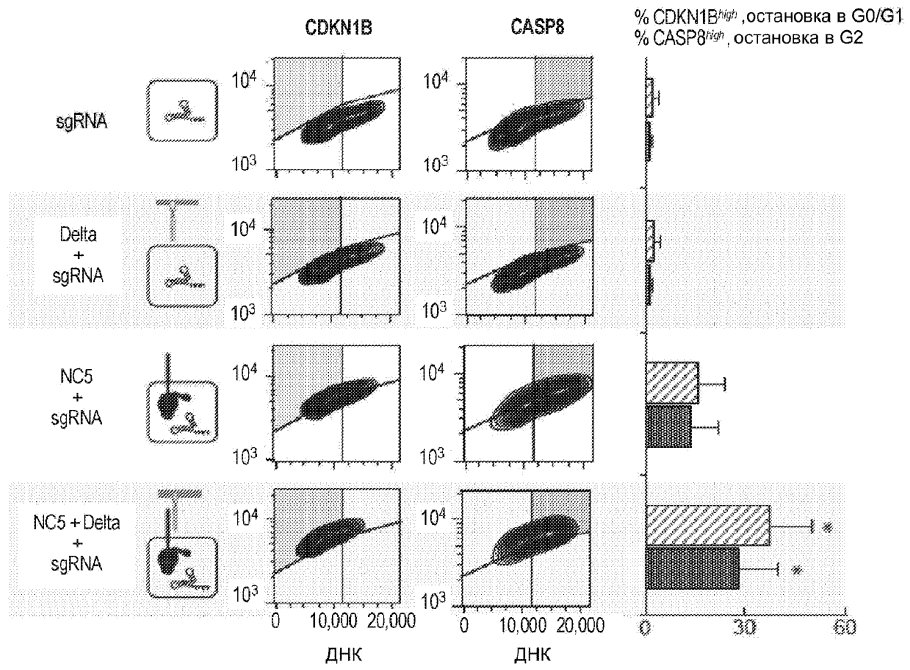


Фиг. 32A

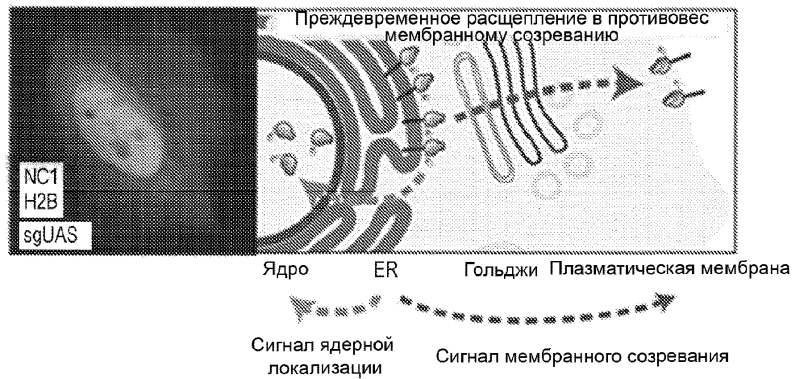
Новое поведение
Delta-зависимая остановка
клеточного цикла



Клеточный цикл
Фиг. 32В

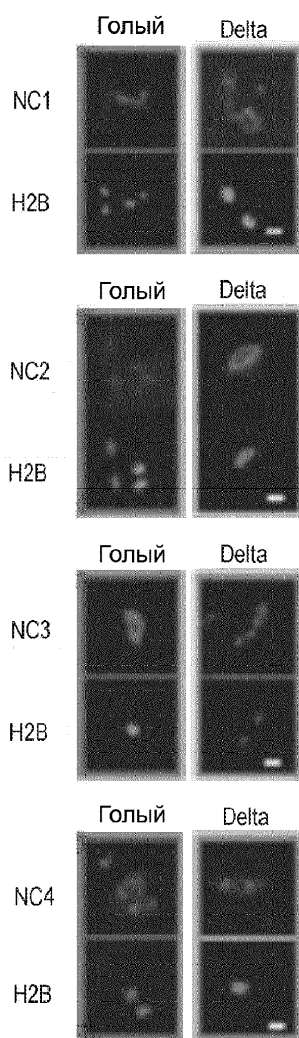


Фиг. 32С



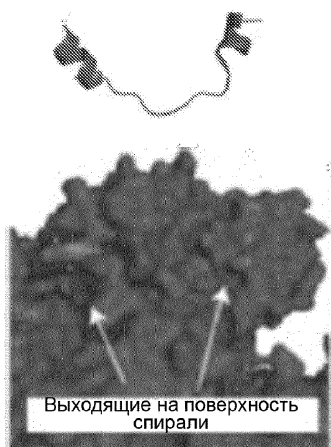
Фиг. 33

Delta-независимая активация

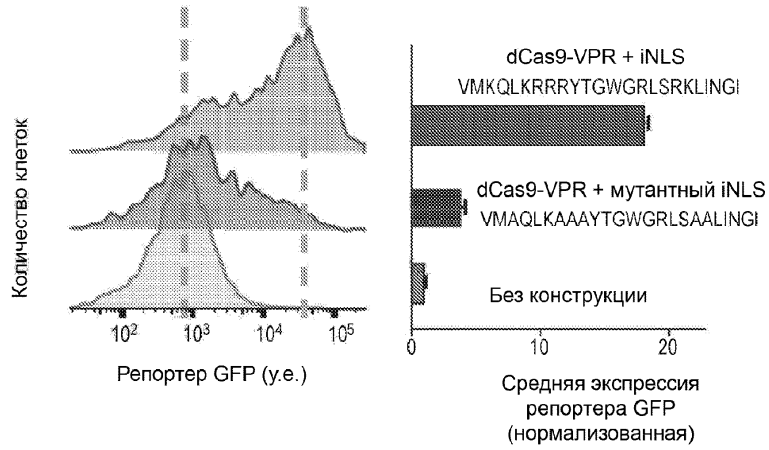


Фиг. 34

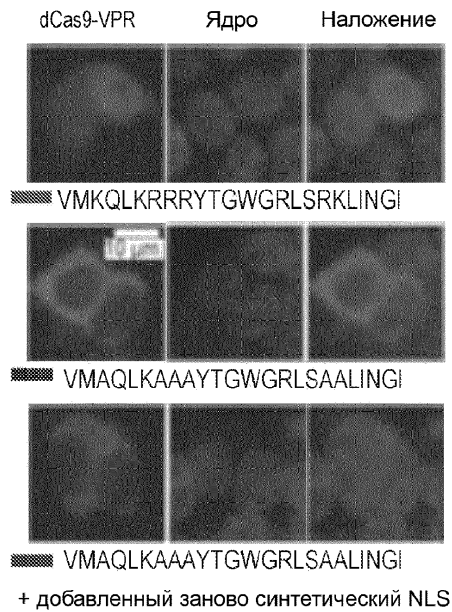
Предполагаемый двухчастный собственный NLS (iNLS)



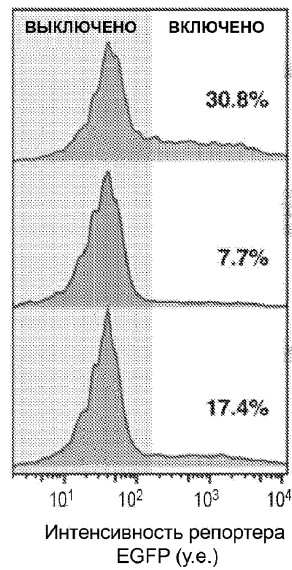
Фиг. 35А



Фиг. 35B

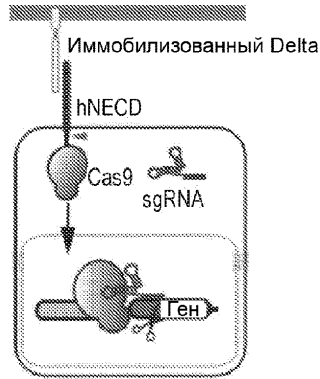


Фиг. 35C

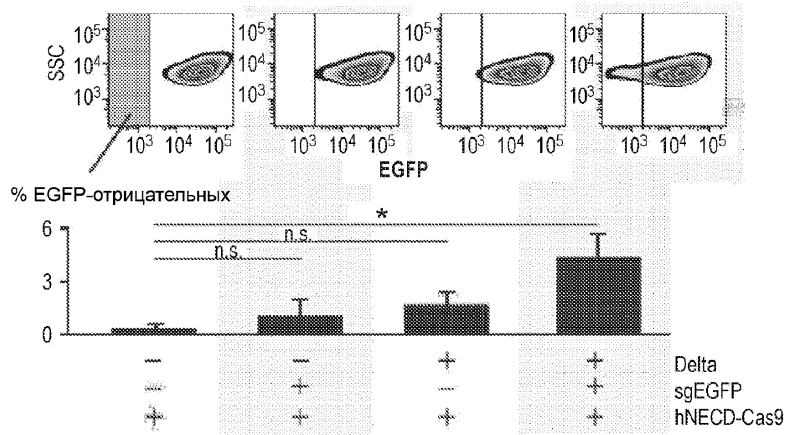


Фиг. 35D

**Delta-зависимое
разрезание гена**

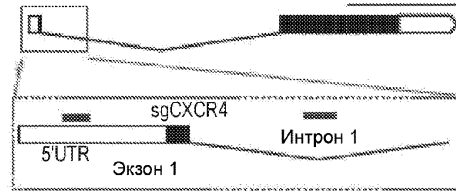


Фиг. 36А

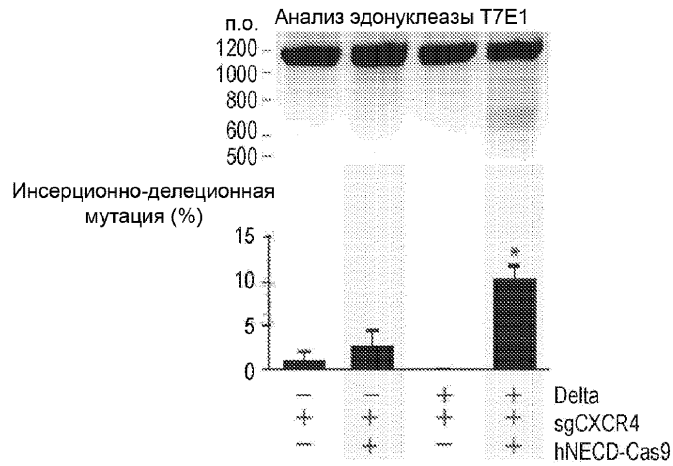


Фиг. 36В

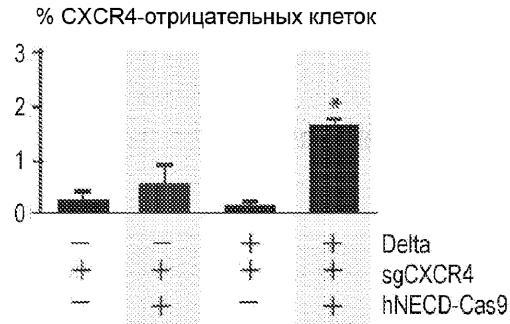
CXCR4



Фиг. 36С



Фиг. 36D

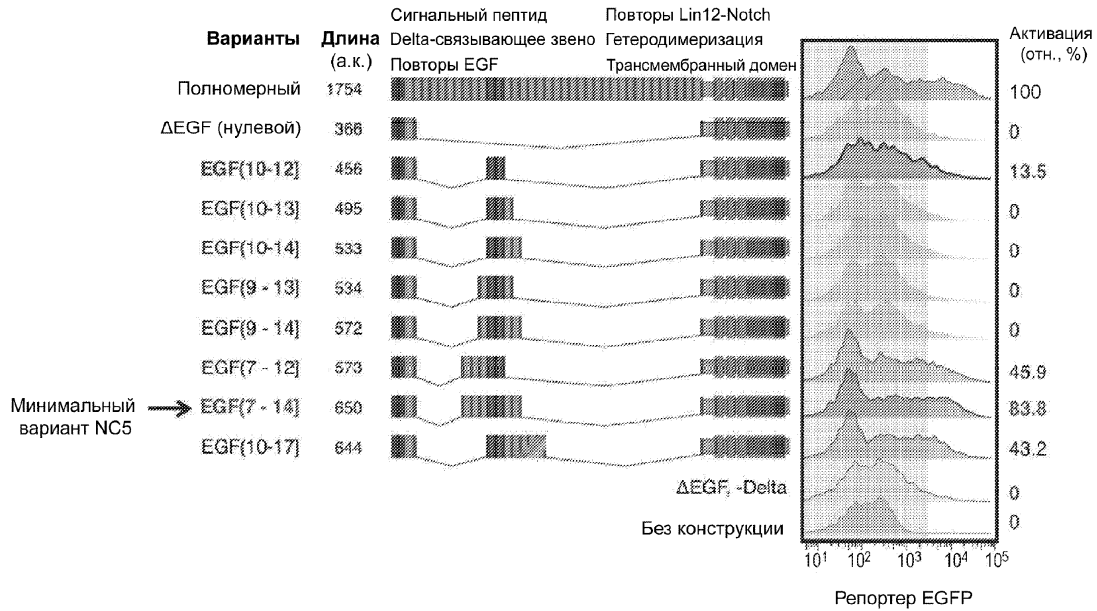


Фиг. 36Е

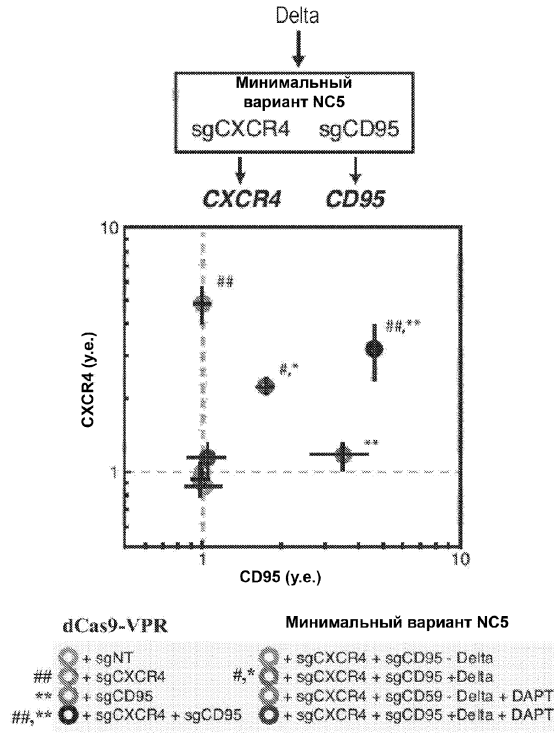
Delta-связывающее звено NECD: повторы EGF 11 и 12

Человек QDVDECSLGNPCEHAGKCIINTLGSFECQCLQGYTGPRCEIDVNECVSNPCQNDATCLDQIGEFQCMCMFPGYEGVHCE
 Xenopus NDVDECSLGNPCEHGGRCNTLGSFQCNCPQGYAGPRCEIDVNECLSNPCQNDSTCLDQIGEFQCIQMPGYEGLYCE
 Данио QDIDECSLGNPCEHGGRCNLTKGSFQCKLQGYEGPRCEMDVNECKSNPCQNDATCLDQIGGFHCICMPGYEGVFCQ
 Drosophila EDIDECDQG SPCEHNGICVNTPGSYRCNCSQGF TGPRCETNINECESHPQNEGSCLDDPGTFRVCMPGFTGTQCE

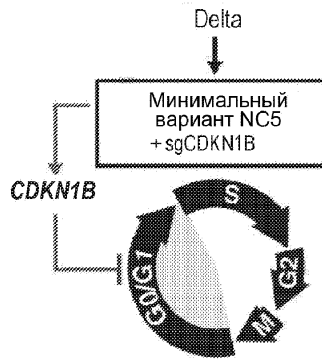
Фиг. 37А



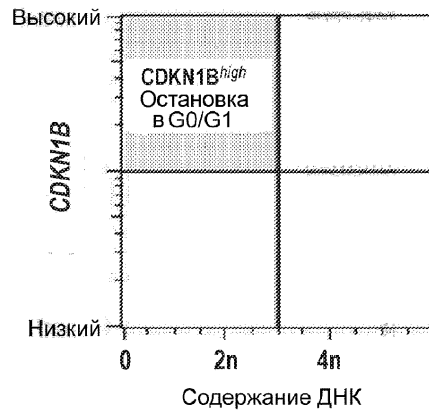
Фиг. 37В



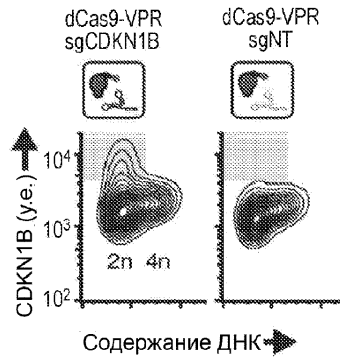
Фиг. 38



Фиг. 39А

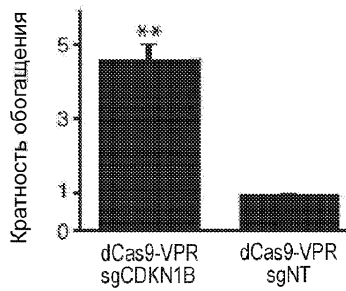


Фиг. 39В

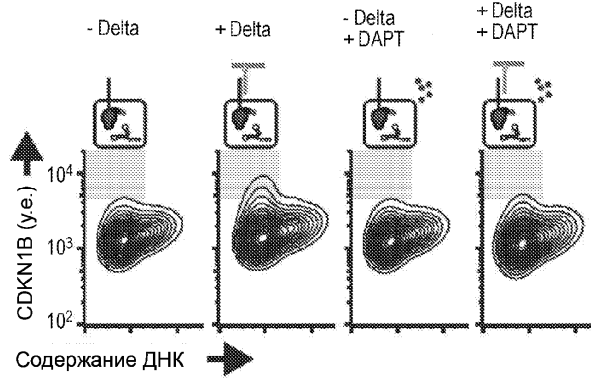


Фиг. 39С

Клетки CDKN1B^{high} с остановкой в G0/G1

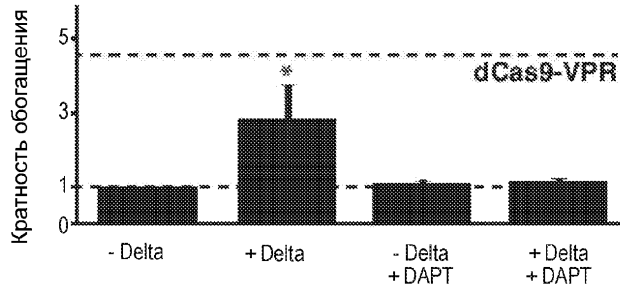


Фиг. 39D



Фиг. 39Е

Клетки CDKN1B^{high} с остановкой в G0/G1



Фиг. 39F

