

(19)



**Евразийское
патентное
ведомство**

(11) **036375**(13) **B1**(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ**(45) Дата публикации и выдачи патента
2020.11.02(51) Int. Cl. *C12N 9/22* (2006.01)
C12N 15/90 (2006.01)(21) Номер заявки
201650071(22) Дата подачи заявки
2015.05.04(54) **TAL-ЭФФЕКТОРНАЯ НУКЛЕАЗА ДЛЯ ЦЕЛЕВОГО НОКАУТА ВИЧ-КОРЕЦЕПТОРА CCR5**(31) **10 2014 106 327.9**(32) **2014.05.07**(33) **DE**(43) **2017.03.31**(86) **PCT/DE2015/200295**(87) **WO 2015/169314 2015.11.12**

(71)(73) Заявитель и патентовладелец:

**ОБЩЕСТВО С ОГРАНИЧЕННОЙ
ОТВЕТСТВЕННОСТЬЮ
"ЭйДжиСиТи" (RU)**

(72) Изобретатель:

Мок Ульрике (GB), Фесе Борис (DE)

(74) Представитель:

**Пустовалова М.Л., Котлов Д.В.,
Черняев М.А., Яремчук А.А. (RU)**

(56) **C. MUSSOLINO ET AL.: "A novel TALE
nuclease scaffold enables high genome editing activity**

in combination with low toxicity", NUCLEIC ACIDS RESEARCH, vol. 39, no. 21, 3 August 2011 (2011-08-03), pages 9283-9293, XP055021128, ISSN: 0305-1048, DOI: 10.1093/nar/gkr597, abstract; figure 5

WO-A1-2011146121

WO-A2-2012093833

WO-A1-2013074999

SANJANA NEVILLE E. ET AL.: "A transcription activator-like effector toolbox for genome engineering", NATURE PROTOCOLS, NATURE PUBLISHING GROUP, GB, vol. 7, no. 1, 1 January 2012 (2012-01-01), pages 171-192, XP009170390, ISSN: 1750-2799, the whole document

E.J. FINE ET AL.: "An online bioinformatics tool predicts zinc finger and TALE nuclease off-target cleavage", NUCLEIC ACIDS RESEARCH, vol. 42, no. 6, 30 December 2013 (2013-12-30), pages e42-e42, XP055218748, ISSN: 0305-1048, DOI: 10.1093/nar/gkt1326, abstract

(57) Изобретение относится к новой TAL-эффektorной нуклеазе (TALEN) для прицельного нокаута ВИЧ-корцептора CCR5. В одном аспекте изобретение представляет собой набор из двух мономеров TAL-эффektorной нуклеазы для инактивации ВИЧ корцептора CCR5 человека, характеризующийся тем, что каждый мономер TAL-эффektorной нуклеазы включает в себя домен с эндонуклеазной активностью II типа и ДНК-связывающий домен TAL-эффektorа, содержащий 19 единиц повтора, каждая из которых имеет пару варьируемых аминокислот RVD (Repeat Variable Diresidue), при этом а) ДНК-связывающий домен TAL-эффektorа одного мономера TAL-эффektorной нуклеазы связывается с последовательностью-мишенью GCTGGTCATCCTCATCCTG и содержит следующую последовательность RVD в составе следующих друг за другом единиц повтора: NH, HD, NG, NH, NH, NG, HD, NI, NG, HD, HD, NG, HD, NI, NG, HD, HD, NG, NN, и б) ДНК-связывающий домен TAL-эффektorа другого мономера TAL-эффektorной нуклеазы связывается с последовательностью-мишенью AGATGTCAGTCATGCTCTT и содержит следующую последовательность RVD в составе следующих друг за другом единиц повтора: NI, NN, NI, NG, NN, NG, HD, NI, NH, NG, HD, NI, NG, NH, HD, NG, HD, NG, NG; при этом домен с эндонуклеазной активностью в каждом мономере TAL-эффektorной нуклеазы является С-концевым по отношению к ДНК-связывающему домену TAL-эффektorа и каждая единица повтора содержит от 33 до 35 аминокислот и при этом указанные мономеры, связываясь с двуцепочечной ДНК в локусе гена ВИЧ корцептора CCR5 человека, образуют TAL-эффektorную нуклеазу, способную образовывать двуцепочечный разрыв в указанной ДНК.

B1**036375****036375 B1**

Изобретение относится к новой TAL-эффektorной нуклеазе (TALEN) для целевого нокаута ВИЧ-корцептора CCR5.

Помимо своей основной функции в клетке хемокиновый рецептор CCR5 играет важную роль при ВИЧ-инфекции. Он выступает в роли корцептора для так называемых CCR5-тропных штаммов ВИЧ, являясь посредником начальной ВИЧ-инфекции. Если на поверхности клетки Т-хелпера отсутствует CCR5, ВИЧ не может связаться с клеткой-хозяином, и инфицирование не происходит. Таким образом, гомозиготная делеция (CCR5Δ32) в гене CCR5, которая встречается приблизительно у 1% западноевропейцев и "белых" американцев ("европеоиды"), практически полностью защищает от ВИЧ-инфицирования CCR5-тропными штаммами. Следовательно, CCR5 является очень интересной мишенью при лечении ВИЧ.

Прежние фармакологические подходы, направленные на блокаду CCR5, требуют пожизненного лечения в рамках комбинированной антиретровирусной терапии, АРТ. В долгосрочной перспективе это связано с возникновением потенциально тяжелых побочных эффектов, а также с отсутствием комплайенса пациентов и развитием резистентности. Напротив, генетического разрушения ("нокаута") CCR5 (в плане генной терапии), в идеальном случае, было бы достаточно в качестве однократного лечения, так как генетическая защита распространяется на все дочерние клетки. Это подтверждается не только естественной резистентностью CCR5Δ32-гомозиготных индивидов, но и описанными случаями успешного лечения ВИЧ-инфекции у так называемого "берлинского пациента" после аллогенной трансплантации стволовых клеток с CCR5Δ32-гомозиготными донорскими клетками (Hütter G. et al. Long-term control of HIV by CCR5Δ32/Delta32 stem-cell transplantation. *N Engl J Med*, 2009, 360: 692-698; Allers K. et al. Evidence for the cure of HIV infection by CCR5Δ32/Δ32 stem cell transplantation. *Blood*, 2011; 117: 2791-2799).

На основе этих наблюдений были разработаны концепции генетического нокаута CCR5 у ВИЧ-инфицированных пациентов. В настоящее время наиболее перспективные стратегии основаны на так называемых "дизайнерских нуклеазах" (см., например, Manjunath N. et al., *Newer Gene Editing Technologies toward HIV. Gene Therapy, Viruses*, 2013, 5, 2748-2766). Эти дизайнерские нуклеазы состоят из двух компонентов: распознающего домена, который определяет специфичность в геноме и может быть практически полностью разработан без ограничений, и нуклеазного домена, который индуцирует двухцепочечный разрыв в выбранном месте генома. Негомологичное восстановление концов данного разрыва с помощью внутриклеточных систем быстрой репарации ДНК приводит к смещению (нарушению) открытой рамки считывания целевого гена и, таким образом, в идеальном случае, к нокауту. Первыми широко применяемыми дизайнерскими нуклеазами были нуклеазы "цинковые пальцы" (ZFN). Например, "Sangamo BioSciences, Inc." в данное время тестирует разработанную ZFN ими CCR5-специфическую нуклеазу цинкового пальца под названием SB-728 (<http://www.sangamo.com/pipeline/sb-728.html>) для клинического применения (Tebas et al., *Gene Editing of CCR5 in Autologous CD4 T Cells of Persons Infected with HIV. N Engl J Med*, 2014;370:901-10). Клинические исследования продемонстрировали осуществимость подхода, однако долгосрочный клинический эффект, выражаемый в виде вирусной нагрузки, наблюдался только у испытуемого, который оказался гетерозиготным по естественной мутации CCR5Δ32.

TAL-эффektorные нуклеазы (эффektorные нуклеазы, подобные активаторам транскрипции, TALEN) являются следующим поколением дизайнерских нуклеаз (см., например, Mussolino, C., Cathomen T. TALE nucleases: tailored genome engineering made easy. *Curr Opin Biotechnol.*, 2012, 23(5): 644-50; WO 2011/072246 A2; EP 2510096 A2; WO 2011/154393 A1; WO 2011/159369 A1; WO 2012/093833 A2; WO 2013/182910 A2).

По сравнению с ZFN они отличаются в том числе более высокой специфичностью, что значительно снижает риск таких нецелевых эффектов, как закрепление мутации в другом, отличном от желаемого, месте генома (Handel E.-M., Cathomen T. Zinc-finger nuclease based genome surgery: it's all about specificity. *Curr Gene Ther*, 2011, 11: 28-37; Mussolino C. et al. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.*, 2011; 39: 9283-9293).

CCR5-специфические TALEN уже известны (см., например, Mussolino C. et al. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.*, 2011; 39: 9283-9293; WO 2011/146121 A1; WO 2012/093833 A2; US 2013/0217131 A1), клиническое исследование, однако пока еще не описано.

По-прежнему существует потребность в средствах для эффективного лечения ВИЧ-инфекции. Поэтому задачей изобретения является создание такого средства. В частности, задачей данного изобретения является создание препарата, при помощи которого может быть достигнут более эффективный нокаут ВИЧ-корцептора CCR5, чем созданный ранее.

Задача решается при осуществлении изобретения по п.1, а также других независимых пунктов формулы изобретения. Частные дополнительные варианты осуществления настоящего изобретения представлены в зависимых пунктах формулы.

В первом аспекте изобретение представляет пару TAL-эффektorных нуклеаз, которая включает в себя набор из двух мономеров TAL-эффektorной нуклеазы для инактивации ВИЧ корцептора CCR5 человека, характеризующийся тем, что каждый мономер TAL-эффektorной нуклеазы включает в себя

домен с эндонуклеазной активностью II типа и ДНК-связывающий домен TAL-эффектора, содержащий 19 единиц повтора, каждая из которых имеет пару переменных аминокислот RVD (Repeat Variable Di-residue), при этом:

а) ДНК-связывающий домен TAL-эффектора одного мономера TAL-эффекторной нуклеазы связывается с последовательностью-мишенью GCTGGTCACTCCATCCTG (SEQ ID NO: 1) и содержит следующую последовательность RVD в составе следующих друг за другом единиц повтора: NH, HD, NG, NH, NH, NG, HD, NI, NG, HD, HD, NG, HD, NI, NG, HD, HD, NG, NN, и

б) ДНК-связывающий домен TAL-эффектора другого мономера TAL-эффекторной нуклеазы связывается с последовательностью-мишенью AGATGTCACTCATGCTCTT (SEQ ID NO: 2) и содержит следующую последовательность RVD в составе следующих друг за другом единиц повтора: NI, NN, NI, NG, NN, NG, HD, NI, NH, NG, HD, NI, NG, NH, HD, NG, HD, NG, NG;

при этом домен с эндонуклеазной активностью в каждом мономере TAL-эффекторной нуклеазы является С-концевым по отношению к ДНК-связывающему домену TAL-эффектора и каждая единица повтора содержит от 33 до 35 аминокислот;

при этом указанные мономеры, связываясь с двуцепочечной ДНК в локусе гена ВИЧ корецептора CCR5 человека, образуют TAL-эффекторную нуклеазу, способную образовывать двуцепочечный разрыв в указанной ДНК.

Пара TAL-эффекторных нуклеаз в соответствии с настоящим изобретением способна вызвать нокаут CCR5 в первичных Т-лимфоцитах с не имеющей себе равных эффективностью >50%. К тому же, как ни удивительно, изобретение делает возможным последовательный биаллельный нокаут обеих CCR5-аллелей и, таким образом, обеспечивает полную защиту модифицированных клеток от проникновения ВИЧ, вопреки мнению ведущих экспертов в данной области техники, что это в настоящее время невозможно ("Consistent nuclease-mediated biallelic knockdown is not yet tenable", см. Kay, M.A. und Walker, B.D., 2014, Engineering Cellular Resistance to HIV, N Engl J Med, 370:968-969). Кроме того, было установлено, что TALEN-пара в соответствии с изобретением идеально подходит для доставки в клетку, основанной на трансфекции в виде мРНК (щадящего, безопасного и GMP-совместимого метода). Таким образом, изобретение впервые представляет основанное на дизайнерской нуклеазе средство для лечения ВИЧ, которое сочетает в себе высокую нокаут-эффективность и нокаут-селективность с низким уровнем нецелевых эффектов и другие, полезные с фармакологической точки зрения, свойства.

Под "TAL-эффекторная нуклеаза" или "TALEN" (эффекторные нуклеазы, подобные активаторам транскрипции) подразумевается гибридный белок, который содержит ДНК-связывающий домен TAL-эффектора (TALE) и ДНК-расщепляющий домен эндонуклеазы рестрикции. TAL-эффекторы представляют собой ДНК-связывающие белки, которые получают из таких фитопатогенов, как *Xanthomonas spp.*

ДНК связывание TAL-эффекторов опосредовано через домен с переменным количеством (как правило, от 5 до 30) единиц повтора ("повторов"), которые, как правило, формируются из 33-35 аминокислот. Каждая единица повтора имеет два высоко переменных аминокислотных остатка (repeat-variable di-residue, RVD), расположенных, как правило, в позициях 12 и 13, и связывающихся только с одним основанием целевой последовательности ДНК. Отношение между RVD и нуклеотидами целевой молекулы ДНК приводятся ниже.

RVD (однолитерный код)	RVD (трехлитерный код)	Нуклеотид(ы)
NH	Asn-His	G
HD	His-Asp	C
NG	Asn-Gly	T
NI	Asn-Ile	A
NN	Asn-Asn	R (G, A)
NK	Asn-Lys	G
NS	Asn-Ser	N (A, C, G, T)

Под "RVD-последовательностью" здесь подразумевается непрерывная последовательность RVD в одном ДНК-связывающем домене TAL-эффектора, в котором последовательность указана, если не указано иное, в направлении N-C, т.е. от N-конца к C-концу.

Очевидно, что в данном случае специалисту в данной области техники известно, что "RVD-последовательности" не следуют непосредственно друг за другом, поскольку они не связаны непосредственно друг с другом ковалентными связями, а непосредственно связаны между собой единицы повтора ("повторы"), в которых содержатся RVD, таким образом, соответствующие RVD отделены аминокислотами основной структуры единиц повтора.

Под "последовательностью-мишенью" здесь подразумевается нуклеотидная последовательность, как правило, последовательность ДНК, которая связана связывающим доменом TAL-эффектора.

Раскрытые в данном документе RVD-последовательности, каждая из которых состоит из 19 RVD, имеют следующие последовательности-мишени (в направлении 5'-3'; однолитерный код для аминокислот):

NH HD NG NH NH NG HD NI NG HD HD NG HD NI NG HD HD NG NN
GCTGGTCATCCTCATCCTG (SEQ ID NO: 1)

NI NN NI NG NN NG HD NI NH NG HD NI NG NH HD NG HD NG NG
AGATGTCAGTCATGCTCTT (SEQ ID NO: 2)

Под "единицей повтора" ("повтором") относительно связывающего домена TAL-эфффектора здесь подразумевается непрерывная последовательность, как правило, из 33-35, в основном 34 аминокислот, которые, кроме высоко варьируемых аминокислот RVD в позициях 12 и 13, имеют практически идентичные аминокислотные последовательности. Вполне возможно, что и в консервативной основной структуре единицы повтора, т.е. по сути в однородной структуре, в которую включены высоко варьируемые аминокислоты RVD, варьируют единичные аминокислоты, например, в позициях 4, 10 и/или 32 в единице повтора из 34 аминокислот. Типичная единица повтора может иметь, например, следующую последовательность аминокислот (подстрочные цифры определяют позицию внутри единицы повтора):

LTPX₄QVVAIX₁₀SX₁₂X₁₃GGKQALETVQRLLPVLCQX₃₂HG (SEQ ID NO: 5)

X обозначает любую аминокислоту, при этом в позициях 12 и 13 стоят гиперварируемые аминокислоты RVD. В позиции 4 (X4) могут стоять, например, аминокислоты E, Q, D или A, в позиции 32 (X32) - аминокислоты A или D. В позиции 10 может стоять, например, A или V. Примеры единиц повтора приводятся ниже (XX обозначает гиперварируемые аминокислоты RVD, варьируемые аминокислоты подчеркнуты):

LTPEQVVAIASXXGGKQALETVQRLLPVLCQANG (SEQ ID NO: 6)

LTPQQVVAIASXXGGKQALETVQRLLPVLCQANG (SEQ ID NO: 7)

LTPDQVVAIASXXGGKQALETVQRLLPVLCQDHG (SEQ ID NO: 8)

LTPAQVVAIASXXGGKQALETVQRLLPVLCQDHG (SEQ ID NO: 9)

LTPEQVVAIVSXXGGKQALETVQRLLPVLCQANG (SEQ ID NO: 10)

Связывающий домен TAL-эфффектора может включать в себя один или несколько подобных вариантов единиц повтора, причем в расчет принимаются также комбинации различных вариантов.

Крайняя, ближайшая к нуклеазному домену единица повтора может содержать меньше, например лишь первые 15, 16, 17, 18, 19 или 20 аминокислот остальных единиц повтора. Такая единица повтора называется "половинная единица повтора" или "полуповтор".

Под "ДНК-связывающим доменом" здесь подразумевается участок белка, который индуцирует связывание белка с молекулой ДНК. В случае ДНК-связывающего домена TAL-эфффектора это происходит посредством описанных выше единиц повтора ("повторов").

Формулировку, согласно которой ДНК-связывающий домен TAL-эфффектора связывается с последовательностью-мишенью, следует понимать так, что ДНК-связывающий домен TAL-эфффектора за счет его RVD-последовательности специфично связывается с ДНК-последовательностью-мишенью. Вместе с тем не обязательно, хотя предпочтительно, чтобы с каждым нуклеотидом последовательности-мишени был связан RVD в связывающем домене. Связь между RVD-последовательностью ДНК-связывающего домена TAL-эфффектора и последовательностью-мишенью должна быть именно такой, чтобы произошло связывание с последовательностью-мишенью. "Специфично" в данном контексте означает, что связывание происходит в основном только с последовательностью-мишенью.

Под "мономером TAL-эффекторной нуклеазы" подразумевается TAL-эффекторная нуклеаза, которая состоит из одной единственной полипептидной цепи. Под "парой TAL-эффекторных нуклеаз" или "TALEN-парой" подразумевается TALEN, состоящая из двух мономеров TAL-эффекторной нуклеазы. Мономеры представляют собой левое или правое плечо TALEN, которые связывают противоположные нити ДНК и вместе вызывают разрыв ДНК в одном сайте.

Ссылка на TALEN-пару из "левой" или "правой" TALEN, или "левого" или "правого" TALEN-плеча отображает тот факт, что в одной TALEN-паре TALEN-мономеры используются попарно, т.е. вызывают разрыв нитей внутри двухцепочечной ДНК, так как один мономер связывается с последовательностью-мишенью смысловой нити, в то время как другой TALEN-мономер TALEN-пары связывается с последовательностью-мишенью комплементарной антисмысловой нити, а именно таким образом, что нуклеазные домены направлены друг к другу, в одном общем участке ДНК, называемом "спейсер", располагаются между последовательностями-мишенями и каждый из нуклеазных доменов вызывает одноцепочечный разрыв. "Левый" и "правый" TALEN-мономеры являются частями определенной TALEN-пары, причем "левой" TALEN зачастую обозначают ту TALEN, которая связывается со смысловой нитью, в то время как "правая" TALEN связывается с комплементарной нитью. Тем не менее, ссылка на "левую" или "правую" TALEN, при этом, необязательно указывает на то, что "левая" TALEN связывается со смысловой нитью, а "правая" - с комплементарной нитью.

Таким образом, к данному изобретению относится также "TALEN-пара", т.е. пара из двух относя-

щихся к изобретению мономеров, каждый из которых представляет левое или правое плечо TALEN. Также к данному изобретению относится пара TAL-эффекторных нуклеаз, содержащая мономер TAL-эффекторной нуклеазы, ДНК-связывающий домен TAL-эффектора которого связывается с последовательностью-мишенью GCTGGTCATCCTCATCCTG (SEQ ID NO: 1) и/или содержит RVD-последовательность NH HD NG NH NH NG HD NI NG HD HD NG HD NI NG HD HD NG NN; и мономер TAL-эффекторной нуклеазы, ДНК-связывающий домен TAL-эффектора которого связывается с последовательностью-мишенью AGATGTCAGTCATGCTCTT (SEQ ID NO: 2) и/или содержит RVD-последовательность NI NN NI NG NN NG HD NI NH NG HD NI NG NH HD NG HD NG NG.

Под "эндонуклеазным доменом со II типом эндонуклеазной активности" подразумевается полипептид, который обладает ДНК-расщепляющей активностью нуклеазы рестрикции и разрезает ДНК внутри или в непосредственной близости от последовательности узнавания, не требующий АТФ и не обладающий активностью метилтрансфераз. Под "эндонуклеазным доменом с IIS типом эндонуклеазной активности" подразумевается домен эндонуклеазы II типа, сайт расщепления которого находится в непосредственной близости от последовательности узнавания, но не внутри нее.

Под "CCR5" подразумевается человеческий C-C рецептор хемокина 5 (обозначаемый также CD195, CMKBR5 или CC-CKR5). Последовательность человеческого CCR5 показана в SEQ ID NO: 11 (см. номер доступа NC_018914.2 в базе данных NCBI).

Под "вектором" подразумевается система транспортировки для доставки в основном чужеродных нуклеиновых кислот в живую клетку-реципиента путем трансфекции или трансдукции. Под "вектором-переносчиком генов" подразумевается вектор, при помощи которого ген может быть введен в клетку. Векторы (-переносчики генов) хорошо известны специалистам. Примерами векторов-переносчиков генов являются плазмиды, вирусные векторы или мРНК.

Под "нуклеиновой кислотой" подразумевается полимер, мономерами которого являются нуклеотиды. Нуклеотид является соединением, состоящим из остатка сахара, азотсодержащего гетероциклического органического основания (нуклеотидного основания или нуклеооснования) и фосфатной группы. Остатком сахара, как правило, является пентоза, в случае ДНК это дезоксирибоза, в случае РНК это рибоза. Связь между нуклеотидами осуществляется через фосфатную группу при помощи фосфодиэфирных мостиков, как правило, между 3'-углеродным атомом сахара нуклеозида (соединения сахара и нуклеооснования) и 5'-углеродным атомом сахара следующего нуклеозида. Термин "нуклеиновая кислота" включает, например, ДНК, РНК и ДНК/РНК-гибриды. В том смысле, в котором он используется здесь, термин "нуклеиновая кислота" относится, прежде всего, к изолированной нуклеиновой кислоте. Под "изолированной нуклеиновой кислотой" подразумевается синтетически изготовленная или извлеченная из ее естественной или изначальной среды нуклеиновая кислота.

Термин "включающий в себя" используется таким образом, что он определяет не только объект, который проявляет признаки, соответствующие исключительно данному понятию, но и объект, который проявляет признаки, соответствующие данному определению, а также дополнительные признаки. Определение объекта благодаря тому, что он включает в себя определенные признаки, включает также определение данного объекта по окончательному списку этих признаков, т.е. по наличию исключительно этих признаков.

В предпочтительном варианте относящейся к изобретению пары TAL-эффекторных нуклеаз, эндонуклеазный домен в каждом мономере TAL-эффекторной нуклеазы является С-концевым по отношению к ДНК-связывающему домену TAL-эффектора. Предпочтительно каждый повтор, за исключением того, который примыкает непосредственно к эндонуклеазному домену, включает в себя 33-35 аминокислот, преимущественно 34 аминокислоты, причем RVD находятся в каждой единице повтора в позициях 12 и 13. Особое предпочтение отдается всем единицам повтора, за исключением "полуповторов" аминокислотной последовательности соответствующей SEQ ID NO: 5, при чем в позиции 4 может стоять E, Q, D или A, в позиции 10 может стоять A или V и в позиции 32 может стоять A или D. Основная структура единиц повтора может быть одинаковой или различной для всех единиц повтора. Относительно аминокислот возможны изменения одной или нескольких единиц повтора в позициях внутри основной структуры, например в позициях 4, 10 и/или 32. Единица повтора, непосредственно примыкающая к эндонуклеазному домену, может включать меньшее число аминокислот, например 15, 16, 17, 18, 19 или 20, причем аминокислоты в этом случае соответствуют преимущественно первым 15, 16, 17, 18, 19 или 20 аминокислотам в других повторах. При этом, например, аминокислота в позиции 4 может отличаться, быть, например, E, Q, D или A, и/или может отличаться аминокислота в позиции 10, быть, например, V вместо A.

Особое предпочтение отдается эндонуклеазному домену мономера TAL-эффекторной нуклеазы эндонуклеазного домена IIS типа, особенно ДНК-расщепляющему домену эндонуклеазы FokI. Аминокислотная последовательность для подходящего расщепляющего домена FokI приведена в SEQ ID NO: 12. В расчет берется также другой расщепляющий домен эндонуклеазы II типа. Эндонуклеазы II типа известны специалистам, и подходящие расщепляющие домены могут быть определены при помощи рутинных исследований.

В особо предпочтительном варианте первый мономер TAL-эффекторной нуклеазы содержит ами-

нокислотную последовательность, соответствующую SEQ ID NO: 3, а второй мономер TAL-эффекторной нуклеазы включает в себя аминокислотную последовательность, соответствующую SEQ ID NO: 4. В SEQ ID NO: 3 приведена левая TALEN (далее обозначаемая также CCR5-Uco-L или левое плечо CCR5-Uco), а в SEQ ID NO: 4 - правая TALEN (далее обозначаемая также CCR5-Uco-R или правое плечо CCR5-Uco) TALEN-пары, которые вместе вызывают двухцепочечный разрыв в ДНК-последовательности CCR5 внутри спейсера, расположенного между последовательностями-мишенями, соответствующими SEQ ID NO: 1 и SEQ ID NO: 2. Восстановление данного двухцепочечного разрыва происходит за счет клеточной репарационной системы (негомологичного соединения концов, НСК) с высокой вероятностью нарушения рамки считывания и, таким образом, к нокауту CCR5.

Во втором аспекте данное изобретение представляет также нуклеиновую кислоту, включающую в себя:

а) первую нуклеиновую кислоту, которая кодирует первый мономер TAL-эффекторной нуклеазы, где первый мономер TAL-эффекторной нуклеазы содержит эндонуклеазный домен со II типом эндонуклеазной активности и ДНК-связывающий домен TAL-эффектора с множеством единиц повтора, каждая из которых имеет пару вариабельных аминокислот (RVD), и при этом ДНК-связывающий домен TAL-эффектора связывается с последовательностью-мишенью GCTGGTCATCCTCATCCTG (SEQ ID NO: 1) и/или содержит RVD-последовательность NH HD NG NH NH NG HD NI NG HD HD NG HD NI NG HD HD NG NN, и

б) вторую нуклеиновую кислоту, которая кодирует второй мономер TAL-эффекторной нуклеазы, где второй мономер TAL-эффекторной нуклеазы содержит эндонуклеазный домен со II типом эндонуклеазной активности и ДНК-связывающий домен TAL-эффектора с множеством единиц повтора, каждая из которых имеет пару вариабельных аминокислот (RVD), и при этом ДНК-связывающий домен TAL-эффектора связывается с последовательностью-мишенью AGATGTCAGTCATGCTCTT (SEQ ID NO: 2) и/или содержит RVD-последовательность NI NN NI NG NN NG HD NI NH NG HD NI NG NH HD NG HD NG NG TALEN-мономеры, составляющие относящуюся к изобретению TALEN-пару, в данном аспекте изобретения кодируются вместе в общей нуклеиновой кислоте. На примере нуклеиновой кислоты, речь может идти о плазмиде или другом подходящем (переносчике генов) векторе. Подходящие векторы и способы их получения и применения хорошо известны на современном техническом уровне. При известных условиях нуклеиновая кислота наряду с TALEN-кодом может содержать и другие элементы, например один или несколько промоторов, сигналы полиаденилирования и др.

TALEN-мономеры, составляющие относящуюся к изобретению TALEN-пару, могут также кодироваться в двух отдельных нуклеиновых кислотах. В третьем аспекте данное изобретение представляет комбинацию нуклеиновых кислот, включающую в себя:

а) первую нуклеиновую кислоту, которая кодирует первый мономер TAL-эффекторной нуклеазы, где первый мономер TAL-эффекторной нуклеазы содержит эндонуклеазный домен со II типом эндонуклеазной активности и ДНК-связывающий домен TAL-эффектора с множеством единиц повтора, каждая из которых имеет пару вариабельных аминокислот (RVD), и при этом ДНК-связывающий домен TAL-эффектора связывается с последовательностью-мишенью GCTGGTCATCCTCATCCTG (SEQ ID NO: 1) и/или содержит RVD-последовательность NH HD NG NH NH NG HD NI NG HD HD NG HD NI NG HD HD NG NN, и

б) вторую нуклеиновую кислоту, которая кодирует второй мономер TAL-эффекторной нуклеазы, где второй мономер TAL-эффекторной нуклеазы содержит эндонуклеазный домен со II типом эндонуклеазной активности и ДНК-связывающий домен TAL-эффектора с множеством единиц повтора, каждая из которых имеет пару вариабельных аминокислот (RVD), и при этом ДНК-связывающий домен TAL-эффектора связывается с последовательностью-мишенью AGATGTCAGTCATGCTCTT (SEQ ID NO: 2) и/или содержит RVD-последовательность NI NN NI NG NN NG HD NI NH NG HD NI NG NH HD NG HD NG NG.

В случае и первой, и второй нуклеиновой кислоты, речь идет, предпочтительно о мРНК, особенно предпочтительно о консервированной мРНК (см., например, Kallen K.-J. et al., A novel, disruptive vaccination technology, *Hum Vaccin Immunother.* Oct 1, 2013; 9(10): 2263-2276, doi: 10.4161/hv.25181; Kallen K.-J. und Theß A., A development that may evolve into a revolution in medicine: mRNA as the basis for novel, nucleotide-based vaccines and drugs, *Ther Adv Vaccines.*, Jan 2014; 2(1): 10-31, doi: 10.1177/2051013613508729). При известных условиях первая и вторая нуклеиновые кислоты наряду с TALEN-кодом могут содержать и другие элементы, например один или несколько промоторов, сигналы полиаденилирования и др. мРНК, подходящие для левого и правого плеча, относящегося к изобретению TALEN, примерно указаны в SEQ ID NO: 17 и 18.

В случае мРНК, она транспортируется в клетку(и)-мишень(и), например Т-лимфоциты, особое предпочтение отдается методу, описанному Berdien и др. (Berdien B. et al., TALEN-mediated editing of endogenous T-cell receptors facilitates efficient reprogramming of T lymphocytes by lentiviral gene transfer, *Gene Therapy*, 2014, doi:10.1038/gt.2014.26). Особое предпочтение отдается одновременному введению обоих плеч TALEN (правому и левому плечу) в клетку.

Введение относящихся к изобретению TALEN-пар при помощи мРНК имеет ряд несомненных пре-

имущества для клинического применения. Можно отказаться от использования вектора-переносчика генов на основе ДНК, что несомненно упростит изготовление и практическое использование. мРНК-опосредованная экспрессия TALEN настолько коротка, что мРНК очень быстро деградирует в клетке-мишени. Это дополнительно уменьшает риск возникновения нецелевых эффектов. Кроме того, клетки-мишени должны культивироваться весьма короткое время *in vitro*. Соответствующая технология легко приспособляется к требованиям GMP. Ожидается, что, в противоположность к вирусным или плазмидным векторам, будут отсутствовать побочные эффекты, возникающие вследствие самого переноса генов (например, инсерционный мутагенез) или, возникшей в результате нежелательного внедрения вектора, длительной TALEN-экспрессии (нецелевые эффекты, активация TALEN-специфического иммунного ответа).

В дополнительных аспектах данное изобретение представляет вектор, в частности вектор-переносчик генов, включающий в себя относящуюся к изобретению нуклеиновую кислоту и изолированную клетку-хозяина, включающую в себя относящийся к изобретению вектор, относящуюся к изобретению нуклеиновую кислоту или комбинацию нуклеиновых кислот, причем в отношении изолированных клеток-хозяев речь не идет о зародышевых клетках живого организма человека, в частности о зародышевых клетках человека или эмбриональных зародышевых клетках человека, а также и об эмбриональных стволовых клетках, при получении которых был или будет уничтожен человеческий эмбрион. Кроме того, в одном дополнительном аспекте данное изобретение представляет фармацевтическую композицию, включающую в себя нуклеиновую кислоту, комбинацию нуклеиновых кислот или вектор, которые соответствуют данному изобретению. Фармацевтический состав может содержать вспомогательные вещества, например, такие как растворители, усилители растворимости, растворы-ускорители, солеобразующие агенты, соли, буферы, вещества, влияющие на вязкость и консистенцию, гелеобразователи, эмульгаторы, солилизаторы, увлажнители, сепараторы, антиоксиданты, консерванты, наполнители, носители и т.д.

Еще в одном дополнительном аспекте данное изобретение представляет лекарственное средство, включающее в себя нуклеиновую кислоту, комбинацию нуклеиновых кислот или фармацевтическую композицию, которые соответствуют данному изобретению.

Далее приводятся пояснения к изобретению при помощи примеров выполнения и прилагаемых для большей наглядности чертежей.

Фиг. 1 - схематическое изображение ДНК-связывающего домена относящейся к изобретению CCR5-специфической TALEN-пары ("CCR5-Uco") и ее последовательности-мишени в гене CCR5. Каждая нижняя строка указывает последовательность-мишень для а) левого и б) правого TALEN-плеча, каждая верхняя строка - RVD (repeat variable diresidues) соответствующих TALEN-мономеров в рамочке (аминокислоты указаны однолитерным кодом); в) отрезок (нуклеотиды 135-221 последовательности SEQ ID NO: 11) ДНК CCR5 с комплементарной цепочкой. Последовательности-мишени левого (верхняя) и правого (нижняя, на комплементарной цепочке) плеч CCR5-Uco-TALEN выделены рамками.

Фиг. 2 - сравнение эффективности между относящейся к изобретению CCR5-TALEN ("Uco") и контрольной CCR5-TALEN ("Mco") из предшествующего уровня техники. Проводилось тестирование посредством трансфекции плазмид в CCR5-положительную клетку-репортер линии клеток 293Т. Для всех проверенных конструкций наблюдалась сравнимая эффективность трансфекции (на основе контртрансфекции eGFP). Нокаут CCR5 определялся на шестой день после трансфекции CCR5+ 293Т клетки клона при помощи специфических антител (анти-CD195-APC-Cy-7 антитела) (n=2). Для "mock"-контроля проводилась трансфекция клеток с irrelevantной контрольной плазмидой (pUC), которая не кодирует TALEN.

Фиг. 3 - CCR5 нокаут в первичных Т-лимфоцитах с CCR5-Uco после мРНК-трансфекции. После активации *ex-vivo* приблизительно половина первичных Т-лимфоцитов здорового пробанда экспрессировали CCR5 (= клетки справа от пунктирной линии, см. "нетрансфицированные"): а) после трансфекции CCR5-специфической TALEN (Uco) уменьшилась доля CCR5-положительных клеток с нарастающим количеством трансфицированных мРНК (обратно пропорционально), б) Контрольная трансфекция единичных TALEN-плеч, напротив, не приводит к уменьшению доли CCR5-положительных клеток. Нокаут CCR5 определялся на шестой день после мРНК-трансфекции при помощи специфических антител (anti-CD195-PerCP-Cy5.5 антитела), в) анализ сайта-мишени CCR5-Uco-TALEN показал генетический нокаут в 9 из 17 (>50%) проанализированных первичных Т-лимфоцитов.

Примеры

Была изготовлена и исследована относящаяся к изобретению CCR5-специфическая TALEN (далее "CCR5-Uco"). Относящаяся к изобретению TALEN отличается от описанной ранее TALEN в отношении последовательности-мишени, распознаваемой в гене CCR5 (см. фиг. 1).

По сравнению с кодон-оптимизированной CCR5-TALEN ("Mco"; см. SEQ ID NO: 13, 14), основываясь на опубликованных работах лаборатории проф. Toni Cathomen (Фрайбург) (см. Mussolino C. et al., A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res.*, 2011, 39: 9283-929), относящаяся к изобретению CCR5-Uco-TALEN показывает более высокий уровень индукции CCR5-нокаута после трансфекции плазмиды в линии клеток-репортеров (см. фиг.

2). Последовательность нуклеиновых кислот кодон-оптимизированных TALEN-элементов для использования в клетках человека основывается на публикациях группы ученых во главе с Feng Zhang (Zhang et al., Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription, *Nature Biotechnology*, 2011, 29, 149-153; Sanjana N.E. et al., A TAL Effector Toolbox for Genome Engineering, *Nature Protocols*, 2012, 7: 171-192).

Ниже приведены RVD-последовательности CCR5-Mco-TALEN в соответствии с предшествующим техническим уровнем:

левое плечо (L) = NN NG NN NN NN HD NI NI HD NI NG NN HD NG NN NN NG HD; правое плечо (R) = HD NG NG HD NI NN HD HD NG NG NG NG NN HD NI NN NG NG.

Соответствующие последовательности распознавания на уровне ДНК: L на смысловой нити = GTGGGCAACATGCTGGTC (SEQ ID NO: 15); R на антисмысловой нити = CTTCAGCCTTTTGCAGTT (SEQ ID NO: 16).

Длина спейсера составляет 15 нуклеотидов. Производство TALEN-плазмид также основывается на публикациях группы ученых во главе с Zhang F. или Sanjana N.E. и др. (см. выше).

При помощи мРНК-трансфекции (см. Berdien B. et al, TALEN-mediated editing of endogenous T-cell receptors facilitates efficient reprogramming of T lymphocytes by lentiviral gene transfer, *Gene Therapy*, 2014, doi:10.1038/gt.2014.26) с относящейся к изобретению CCR5-Uco производится нокаут в первичных Т-лимфоцитах. (см. фиг. 3). Последовательности нуклеиновых кислот используемых здесь мРНК указаны в SEQ ID NO: 17 и 18. В SEQ ID NO: 17 указана мРНК левого TALEN-плеча, в SEQ ID NO: 18 указана мРНК правого TALEN-плеча. Нуклеотиды 10-3225 мРНК в SEQ ID NO: 17 и 18 кодируют каждое из TALEN-плеч (мономеров), аминокислотные последовательности которых указаны в SEQ ID NO: 3, 4.

Трансфецированная мРНК была создана при помощи T7-промотора после AvrII-линеаризации вектора. Так как сайт расщепления AvrII 563bp находится за стоп-кодоном, указанная последовательность больше чем открытая рамка считывания. После линеаризации соответствующие Uco-TALEN-ДНК используются в качестве матрицы для синтеза мРНК при помощи T7 mScript™ Standard mRNA Production System от Cellscript (Madison, WI 53713 USA). мРНК была оснащена 5'Cap-концом и Poly-A-концом в соответствии с руководством производителя. Трансфекция мРНК производится путем электропорации первичных Т-клеток в течение 10 мс при 300 V. При помощи Mco-CCR5-TALEN, однако, не удалось реализовать CCR5-нокаут в первичных Т-клетках или линиях Z-клеток (в то время как нокаут рецепторов Т-клеток был возможным, см. Berdien et al. 2014, см. выше). Только с относящейся к изобретению CCR5-Uco-TALEN, при помощи мРНК-переноса, удалось выключить значительную часть >50% CCR5-аллелей. Это позволяет предположить, что только достаточно активные TALEN способны выполнять свою функцию в первичных Т-клетках после мРНК-трансфекции. Таким образом, относящаяся к изобретению CCR5-TALEN особенно хорошо подходит для использования с помощью мРНК-трансфекции, что делает ее крайне привлекательной с точки зрения клинического применения.

Обзор последовательностей.

SEQ ID NO:	Тип	Описание
01	ДНК	Последовательность-мишень TALEN CCR5-Uco L
02	ДНК	Последовательность-мишень TALEN CCR5-Uco R
03	БЕЛОК	TALEN CCR5-Uco L
04	БЕЛОК	TALEN CCR5-Uco R
05	БЕЛОК	Последовательность повтора (Консенсус)
06	БЕЛОК	Последовательность повтора
07	БЕЛОК	Последовательность повтора
08	БЕЛОК	Последовательность повтора
09	БЕЛОК	Последовательность повтора
10	БЕЛОК	Последовательность повтора
11	ДНК	hCCR5
12	БЕЛОК	FokI-домен расщепления
13	БЕЛОК	TALEN CCR5-Mco L
14	БЕЛОК	TALEN CCR5-Mco R
15	ДНК	Последовательность-мишень TALEN CCR5-Mco L
16	ДНК	Последовательность-мишень TALEN CCR5-Mco R
17	мРНК	мРНК CCR5-Uco L
18	мРНК	мРНК CCR5-Uco R

Перечень последовательностей - произвольный текст

TALEN repeat = TALEN-единица повтора.

Any amino acid, or E, Q, D or A = любая аминокислота, или E, Q, D или A.

Any amino acid, or A or V = любая аминокислота, или A или V.

Any amino acid, or A or D = любая аминокислота, или A или D.

Repeat variable diresidue (RVD) = вариабельные последовательности из двух аминокислотных остатков (RVD).

FokI cleavage domain = FokI - домен расщепления.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Набор из двух мономеров TAL-эффекторной нуклеазы для инактивации ВИЧ корцептора CCR5 человека, характеризующийся тем, что каждый мономер TAL-эффекторной нуклеазы включает в себя домен с эндонуклеазной активностью II типа и ДНК-связывающий домен TAL-эффектора, содержащий 19 единиц повтора, каждая из которых имеет пару вариабельных аминокислот RVD (Repeat Variable Diresidue), при этом:

а) ДНК-связывающий домен TAL-эффектора одного мономера TAL-эффекторной нуклеазы связывается с последовательностью-мишенью GCTGGTCATCCTCATCCTG (SEQ ID NO: 1) и содержит следующую последовательность RVD в составе следующих друг за другом единиц повтора: NH, HD, NG, NH, NH, NG, HD, NI, NG, HD, HD, NG, HD, NI, NG, HD, HD, NG, NN, и

б) ДНК-связывающий домен TAL-эффектора другого мономера TAL-эффекторной нуклеазы связывается с последовательностью-мишенью AGATGTCAGTCATGCTCTT (SEQ ID NO: 2) и содержит следующую последовательность RVD в составе следующих друг за другом единиц повтора: NI, NN, NI, NG, NN, NG, HD, NI, NH, NG, HD, NI, NG, NH, HD, NG, HD, NG, NG;

при этом домен с эндонуклеазной активностью в каждом мономере TAL-эффекторной нуклеазы является С-концевым по отношению к ДНК-связывающему домену TAL-эффектора и каждая единица повтора содержит от 33 до 35 аминокислот;

при этом указанные мономеры, связываясь с двуцепочечной ДНК в локусе гена ВИЧ корцептора CCR5 человека, образуют TAL-эффекторную нуклеазу, способную образовывать двуцепочечный разрыв в указанной ДНК.

2. Набор по п.1, в котором домен с эндонуклеазной активностью мономеров TAL-эффекторных нуклеаз является ДНК-расщепляющим доменом эндонуклеазы FokI.

3. Набор по любому из пп.1, 2, в котором первый мономер TAL-эффекторной нуклеазы содержит аминокислотную последовательность, соответствующую SEQ ID NO: 3, а второй мономер TAL-эффекторной нуклеазы содержит аминокислотную последовательность, соответствующую SEQ ID NO: 4.

4. Нуклеиновая кислота для инактивации ВИЧ корцептора CCR5 человека, включающая в себя:

а) первую нуклеотидную последовательность, кодирующую первый мономер TAL-эффекторной нуклеазы, который содержит домен с эндонуклеазной активностью II типа и ДНК-связывающий домен TAL-эффектора, содержащий 19 единиц повтора, каждая из которых имеет пару вариабельных аминокислот RVD, и при этом ДНК-связывающий домен TAL-эффектора связывается с последовательностью-мишенью GCTGGTCATCCTCATCCTG (SEQ ID NO: 1) и содержит следующую последовательность RVD в составе следующих друг за другом единиц повтора: NH, HD, NG, NH, NH, NG, HD, NI, NG, HD, HD, NG, HD, NI, NG, HD, HD, NG, NN, и

б) вторую нуклеотидную последовательность, кодирующую второй мономер TAL-эффекторной нуклеазы, который содержит домен с эндонуклеазной активностью II типа и ДНК-связывающий домен TAL-эффектора, содержащий 19 единиц повтора, каждая из которых имеет пару вариабельных аминокислот RVD, и при этом ДНК-связывающий домен TAL-эффектора связывается с последовательностью-мишенью AGATGTCAGTCATGCTCTT (SEQ ID NO: 2) и содержит следующую последовательность RVD в составе следующих друг за другом единиц повтора: NI, NN, NI, NG, NN, NG, HD, NI, NH, NG, HD, NI, NG, NH, HD, NG, HD, NG, NG.

5. Вектор, включающий в себя нуклеиновую кислоту по п.4.

6. Комбинация нуклеиновых кислот для инактивации ВИЧ корцептора CCR5 человека, включающая в себя:

а) первую нуклеиновую кислоту, кодирующую первый мономер TAL-эффекторной нуклеазы, который содержит домен с эндонуклеазной активностью II типа и ДНК-связывающий домен TAL-эффектора, содержащий 19 единиц повтора, каждая из которых имеет пару вариабельных аминокислот RVD, и при этом ДНК-связывающий домен TAL-эффектора связывается с последовательностью-мишенью GCTGGTCATCCTCATCCTG (SEQ ID NO: 1) и содержит следующую последовательность RVD в составе следующих друг за другом единиц повтора: NH, HD, NG, NH, NH, NG, HD, NI, NG, HD, HD, NG, HD, NI, NG, HD, HD, NG, NN, и

б) вторую нуклеиновую кислоту, кодирующую второй мономер TAL-эффекторной нуклеазы, который содержит домен с эндонуклеазной активностью II типа и ДНК-связывающий домен TAL-эффектора,

содержащий 19 единиц повтора, каждая из которых имеет пару переменных аминокислот RVD, и при этом ДНК-связывающий домен TAL-эффектора связывается с последовательностью-мишенью AGATGTCAGTCATGCTCTT (SEQ ID NO: 2) и содержит следующую последовательность RVD в составе следующих друг за другом единиц повтора: NI, NN, NI, NG, NN, NG, HD, NI, NH, NG, HD, NI, NG, NH, HD, NG, HD, NG, NG.

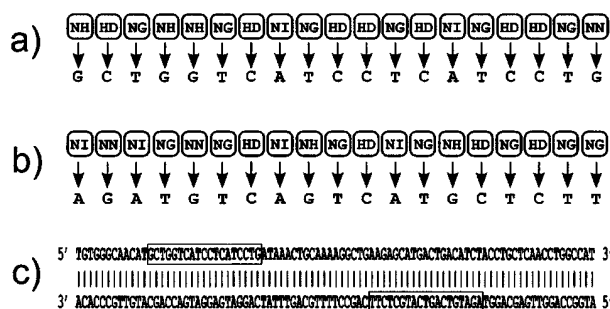
7. Комбинация нуклеиновых кислот по п.6, в которой первая нуклеиновая кислота является первой мРНК и вторая нуклеиновая кислота является второй мРНК.

8. Комбинация нуклеиновых кислот по п.7, в которой первая нуклеиновая кислота включает последовательность, соответствующую SEQ ID NO: 17, а вторая нуклеиновая кислота включает последовательность, соответствующую SEQ ID NO: 18.

9. Изолированная клетка-хозяин, включающая в себя нуклеиновую кислоту по п.4, комбинацию нуклеиновых кислот по пп.6-8 или вектор по п.5 при условии, что клетка-хозяин не является зародышевой клеткой человека, эмбриональной зародышевой клеткой человека или эмбриональной стволовой клеткой, процесс получения которых сопряжен с разрушением эмбриона человека.

10. Фармацевтическая композиция для инактивации ВИЧ корцептора CCR5 человека, включающая в себя нуклеиновую кислоту по п.4, комбинацию нуклеиновых кислот по пп.6-8 или вектор по п.5.

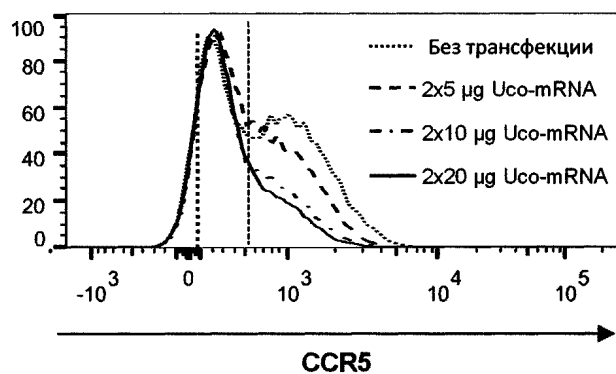
11. Лекарственное средство для инактивации ВИЧ корцептора CCR5 человека, включающее в себя нуклеиновую кислоту по п.4, комбинацию нуклеиновых кислот по любому из пп.6-8, вектор по п.5 или фармацевтическую композицию по п.10.



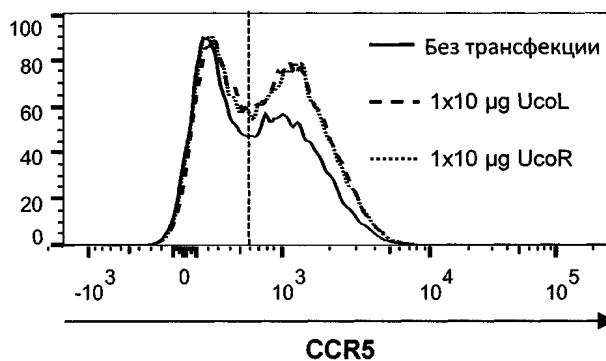
Фиг. 1



Фиг. 2



Фиг. 3а



Фиг. 3b

```

WT TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#1 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#2 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#3 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#4 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#5 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#6 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#7 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#8 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#9 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAAA-----GGCTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#10 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAA-----GGCTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#11 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAAA-----CTGAAGAGCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#12 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAAAC-----GCATGACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#13 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATAA-----ACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#14 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATG-----ACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#15 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTGATG-----ACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG
#16 TTGGTTTGTGGGCAACATGCTGGTCACTCTCATCTCA-----GATAAACT-----GCCATCTCTGAOCTG
#17 TTGGTTTGTGGGCAACANNCTGGTCTCATCTG-----ACTGACATCTACCTGCTCAACTGGCCATCTCTGAOCTG

```

Фиг. 3c

