

(19)



**Евразийское
патентное
ведомство**

(11) **040859**(13) **B1**

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОМУ ПАТЕНТУ

(45) Дата публикации и выдачи патента
2022.08.08

(21) Номер заявки
201992795

(22) Дата подачи заявки
2018.06.08

(51) Int. Cl. **C12N 15/09** (2006.01)
A01H 1/00 (2006.01)
A01K 67/027 (2006.01)
C12N 5/10 (2006.01)

(54) СПОСОБ ПОЛУЧЕНИЯ ЭУКАРИОТИЧЕСКИХ КЛЕТОК С ОТРЕДАКТИРОВАННОЙ ДНК И НАБОР, ИСПОЛЬЗУЕМЫЙ В ЭТОМ СПОСОБЕ

(31) 2017-113747

(32) 2017.06.08

(33) JP

(43) 2020.04.21

(86) PCT/JP2018/022066

(87) WO 2018/225858 2018.12.13

(71)(73) Заявитель и патентовладелец:
ОСАКА ЮНИВЕРСИТИ (JP)

(72) Изобретатель:
**Масимо Томодзи, Такеда Дзундзи,
Морисака Хироюки, Йосими Кадзуго
(JP)**

(74) Представитель:
Медведев В.Н. (RU)

(56) WO-A2-2017066497

JORISAKA, H. et al., "Genome editing in mammalian cells by cascade and Cas3", J. Invest. dermatol., May 2017, vol. 137, issue 5, supplement 1, pp. S84, 490, entire text

SUZUKI, Keiichiro et al., "In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration", Nature, 2016, vol. 540, pp. 144-149, extended data

JP-A-2015503535

WO-A1-2017043573

JP-A-2017512481

MORISAKA, Hiroyuki et al., "Genome editing in mammalian cells by Cascade and Cas3", 02 November 2017, p. 232, P07-20 [04-12] entire text, non-official translation (Programs and abstracts of 42th annual academic conference and general conference of the Japanese Society for Investigate Dermatology)

WO-A1-2017219033

CN-A1-107557378

CN-A-107557373

(57) Система CRISPR-Cas3 была успешно введена в эукариотическую клетку.

040859

B1

040859

B1

Область техники

Настоящее изобретение относится к способу получения эукариотических клеток животных и растений с отредактированной ДНК и к набору, используемому в этом способе.

Уровень техники

Бактерии и археобактерии имеют адаптивные иммунные механизмы, которые специфически распознают и удаляют микроорганизмы, такие как чужеродные фаги, которые могут проникать из окружающей среды. Эти системы, называемые системами CRISPR-Cas, сначала вводят геномную информацию о чужеродных организмах в собственный геном (адаптация). Затем при повторном проникновении тех же самых чужеродных организмов эти системы расщепляют и удаляют чужеродные геномы с использованием дополнительной информации, поступающей в собственный геном и в геномную последовательность (интерференция).

Недавно были разработаны методы редактирования генома (редактирования ДНК) с использованием вышеописанных систем CRISPR-Cas в качестве "средств для редактирования ДНК" (NPL 1).

Системы CRISPR-Cas можно приблизительно подразделить на "класс 1" и "класс 2", где эффекторы, участвующие в процессе расщепления ДНК, состоят из множества Cas и одного Cas соответственно. Из этих систем широко известны системы CRISPR-Cas класса 1, а именно системы "типа I", участвующие в образовании комплексов Cas3 и Cascade (обозначаемых комплексами Cascade-cr-PHK, и это относится к нижеследующим комплексам). Системы CRISPR-Cas класса 2 широко известны как системы "типа II", включающие Cas9 (далее системы CRISPR-Cas "класса 1 типа I" и "класса 2 типа II" могут просто называться системами "типа I" и "типа II" соответственно). Кроме того, в стандартных методах редактирования ДНК широко используются системы CRISPR-Cas класса 2, включающие Cas9 (эти системы далее могут называться "системами CRISPR-Cas9"). Так, например, в NPL 1 описана система CRISPR-Cas класса 2, которая расщепляет ДНК посредством Cas9.

С другой стороны, что касается систем CRISPR-Cas класса 1, которые расщепляют ДНК посредством комплексов Cas3 и Cascade (которые далее могут также называться "системами CRISPR-Cas3"), то несмотря на ряд попыток, какого-либо успешного редактирования генома в эукариотических клетках не достигалось. Так, например, в NPL 1 и в NPL 3 сообщалось, что просто использование системы CRISPR-Cas3 может полностью разрушать ДНК-мишень в бесклеточной системе и селективно удалять специфические штаммы *E. coli*. Однако это не означает успешное редактирование генома, и вообще не было продемонстрировано в эукариотических клетках. Кроме того, в PTL 1 предлагается осуществлять редактирование генома с использованием нуклеазы FokI вместо Cas3 в эукариотических клетках (пример 7, фиг. 7 и 11), поскольку системы CRISPR-Cas3 разрушают ДНК-мишень в *E. coli* посредством геликазной активности и экзонуклеазной активности Cas3 (пример 5 и фиг. 6). Кроме того, в PTL 2 предлагается вносить делецию Cas3 и проводить повторную репрессию программируемого гена благодаря использованию инактивированного Cas3 (Cas3' и Cas3'') (например, пример 15 и п.4(е)), поскольку системы CRISPR-Cas3 разрушают ДНК-мишень в *E. coli* (фиг. 4).

Список цитируемой литературы

Патентная литература.

PTL 1. Опубликованный перевод японской международной заявки PCT No. 2015-503535.

PTL 2. Опубликованный перевод японской международной заявки PCT No. 2017-512481.

Непатентная литература.

[NPL 1] Jinek M et al. (2012) A Programmable Dual-RNA Guided DNA Endonuclease in Adaptive Bacterial Immunity, *Science*, Vol. 337 (Issue 6096), pp. 816–821

[NPL 2] Mulepati S & Bailey S (2013) In Vitro Reconstitution of an Escherichia coli RNA-guided Immune System Reveals Unidirectional, ATP-dependent Degradation of DNA Target, *Journal of Biological Chemistry*, Vol. 288 (No. 31), pp. 22184–22192

[NPL 3] Ahmed A. Goma et al. (2014) Programmable Removal of Bacterial Strains by Use of Genome Targeting CRISPR-Cas Systems, *mbio.asm.org*, Volume 5, Issue 1, e00928–13.

Сущность изобретения

Техническая задача.

Настоящее изобретение было разработано с учетом вышеупомянутых факторов, и его целью является создание системы CRISPR-Cas3 в эукариотических клетках.

Решение задачи.

Авторами настоящего изобретения были проведены серьезные исследования для достижения вышеуказанной цели и, наконец, был достигнут успех в создании системы CRISPR-Cas3 в эукариотических клетках. Наиболее широко используемая система CRISPR-Cas9 позволяет успешно осуществлять редактирование генома в различных эукариотических клетках, но в этой системе в качестве cr-PHK обычно используют зрелую cr-PHK. Однако было неожиданно обнаружено, что в системах CRISPR-Cas3, редактирование генома в эукариотических клетках было затруднено в случае использования зрелой cr-PHK, и что эффективное редактирование генома было возможно только при использовании пре-cr-PHK, которая

обычно не используется как составной элемент системы. То есть для того чтобы сделать системы CRISPR-Cas3 функциональными в эукариотических клетках, было важно понять механизм расщепления cr-PHK белками, составляющими Cascade. Системы CRISPR-Cas3, использующие эту pre-cr-PHK, широко применялись не только к системе типа I-E, но также к системам типа I-F и типа I-G. Кроме того, присоединение сигнала локализации в ядре, а в частности, двухкомпонентного сигнала локализации в ядре, к Cas3, позволило дополнительно повысить эффективность редактирования генома для систем CRISPR-Cas3 в эукариотических клетках. Кроме того, авторами настоящего изобретения было обнаружено, что системы CRISPR-Cas3, в отличие от систем CRISPR-Cas9, могут давать крупную делецию в области, содержащей последовательность PAM, или в вышерасположенной области. На основании этих результатов было создано настоящее изобретение.

В частности, настоящее изобретение относится к системе CRISPR-Cas3 в эукариотических клетках, а более конкретно к нижеследующим предметам изобретения.

1. Способ получения эукариотической клетки с отредактированной ДНК, включающий введение системы CRISPR-Cas3 в эукариотическую клетку, где система CRISPR-Cas3 включает следующие (A)-(C):

A) белок Cas3, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид,

B) белок Cascade, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид, и

C) cr-PHK, полинуклеотид, кодирующий cr-PHK, или экспрессионный вектор, содержащий этот полинуклеотид.

2. Способ получения животного (за исключением человека) или растения с отредактированной ДНК, включающий: введение системы CRISPR-Cas3 животному (за исключением человека) или в растение, где система CRISPR-Cas3 включает следующее (A)-(C):

A) белок Cas3, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид,

B) белок Cascade, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид, и

C) cr-PHK, полинуклеотид, кодирующий cr-PHK, или экспрессионный вектор, содержащий этот полинуклеотид.

3. Способ в соответствии с п.1 или 2, дополнительно включающий расщепление cr-PHK белком, составляющим белок Cascade, после введения системы CRISPR-Cas3 в эукариотическую клетку.

4. Способ в соответствии с п.1 или 2, где cr-PHK представляет собой pre-cr-PHK.

5. Способ в соответствии с любым из пп.1-4, где сигнал локализации в ядре присоединяют к белку Cas3 и/или к белку Cascade.

6. Способ в соответствии с п.5, где сигнал локализации в ядре представляет собой двухкомпонентный сигнал локализации в ядре.

7. Набор для его применения в способе в соответствии с любым из пп.1-6, где набор включает следующие A) и B):

A) белок Cas3, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид,

B) белок Cascade, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид.

8. Набор в соответствии с п.7, дополнительно содержащий cr-PHK, полинуклеотид, кодирующий cr-PHK или экспрессионный вектор, содержащий этот полинуклеотид.

9. Набор в соответствии с п.8, где cr-PHK представляет собой pre-cr-PHK.

10. Набор в соответствии с любым из пп.7-9, где сигнал локализации в ядре присоединяют к белку Cas3 и/или к белку Cascade.

11. Набор в соответствии с п.10, где сигнал локализации в ядре представляет собой двухкомпонентный сигнал локализации в ядре.

Следует отметить, что в настоящем описании термин "полинуклеотид" означает полимер, состоящий из нуклеотидов, и этот термин используется здесь как синоним термина "ген", "нуклеиновая кислота" или "молекула нуклеиновой кислоты". Полинуклеотид может также присутствовать в форме ДНК (например, к ДНК или геномной ДНК) или в форме РНК (например, мРНК). Используемый здесь термин "белок" также является синонимом термина "пептид" или "полипептид".

Преимущественные эффекты изобретения

Использование системы CRISPR-Cas3 согласно изобретению позволяет осуществлять редактирование ДНК в эукариотических клетках.

Краткое описание фигур

На фиг. 1 представлены результаты анализа SSA для измерения активности расщепления экзогенной ДНК.

На фиг. 2 схематически представлена диаграмма, иллюстрирующая положение последовательно-

сти-мишени в гене CCR5.

На фиг. 3A представлена диаграмма, иллюстрирующая ген CCR5 (клон 1), в котором часть последовательности нуклеиновой кислоты была делетирована с использованием системы CRISPR-Cas3.

На фиг. 3B представлена диаграмма, иллюстрирующая ген CCR5 (клон 2), в котором часть последовательности нуклеиновой кислоты была делетирована с использованием системы CRISPR-Cas3.

На фиг. 3C представлена диаграмма, иллюстрирующая ген CCR5 (клон 3), в котором часть последовательности нуклеиновой кислоты была делетирована с использованием системы CRISPR-Cas3.

На фиг. 3D представлена диаграмма, иллюстрирующая ген CCR5 (клон 4), в котором часть последовательности нуклеиновой кислоты была делетирована с использованием системы CRISPR-Cas3.

На фиг. 4(a) схематически представлена диаграмма, иллюстрирующая структуру плазмиды Cascade.

На фиг. 4(b) схематически представлена диаграмма, иллюстрирующая структуру плазмиды Cas3.

На фиг. 4(c) схематически представлена диаграмма, иллюстрирующая структуру плазмиды pre-cr-PHK.

На фиг. 4(d) схематически представлена диаграмма, иллюстрирующая структуру репортерного вектора (включающего последовательность-мишень).

На фиг. 5 схематически представлена диаграмма, иллюстрирующая положение последовательности-мишени в гене EMX1.

На фиг. 6A представлена диаграмма, иллюстрирующая ген EMX1 (клон 1), делетированный в части последовательности нуклеиновой кислоты посредством системы CRISPR-Cas3.

На фиг. 6B представлена диаграмма, иллюстрирующая ген EMX1 (клон 2), делетированный в части последовательности нуклеиновой кислоты посредством системы CRISPR-Cas3.

На фиг. 7 схематически представлена диаграмма, иллюстрирующая структуру плазмиды Cas3/Cascade, присоединенной к bpNLS.

На фиг. 8 схематически представлена диаграмма, иллюстрирующая структуру плазмиды Cascade (2A).

На фиг. 9 представлены результаты анализа SSA для измерения активности расщепления экзогенной ДНК.

На фиг. 10A представлена диаграмма, иллюстрирующая структуры pre-cr-PHK (LRSR и RSR) и зрелой cr-PHK, используемой в примерах. На этой фигуре 5'-фланкирующая последовательность (фланкирующая последовательность Cas5) подчеркнута, а 3'-фланкирующая последовательность (фланкирующая последовательность Cas3) показана двойным подчеркиванием.

На фиг. 10B представлена диаграмма, иллюстрирующая результаты анализа SSA с использованием pre-cr-PHK (LRSR и RSR) и зрелой cr-PHK.

На фиг. 11 показаны результаты анализа SSA с использованием одного NLS или двух NLS (bpNLS) в плазмиде для экспрессии гена Cas3/Cascade.

На фиг. 12 представлена диаграмма, иллюстрирующая влияние последовательности PAM на активность расщепления ДНК системой CRISPR-Cas3.

На фиг. 13 представлена диаграмма, иллюстрирующая влияние одного ошибочного спаривания спейсера на активность расщепления ДНК системой CRISPR-Cas3.

На фиг. 14 представлена диаграмма, иллюстрирующая эффекты мутации Cas3 в нуклеазном домене HD (H74A), в мотиве 1 домена геликазы SF2 (K320A) и в мотиве 3 (S483/T485A).

На фиг. 15 показано сравнение активности расщепления ДНК системами типа I-E, типа I-F и типа I-G CRISPR-Cas3.

На фиг. 16 представлена диаграмма, иллюстрирующая величину делеции, вносимой системой CRISPR-Cas3, детектируемой путем секвенирования образца ПНП-продукта для клонирования ТА.

На фиг. 17 представлена диаграмма, иллюстрирующая положение делеции, вносимой системой CRISPR-Cas3, детектируемой путем крупномасштабной обработки в процессе секвенирования клона ТА (n=49).

На фиг. 18A представлена диаграмма, где показано число для каждого размера делеции, детектируемой системой CRISPR-Cas3 с использованием последовательности для захвата на основе микромасивов размером 1000 т.п.о. или более, вокруг локуса-мишени EMX1.

На фиг. 18B представлена диаграмма, где показано число для каждого размера делеции, детектируемой системой CRISPR-Cas3 с использованием последовательности для захвата на основе микромасивов размером 1000 т.п.о. или более, вокруг локуса-мишени CCR5.

Описание вариантов осуществления изобретения

1. Способ получения эукариотических клеток, животных и растений с отредактированной ДНК.

Способ согласно изобретению включает введение системы CRISPR-Cas3 в эукариотическую клетку, где система CRISPR-Cas3 включает следующие А)-С).

А) белок Cas3, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид,

В) белок Cascade, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид, и

С) сг-ПНК, полинуклеотид, кодирующий сг-ПНК, или экспрессионный вектор, содержащий этот полинуклеотид.

Системы CRISPR-Cas класса 1 подразделяются на тип I и тип III, и в зависимости от типа белков, составляющих Cascade (далее называемых просто "Cascade" или "белками Cascade"), тип I дополнительно подразделяется на шесть типов: тип I-A, тип I-B, тип I-C, тип I-D, тип I-E и тип I-F, а также типа I-G, подтип типа I-B (см., например, van der Oost J. et al. (2014) Unravelling the structural and mechanistic basis of CRISPR-Cas systems, *Nature Reviews Microbiology*, Vol. 12 (No. 7), pp. 479-492 и Jackson R.N. et al. (2014) Fitting CRISPR-associated Cas3 into the Helicase Family Tree, *Current Opinion in Structural Biology*, Vol. 24, pp. 106-114).

Системы CRISPR-Cas типа I выполняют функцию расщепления ДНК посредством совместного действия Cas3 (белка, обладающего нуклеазной активностью и геликазной активностью), Cascade и сг-ПНК. В настоящем изобретении они упоминаются как "система CRISPR-Cas3", поскольку Cas3 используется в качестве нуклеазы.

Использование системы CRISPR-Cas3 согласно изобретению позволяет получить, например, следующие преимущества.

Во-первых, сг-ПНК, используемая в системе CRISPR-Cas3, обычно распознает последовательность-мишень размером 32-37 оснований (Ming Li et al., *Nucleic Acids Res.* 2017 May 5; 45(8): 4642-4654). С другой стороны, сг-ПНК, используемая в системе CRISPR-Cas9, обычно распознает последовательность-мишень размером 18-24 оснований. Следовательно, считается, что система CRISPR-Cas3 может распознавать последовательности-мишени более точно, чем система CRISPR-Cas9.

Кроме того, последовательность PAM системы класса 2 типа II системы CRISPR-Cas9 представляет собой "NGG (N означает произвольно выбранное основание)", смежную с 3'-стороной последовательности-мишени. Кроме того, последовательность PAM системы класса 2 типа V системы CRISPR-Cpf1 представляет собой "AA", смежную с 5'-стороной последовательности-мишени. С другой стороны, последовательность PAM системы CRISPR-Cas3 согласно изобретению представляет собой "AAG", смежную с 5'-стороной последовательности-мишени или последовательности нуклеиновой кислоты, подобной этой последовательности (например, "AGG", "GAG", "TAC", "ATG", "TAG" и т.п.) (фиг. 12). Таким образом, считается, что с использованием системы CRISPR-Cas3 согласно изобретению области, которые не могут распознаваться обычными способами, могут быть подвергнуты редактированию ДНК.

Кроме того, в отличие от вышеупомянутых систем CRISPR-Cas класса 2, система CRISPR-Cas3 расщепляет ДНК во многих положениях. Следовательно, использование системы CRISPR-Cas3 согласно изобретению позволяет получить делеционные мутации широкого спектра в интервале от ста до нескольких тысяч оснований и возможно даже более (фиг. 3, 6 и 16-18). Считается, что эту функцию можно использовать для нокаута длинной геномной области или для нокина длинной ДНК. При осуществлении нокина обычно используется донорная ДНК, и эта донорная ДНК также представляет собой молекулу, составляющую систему CRISPR-Cas3 согласно изобретению.

Следует отметить, что если в описании изобретения приводится просто обозначение "Cas3", то это означает "белок Cas3". То же самое относится и к белкам Cascade.

Система CRISPR-Cas3 согласно изобретению включает все шесть подтипов типа I. То есть хотя белки, составляющие систему CRISPR-Cas3, могут незначительно отличаться по структуре и т.п. в зависимости от подтипа (например, белки, составляющие Cascade, отличаются друг от друга), однако настоящее изобретение включает все эти белки. Действительно, в настоящем примере было обнаружено, что редактирование генома возможно не только для систем типа I-E, но также для систем типа I-G и типа I-F (фиг. 15).

Система CRISPR-Cas3 типа I-E, которая является самой распространенной среди систем CRISPR-Cas3 типа I, расщепляет ДНК, если сг-ПНК взаимодействует с Cas3 и Cascade (Cse1 (Cas8), Cse2 (Cas11), Cas5, Cas6 и Cas7).

Система типа I-A имеет составные элементы Cascade Cas8a1, Csa5 (Cas11), Cas5, Cas6 и Cas7, тип I-B имеет составные элементы Cascade Cas8b1, Cas5, Cas6 и Cas7, тип I-C имеет составные элементы Cascade Cas8c, Cas5 и Cas7, тип I-D имеет составные элементы Cascade Cas10d, Csc1 (Cas5), Cas6 и Csc2 (Cas7), тип I-F имеет составные элементы Cascade Csy1 (Cas8f), Csy2 (Cas5), Cas6 и Csy3 (Cas7), а система типа I-G имеет составные элементы Cascade Cst1 (Cas8a1), Cas5, Cas6 и Cst2 (Cas7). В настоящем изобретении Cas3 и Cascade называется общим термином "группа белков Cas".

Далее система CRISPR-Cas3 типа I-E описана в качестве репрезентативного примера. Для систем CRISPR-Cas3 других типов Cascade, составляющий эти системы, может быть интерпретирован соответствующим образом.

Группа белков Cas.

В системе CRISPR-Cas3 согласно изобретению группа белков Cas может быть введена в эукариотические клетки в форме белка, в форме полинуклеотида, кодирующего белок, или в форме экспрессионного вектора, содержащего полинуклеотид. Если группа белков Cas вводится в эукариотические клетки в форме белка, то можно получить соответствующее количество и т.п. каждого белка, что является превосходным с точки зрения обработки. Кроме того, принимая во внимание, например, эффективность

расщепления в клетках, можно также сначала получить комплекс группы белков Cas, а затем ввести его в эукариотические клетки.

В настоящем изобретении предпочтительно присоединить сигнал локализации в ядре к группе белков Cas. Сигнал локализации в ядре может быть присоединен к N-концевой стороне и/или к С-концевой стороне группы белков Cas (к 5'-концевой стороне и/или к 3'-концевой стороне полинуклеотида, кодирующего каждую группу белков Cas). Таким образом, присоединение сигнала локализации в ядре к группе белков Cas стимулирует локализацию в ядре клетки, что, в результате, позволяет эффективно осуществлять редактирование ДНК.

Вышеописанный сигнал локализации в ядре представляет собой пептидную последовательность, состоящую из нескольких основных аминокислот до нескольких десятков основных аминокислот, и такая последовательность не имеет конкретных ограничений при условии, что белки будут перенесены в ядро. Конкретный пример такого сигнала локализации в ядре описан, например, в литературе [Wu J. et al. (2009) The Intracellular Mobility of Nuclear Import Receptors and NLS Cargoes, Biophysical journal, Vol. 96 (Issue 9), pp. 3840-3849]. В настоящем изобретении может быть использован любой сигнал локализации в ядре, обычно применяемый в данной области техники.

Сигнал локализации в ядре может представлять собой, например, PKKKRKV (SEQ ID NO: 52) (кодируется последовательностью нуклеиновой кислоты CCCAAGAAGAAGCGGAAGGTG (SEQ ID NO: 53)).

Если используется вышеуказанный сигнал локализации в ядре, то предпочтительно, чтобы он представлял собой, например, полинуклеотид, состоящий из последовательности нуклеиновой кислоты с SEQ ID NO: 53 на 5'-концевой стороне полинуклеотида, кодирующего каждую группу белков Cas. Кроме того, сигналом локализации в ядре может быть, например, KRTADGSEFESPKKKRKVE (SEQ ID NO: 54) (кодируемый последовательностью нуклеиновой кислоты AAGCGGACTGCTGATGGCAGTGAATTTGAGTCCCCAAAGAAGAAGAGAAAGGTGGA (SEQ ID NO: 55)).

Если используется вышеуказанный сигнал локализации в ядре, то предпочтительно, чтобы он представлял собой, например, полинуклеотиды, состоящие из последовательностей нуклеиновой кислоты с SEQ ID NO: 55, на обеих сторонах полинуклеотида, кодирующего каждую группу белков Cas (в частности, для использования "двухкомпонентного сигнала локализации в ядре (bpNLS)").

Такие модификации играют важную роль в экспрессии системы CRISPR-Cas3 согласно изобретению и в ее эффективном функционировании в эукариотических клетках вместе с использованием пре-сгРНК, описанных ниже.

Один предпочтительный вариант группы белков Cas, используемой в настоящем изобретении, представлен ниже.

Cas3; белок, кодируемый полинуклеотидом, состоящим из последовательности нуклеиновой кислоты, представленной SEQ ID NO: 1 или SEQ ID NO: 7;

Cse1 (Cas8); белок, кодируемый полинуклеотидом, состоящим из последовательности нуклеиновой кислоты, представленной SEQ ID NO: 2 или SEQ ID NO: 8;

Cse2 (Cas11); белок, кодируемый полинуклеотидом, состоящим из последовательности нуклеиновой кислоты, представленной SEQ ID NO: 3 или SEQ ID NO: 9;

Cas5; белок, кодируемый полинуклеотидом, состоящим из последовательности нуклеиновой кислоты, представленной SEQ ID NO: 4 или SEQ ID NO: 10;

Cas6; белок, кодируемый полинуклеотидом, состоящим из последовательности нуклеиновой кислоты, представленной SEQ ID NO: 5 или SEQ ID NO: 11;

Cas7; белок, кодируемый полинуклеотидом, состоящим из последовательности нуклеиновой кислоты, представленной SEQ ID NO: 6 или SEQ ID NO: 12.

Вышеуказанная группа белков Cas представляет собой (1) белок, полученный путем присоединения PKKKRKV (SEQ ID NO: 52) в качестве сигнала локализации в ядре к N-концам Cas3, Cse1 (Cas8), Cse2 (Cas11), Cas5, Cas6 и Cas7 E.coli дикого типа, или (2) белок, полученный путем присоединения KRTADGSEFESPKKKRKVE (SEQ ID NO: 54) в качестве сигнала локализации в ядре к N-концам и С-концам Cas3, Cse1 (Cas8), Cse2 (Cas11), Cas5, Cas6 и Cas7 E. coli дикого типа. В случае белков, имеющих такие аминокислотные последовательности, вышеуказанная группа белков Cas может быть перенесена в ядро эукариотической клетки. Указанная выше группа белков Cas, перенесенная в ядро таким образом, расщепляет ДНК-мишень. Кроме того, можно редактировать ДНК-мишень даже в области ДНК, имеющей стерическую структуру, которая, как считается, трудно поддается расщеплению системой CRISPR-Cas9 (гетерохроматин и т.п.).

Другой вариант белков в группе белков Cas, используемых в настоящем изобретении, представляет собой белок, кодируемый последовательностью нуклеиновой кислоты, которая на 90% или более идентична последовательности нуклеиновой кислоты вышеуказанной группы белков Cas. Другой вариант белков в группе белков Cas, используемых в настоящем изобретении, является белок, кодируемый полинуклеотидом, который гибридизуется с полинуклеотидом, состоящим из последовательности нуклеино-

вой кислоты, комплементарной последовательности нуклеиновой кислоты группы белка Cas, описанной выше, в жестких условиях. Каждый из вышеуказанных белков обладает активностью расщепления ДНК при образовании комплекса с другим белком, составляющим группу белков Cas. Значения таких терминов, как "идентичность последовательностей" и "условия жесткости", описаны ниже.

Полинуклеотид, кодирующий группу белков Cas.

Полинуклеотиды, кодирующие белки дикого типа, составляющие систему CRISPR-Cas типа I-E, включают полинуклеотиды, модифицированные для эффективной экспрессии в эукариотических клетках. То есть можно использовать полинуклеотид, который кодирует группу белков Cas, и который был модифицирован. Одним предпочтительным вариантом модификации полинуклеотидов является модификация последовательности нуклеиновой кислоты, подходящей для экспрессии в эукариотических клетках, например, оптимизация кодона, экспрессируемого в эукариотических клетках.

Один предпочтительный вариант полинуклеотидов, кодирующих группу белков Cas, используемых в настоящем изобретении, представляет собой следующие варианты:

Cas3; полинуклеотид, состоящий из последовательности нуклеиновой кислоты, представленной SEQ ID NO: 1 или SEQ ID NO: 7;

Cse1 (Cas8); полинуклеотид, состоящий из последовательности нуклеиновой кислоты, представленной SEQ ID NO: 2 или SEQ ID NO: 8;

Cse2 (Cas11); полинуклеотид, состоящий из последовательности нуклеиновой кислоты, представленной SEQ ID NO: 3 или SEQ ID NO: 9;

Cas5; полинуклеотид, состоящий из последовательности нуклеиновой кислоты, представленной SEQ ID NO: 4 или SEQ ID NO: 10;

Cas6; полинуклеотид, состоящий из последовательности нуклеиновой кислоты, представленной SEQ ID NO: 5 или SEQ ID NO: 11;

Cas7; полинуклеотид, состоящий из последовательности нуклеиновой кислоты, представленной SEQ ID NO: 6 или SEQ ID NO: 12.

Каждый из них представляет собой полинуклеотид, полученный так, чтобы он функционировал и экспрессировался в клетках млекопитающих, путем искусственной модификации последовательностей нуклеиновой кислоты, кодирующих группу белков Cas дикого типа E.coli (Cas3; SEQ ID NO: 13, Cse1 (Cas8); SEQ ID NO 14: Cse2 (Cas11); SEQ ID NO: 15, Cas5; SEQ ID NO: 16, Cas6; SEQ ID NO: 17, Cas7; SEQ ID NO: 18).

Вышеуказанная искусственная модификация полинуклеотидов предназначена для модификации последовательности нуклеиновой кислоты, подходящей для ее экспрессии в эукариотических клетках, и для присоединения сигнала локализации в ядре. Модификация последовательности нуклеиновой кислоты и присоединение сигнала локализации в ядре являются такими, как описано выше. В результате можно ожидать, что для группы белков Cas уровень экспрессии будет достаточно увеличен, а функции будут улучшены.

Другим вариантом полинуклеотидов, кодирующих группу белков Cas, используемую в настоящем изобретении, является полинуклеотид, образованный путем модификации последовательности нуклеиновой кислоты, кодирующей группу белков Cas дикого типа, и состоящий из последовательности нуклеиновой кислоты, которая на 90% или более идентична последовательности нуклеиновой кислоты вышеуказанной группы белков Cas. Белки, экспрессируемые из этих полинуклеотидов, обладают активностью расщепления ДНК при образовании комплекса с белками, экспрессируемыми из других полинуклеотидов, составляющих группу белков Cas.

Идентичность последовательностей нуклеиновых кислот может составлять по меньшей мере 90% или более, а более предпочтительно 95% или более (например, 95%, 96%, 97%, 98% и 99% или более) по всей последовательности нуклеиновой кислоты (или области, кодирующей сайт, необходимый для функционирования Cse3). Идентичность последовательностей нуклеиновых кислот можно определить с помощью такой программы, как BLASTN (см. Altschul SF (1990) Basic local alignment search tool, Journal of Molecular Biology, Vol. 215 (Issue 3), pp. 403-410). Примеры параметров для анализа последовательностей нуклеиновых кислот с помощью BLASTN включают число баллов=100 и длину слова=12. Специалистам в данной области известны конкретные методы анализа с помощью BLASTN. Добавление или удаление (пробел и т.п.) могут быть допустимы при выравнивании последовательностей нуклеиновых кислот для оптимального сравнения.

Более того, термин "обладающий активностью расщепления ДНК" означает способность расщеплять по меньшей мере один сайт полинуклеотидной цепи.

Предпочтительно, чтобы система CRISPR-Cas3 согласно изобретению расщепляла ДНК посредством специфического распознавания последовательности-мишени. Так, например, двойной люциферазный анализ, описанный в примере A-1, позволяет определить, может ли система CRISPR-Cas3 специфически распознавать последовательность-мишень.

Другой вариант полинуклеотидов, кодирующих группу белков Cas, используемую в настоящем изобретении, представляет собой полинуклеотид, который гибридизуется с полинуклеотидом, состоящим из последовательности нуклеиновой кислоты, комплементарной последовательности нуклеиновой

кислоты группы белков Cas, описанной выше, в жестких условиях. Белки, экспрессируемые из этих полинуклеотидов, обладают активностью расщепления ДНК при образовании комплекса с белками, экспрессируемыми из других полинуклеотидов, составляющих группу белков Cas.

Используемый здесь термин "жесткие условия" относятся к условиям, при которых две полинуклеотидные цепи образуют двухцепочечный полинуклеотид, специфичный к последовательности нуклеиновой кислоты, но не образуют неспецифический двухцепочечный полинуклеотид. Другими словами, выражение "гибридируется в жестких условиях" может означать условия, гибридизации при температуре в пределах от температуры плавления (значения T_m) нуклеиновых кислот с высокой идентичностью последовательности (например, идеально соответствующих гибридов) до температуры ниже на 15°C , предпочтительно на 10°C , а более предпочтительно на 5°C .

Примеры жестких условий приведены ниже. Во-первых, два типа полинуклеотидов гибридизуются в течение 16-24 ч при температуре от 60 до 68°C (предпочтительно 65°C , а более предпочтительно 68°C) в буферном растворе (pH 7,2), состоящем из $0,25\text{M Na}_2\text{HPO}_4$, 7% ДСН, 1 мМ EDTA и 1×раствора Денхардта. После этого промывку проводят два раза в течение 15 мин в буферном растворе (pH 7,2), состоящем из $20\text{ mM Na}_2\text{HPO}_4$, 1% ДСН и 1 мМ EDTA при температуре от 60 до 68°C (предпочтительно 65°C , а более предпочтительно 68°C).

Другие примеры включают следующий метод.

Во-первых, предварительную гибридизацию проводят в течение ночи при 42°C в растворе для гибридизации, содержащем 25% формамида (50% формамида в более жестких условиях), 4×SSC (хлорид натрия/цитрат натрия), 50 мМ Hepes (pH 7,0), 10×раствора Денхардта и 20 мкг/мл денатурированной ДНК спермы лосося. После этого добавляют меченые зонды и проводят инкубирование в течение ночи при 42°C для гибридизации полинуклеотидов двух типов.

Далее проводят промывку в любых из нижеисследованных условий. Нормальные условия: 1×SSC и 0,1% ДСН используют в качестве промывочных жидкостей для промывки при температуре приблизительно 37°C . Жесткие условия: $0,5\times\text{SSC}$ и 0,1% ДСН используются в качестве промывочных жидкостей для промывки при температуре приблизительно 42°C . Более жесткие условия: $0,2\times\text{SSC}$ и 0,1% ДСН используются в качестве промывочных жидкостей для промывки при температуре приблизительно 65°C .

По мере того как условия промывки для гибридизации становятся более жесткими, специфичность гибридизации возрастает. Следует отметить, что вышеуказанная комбинация условий SSC, ДСН и температур приводится лишь в иллюстративных целях. Жесткость, аналогичная приведенной выше, может быть достигнута путем соответствующего сочетания вышеупомянутых элементов для определения жесткости гибридизации или других элементов (например, концентрации зонда, длины зонда и времени реакции гибридизации). Это описано, например, в литературе Joseph Sambrook & David W. Russell, *Molecular cloning: a laboratory manual* 3rd Ed., New York: Cold Spring Harbor Laboratory Press, 2001.

Экспрессионный вектор, содержащий полинуклеотид, кодирующий группу белков Cas.

В настоящем изобретении можно использовать экспрессионный вектор для экспрессии группы белков Cas. Что касается экспрессионного вектора, то в качестве основного вектора могут быть использованы различные типы обычно используемых векторов, и эти векторы могут быть соответствующим образом выбраны в зависимости от клеток для введения или способа введения. Конкретными примерами, подходящими для использования, являются плазмиды, фаги, космиды и т.п. Тип вектора не имеет конкретных ограничений, и достаточно выбрать соответствующий вектор, который может быть экспрессирован в клетке-хозяине.

Примерами описанных выше экспрессионных векторов являются фаговые векторы, плазмидные векторы, вирусные векторы, ретровирусные векторы, хромосомные векторы, эписомные векторы, векторы, происходящие от вируса (бактериальные плазмиды, бактериофаги, дрожжевые эписомы и т.п.), дрожжевые хромосомные элементы и вирусы (бакуловирусы, паповавирусы, вирусы осповакцины, аденовирусы, трехвалентные поксвирусы, вирусы псевдобешенства, герпесвирусы, лентивирусы, ретровирусы и т.п.) и векторы, полученные из их комбинаций (космиды, фагмиды и т.п.).

Предпочтительно экспрессионный вектор также содержит сайт инициации транскрипции и сайт терминации транскрипции, а также сайт связывания с рибосомой в области транскрипции. Кодированная область зрелого транскрипта в векторе будет содержать кодон инициации транскрипции AUG в начале и соответствующим образом расположенный стоп-кодон в конце транслируемого полипептида.

В настоящем изобретении экспрессионный вектор для экспрессии группы белков Cas может содержать промоторную последовательность. Вышеуказанная промоторная последовательность может быть соответствующим образом выбрана в зависимости от типа эукариотической клетки, служащей в качестве хозяина. Кроме того, экспрессионный вектор может содержать последовательность для усиления транскрипции ДНК, например энхансерную последовательность. Примерами энхансеров являются энхансер SV40 (который расположен на 100-270 п.о. ниже ориджина репликации), энхансер раннего промотора цитомегаловируса, энхансер полиомы и энхансер аденовируса, расположенный ниже ориджина репликации. Кроме того, экспрессионный вектор может содержать последовательность для стабилизации транскрибированной РНК, например последовательность присоединения poly(A) (последовательность поли-

аденилирования, poly(A)). Примерами последовательностей присоединения poly(A) являются последовательности присоединения poly(A), происходящие от гена гормона роста; последовательности присоединения poly(A), происходящие от гена бычьего гормона роста; последовательности присоединения poly(A), происходящие от гена человеческого гормона роста; последовательности присоединения poly(A), происходящие от вируса SV40; и последовательности присоединения poly(A), происходящие от человеческого или кроличьего гена β -глобина.

Количество полинуклеотидов, кодирующих группу белков Cas, которые должны быть включены в один и тот же вектор, не имеет конкретных ограничений, при условии, что они могут выполнять функции систем CRISPR-Cas в клетке-хозяине, в которую был введен экспрессионный вектор. Так, например, можно создать конструкцию, в которой полинуклеотид, кодирующий группу белков Cas, будет присутствовать на векторах одного типа (одного и того же типа). Кроме того, можно также создать такую конструкцию, в которой весь полипептид или некоторые его части, кодирующие группы белков Cas, будут присутствовать на отдельных векторах. Так, например, можно создать такую конструкцию, в которой полинуклеотид, кодирующий белки Cascade, будет расположен на векторах одного типа (одного и того же типа), а полинуклеотид, кодирующий Cas3, будет присутствовать на других векторах. С точки зрения эффективности экспрессии и т.п., предпочтительно использовать метод создания полинуклеотида, кодирующего группы белков Cas, на векторах шести различных типов.

В другом случае, несколько полинуклеотидов, кодирующих одни и те же белки, могут быть встроены в одни и те же векторы для регуляции уровня экспрессии и т.п. Так, например, можно создать такую конструкцию, в которой полинуклеотиды, кодирующие Cas3, будут располагаться в двух сайтах векторов одного типа (одного и того же типа).

Кроме того, можно использовать экспрессионный вектор, который содержит множество последовательностей нуклеиновых кислот, кодирующих группу белков Cas, и который имеет последовательности нуклеиновых кислот, встроенные между этими множественными последовательностями нуклеиновых кислот и кодирующие аминокислотные последовательности (пептиды 2A и т.п.), расщепляемые внутриклеточными протеазами (см., например, структуру вектора на фиг. 8). Если полинуклеотиды, имеющие такие последовательности нуклеиновых кислот, транскрибируются и транслируются, то в клетках экспрессируются связанные полипептидные цепи. Впоследствии, благодаря действию внутриклеточных протеаз группы белков Cas разделяются, становятся отдельными белками, а затем образуют функциональные комплексы. Это позволяет регулировать соотношение количества групп белков Cas, экспрессируемых внутри клеток. Так, например, было предсказано, что Cas3 и Cse1 (Cas8) будут экспрессироваться в равных количествах из "экспрессионного вектора, содержащего одну последовательность нуклеиновой кислоты, кодирующую Cas3, и одну последовательность нуклеиновой кислоты, кодирующую Cse1 (Cas8)". Кроме того, можно экспрессировать множество групп белков Cas с помощью экспрессионного вектора одного типа вектора, что дает преимущества с точки зрения технологических свойств. С другой стороны, обычно лучшим является вариант, в котором группы белков Cas экспрессируются различными экспрессионными векторами с точки зрения высокой активности расщепления ДНК.

Экспрессионные векторы, используемые в настоящем изобретении, можно получить известными методами. Примеры таких методов включают метод, описанный в руководстве, прилагаемом к набору для получения векторов, а также методы, описанные в различных руководствах. Примером исчерпывающего руководства является руководство Joseph Sambrook & David W. Russell, *Molecular cloning: a laboratory manual* 3rd Ed., New York: Cold Spring Harbor Laboratory Press, 2001.

Экспрессионный вектор, содержащий cr-PHK, полинуклеотид, кодирующий cr-PHK, или полинуклеотид.

Система CRISPR-Cas3 согласно изобретению включает cr-PHK, полинуклеотид, кодирующий cr-PHK, или экспрессионный вектор, содержащий полинуклеотид, для нацеливания на ДНК в целях редактирования генома.

cr-PHK представляет собой PHK, которая образует часть системы CRISPR-Cas и имеет последовательность нуклеиновой кислоты, комплементарную последовательности-мишени. Система CRISPR-Cas3 согласно изобретению позволяет с помощью cr-PHK специфически распознавать последовательность-мишень и расщеплять эту последовательность. В системах CRISPR-Cas, типированных системой CRISPR-Cas9, в качестве cr-PHK обычно используются зрелые cr-PHK. Хотя причина этого не ясна, однако было обнаружено, что использование зрелой cr-PHK нежелательно, если система CRISPR-Cas3 предназначена для функционирования в эукариотических клетках. Более того, неожиданно было обнаружено, что высокоэффективное редактирование генома может быть осуществлено в эукариотических клетках с использованием pre-cr-PHK вместо зрелой cr-PHK. Этот факт был очевиден из эксперимента по сравнению зрелой cr-PHK и pre-cr-PHK (фиг. 10). Поэтому особенно предпочтительно использовать pre-cr-PHK в качестве cr-PHK согласно изобретению.

Pre-cr-PHK, используемые в настоящем изобретении, обычно имеют структуры "лидерная последовательность-повторяющаяся последовательность-спейсерная последовательность-повторяющаяся последовательность (структура LRSR)" и "повторяющаяся последовательность-спейсерная последовательность-повторяющаяся последовательность (структура RSR)". Лидерная последовательность представляет

собой АТ-богатую последовательность и функционирует как промотор для экспрессии пре-сг-РНК. Повторяющаяся последовательность представляет собой повторяющуюся последовательность со спейсерной последовательностью между повторами, а спейсерная последовательность представляет собой последовательность, сконструированную в настоящем изобретении как последовательность, комплементарную ДНК-мишени (по своей природе, она представляет собой последовательность, происходящую из чужеродной ДНК, включенной в процессе адаптации). Пре-сг-РНК становится зрелой сг-РНК при расщеплении белками, составляющими Cascade (например, Cas6 для типов I-A, B и D-E и Cas5 для типа I-C).

Как правило, длина цепи лидерной последовательности составляет 86 оснований, а длина цепи повторяющейся последовательности составляет 29 оснований. Длина цепи спейсерной последовательности составляет, например, 10-60 оснований, предпочтительно 20-50 оснований, более предпочтительно 25-40 оснований и обычно 32-37 оснований. Таким образом, в случае структуры LRSR, пре-сг-РНК, используемая в настоящем изобретении, имеет длину цепи, например, 154-204 оснований, предпочтительно 164-194 оснований, более предпочтительно 169-184 оснований, а обычно 176-181 оснований. Кроме того, в случае структуры RSR длина цепи составляет, например, 68-118 оснований, предпочтительно 78-108 оснований, более предпочтительно 83-98 оснований, а обычно 90-95 оснований.

Для того чтобы система CRISPR-Cas3 согласно изобретению функционировала в эукариотических клетках, считается, что важную роль играет процесс, посредством которого повторяющиеся последовательности пре-сг-РНК расщепляются белками, составляющими Cascade. Таким образом, следует отметить, что вышеупомянутые повторяющиеся последовательности могут быть короче или длиннее, чем указанная выше длина цепи, при условии, что будет происходить такое расщепление. В частности, можно сказать, что пре-сг-РНК представляет собой сг-РНК, образованную путем присоединения последовательностей, достаточных для расщепления белками, составляющими Cascade, к обоим концам зрелой сг-РНК, описанной ниже. Таким образом, предпочтительный вариант осуществления способа согласно изобретению включает стадию расщепления сг-РНК белками, составляющими Cascade, после введения системы CRISPR-Cas3 в эукариотические клетки.

С другой стороны, зрелая сг-РНК, образованная расщеплением пре-сг-РНК, имеет структуру "5'-фланкирующая последовательность - спейсерная последовательность - 3'-фланкирующая последовательность". Обычно 5'-фланкирующая последовательность состоит из 8 оснований в положениях 22-29 повторяющейся последовательности и содержится в Cas5. Кроме того, 3'-фланкирующая последовательность обычно состоит из 21 основания в положениях 1-21 в повторяющейся последовательности, образует структуру стебель-петля с основаниями в положениях 6-21 и сохраняется в Cas6. Таким образом, длина цепи зрелой сг-РНК обычно составляет 61-66 оснований. Следует отметить, что, поскольку существуют также зрелые сг-РНК, не имеющие 3'-фланкирующей последовательности, то в зависимости от типа системы CRISPR-Cas3, длина цепи в этом случае сокращается на 21 основание.

Следует отметить, что последовательность РНК может быть соответствующим образом сконструирована в соответствии с последовательностью-мишенью, в которой желательно осуществлять редактирование ДНК. Кроме того, можно синтезировать РНК с применением любого метода, известного специалистам в данной области.

Эукариотическая клетка.

Примеры "эукариотических клеток" в настоящем изобретении включают клетки животных, клетки растений, клетки водорослей и клетки грибов. Кроме того, примеры клеток животных включают клетки млекопитающих, а также, например, клетки рыб, птиц, рептилий, амфибий и насекомых.

Примеры "клеток животных" включают клетки, составляющие организм животных; клетки, составляющие органы/ткани, вырезанные у животных; и культивируемые клетки, полученные из тканей животных. Конкретными примерами являются зародышевые клетки, такие как ооциты и сперма; эмбриональные клетки эмбрионов на различных стадиях развития (такие как 1-клеточные эмбрионы, 2-клеточные эмбрионы, 4-клеточные эмбрионы, 8-клеточные эмбрионы, 16-клеточные эмбрионы и эмбрионы на стадии морулы); стволовые клетки, такие как индуцированные плюрипотентные стволовые (iPS) клетки; эмбриональные стволовые (ES) клетки; и соматические клетки, такие как фибробласты, гемопоэтические клетки, нейроны, мышечные клетки, костные клетки, клетки печени, клетки поджелудочной железы, клетки головного мозга и клетки почек. Ооциты можно использовать до оплодотворения и после оплодотворения в качестве ооцитов, используемых для создания животных с отредактированным геномом, а предпочтительно ооцитов после оплодотворения, т.е. оплодотворенных яйцеклеток. Особенно предпочтительно, чтобы оплодотворенные эмбриональные яйцеклетки происходили от клеток на стадии пронуклеуса. Ооциты могут быть оттаяны и использованы после замораживания.

В настоящем изобретении "млекопитающее" представляет собой понятие, включающее человека и млекопитающих, не являющихся человеком. Примеры млекопитающих, не являющихся человеком, включают парнокопытных млекопитающих, таких как крупный рогатый скот, кабаны, свиньи, овцы и козы, непарнокопытных млекопитающих, таких как лошади, грызуны, такие как мыши, крысы, морские свинки, хомяки и белки, зайцеобразные, такие как кролики, и плотоядные животные, такие как собаки, кошки и хорьки. Описанные выше млекопитающие, не являющиеся человеком, могут представлять собой крупный рогатый скот или животных-компаньонов (домашние питомцы) или диких животных.

Примеры "клеток растений" включают клетки зерновых культур, масличных культур, кормовых культур, плодовых и овощных культур. Примеры "клеток растений" включают клетки, составляющие организм растения; клетки, составляющие органы и ткани, отделенные от растений, и культивируемые клетки, полученные из тканей растений. Примеры органов и тканей растений включают листья, стебли, верхушки побегов (точки роста), корни, клубни и каллусы. Примерами растений являются рис, кукуруза, банан, арахис, подсолнечник, томат, рапс, табак, пшеница, ячмень, картофель, соя, хлопчатник и гвоздика, а также материалы для их размножения (например, семена, клубневидные корни и клубни).

Редактирование ДНК.

В настоящем изобретении "редактирование ДНК эукариотических клеток" может представлять собой стадию, на которой ДНК эукариотической клетки редактируется *in vivo* или *in vitro*. Кроме того, "редактирование ДНК" означает процессы, представленные ниже следующими типами (включая их комбинации).

Следует отметить, что в настоящем описании ДНК, используемая в вышеупомянутом контексте, включает не только ДНК, присутствующую в ядре клетки, но также экзогенную ДНК и ДНК, присутствующую не в ядре клетки, такую как митохондриальная ДНК.

1. Расщепление цепи ДНК в сайте-мишени.
2. Делеция основания цепи ДНК в сайте-мишени.
3. Инсерция основания в цепь ДНК в сайте-мишени.
4. Замена основания цепи ДНК в сайте-мишени.
5. Модификация основания цепи ДНК в сайте-мишени.
6. Модуляция транскрипции ДНК (гена) в сайте-мишени.

В одном варианте системы CRISPR-Cas3 согласно изобретению используется белок, обладающий ферментативной активностью, для модификации ДНК-мишени другим способом, кроме расщепления ДНК. Этот вариант осуществления изобретения может быть достигнут, например, путем присоединения Cas3 или Cascade к гетерологичному белку, имеющему желаемую ферментативную активность, с образованием химерного белка. Таким образом, "Cas3" и "Cascade" согласно изобретению также включают такие гибридные белки. Примерами ферментативной активности гибридного белка являются, но не ограничиваются ими, дезаминазная активность (например, цитидин-дезаминазная активность и аденозин-дезаминазная активность), метилтрансферазная активность, активность деметилирующего фермента, ДНК-репарирующая активность, ДНК-повреждающая активность, дисмутазная активность, алкилирующая активность, депуринизирующая активность, окисляющая активность, активность образования пириимидинового димера, интегразная активность, транспозазная активность, рекомбиназная активность, полимеразная активность, лигазная активность, активность фотореактивирующего фермента и гликозилазная активность. В этом случае нуклеазная активность или геликазная активность Cas3 является необязательной. По этой причине в качестве Cas3 можно использовать мутант, в котором некоторые или все эти активности были удалены (например, мутант домена D H74A (dnCas3), мутант K320N с мотивом 1 домена SF2 (dhCas3) и двойной мутант S483A/T485A с мотивом 3 домена SF2 (dh2Cas3)). Точное редактирование генома можно осуществить путем замены оснований без какой-либо крупной делеции в сайте-мишени, если, например, гибридный белок деаминазы и мутант, в котором некоторые или все нуклеазные активности Cas3 были удалены, используются в качестве составного элемента системы CRISPR-Cas3 согласно изобретению. Способ применения деаминазы к системам CRISPR-Cas хорошо известен (Nishida K. et al., Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems, Science, DOI: 10.1126/science.aaf8729, (2016)), и достаточно просто применить его к системе CRISPR-Cas3 согласно изобретению.

Другой вариант системы CRISPR-Cas3 согласно изобретению регулирует транскрипцию гена в сайте связывания системы согласно изобретению без расщепления ДНК. Этот вариант может быть достигнут, например, путем присоединения Cas3 или Cascade к нужному белку регуляции транскрипции с получением химерного белка. Таким образом, "Cas3" и "Cascade" согласно изобретению также включают такие гибридные белки. Примерами белка регуляции транскрипции являются, но не ограничиваются ими, регуляторы транскрипции, индуцируемые излучением; регуляторы транскрипции, чувствительные к небольшим молекулам/лекарственным средствам, факторы транскрипции и репрессоры транскрипции. В этом случае нуклеазная активность или геликазная активность Cas3 является необязательной. По этой причине в качестве Cas3 можно использовать мутант, в котором некоторые или все эти активности удалены (например, мутант домена DH74A (dnCas3), мутант K320N с мотивом 1 домена SF2 (dhCas3) и двойной мутант S483A/T485A с мотивом 3 домена SF2 (dh2Cas3)). Способы использования белка регуляции транскрипции в системах CRISPR-Cas известны специалистам в данной области.

Кроме того, в системе CRISPR-Cas3 согласно изобретению рассматривается случай, например, использования мутанта, в котором некоторые или все нуклеазные активности Cas3 были удалены. Белки, обладающие другими нуклеазными активностями, могут быть присоединены к Cas3 или Cascade. Такой вариант входит в объем настоящего изобретения.

Кроме того, в системе CRISPR-Cas3 согласно изобретению рассматривается случай использования мутанта, в котором некоторые или все нуклеазные активности Cas3 были удалены, и использования ак-

тивностей других белков при редактировании ДНК. В настоящем описании "активность расщепления ДНК" соответствующим образом интерпретируется как различные активности, которыми обладают эти белки.

Кроме того, редактирование ДНК может быть осуществлено на ДНК, содержащейся в конкретной клетке индивидуума. Такое редактирование ДНК может быть осуществлено, например, на конкретной клетке, выбранной в качестве мишени из клеток, составляющих организм животного или растения.

Метод введения молекул, составляющих систему CRISPR-Cas3 согласно изобретению, в эукариотические клетки в форме полинуклеотида или экспрессионного вектора, содержащего полинуклеотид, не имеет ограничений. Примеры такого метода включают электропорацию, метод с использованием фосфата кальция, липосомный метод, метод с использованием DEAE-декстрана, метод микроинъекции, трансфекцию, опосредованную катионными липидами, электропорацию, трансдукцию и инфицирование вирусными векторами. Такие методы описаны во многих стандартных лабораторных руководствах, таких как "Leonard G. Davis et al., Basic methods in molecular biology, New York: Elsevier, 1986".

Метод введения молекул системы CRISPR-Cas3 согласно изобретению в эукариотические клетки в форме белка не имеет ограничений. Примерами являются электропорация, трансфекция, опосредованная катионными липидами, и микроинъекция.

Редактирование ДНК согласно изобретению может применяться в различных областях. Примеры применения включают генотерапию, улучшение породы животных, получение трансгенных животных или клеток, получение полезных веществ и исследования в области биологических наук.

В качестве методов выращивания особей, не являющихся человеком, из клеток, могут быть применены известные методы. Зародышевые клетки или плюрипотентные стволовые клетки обычно используются в случае выращивания особей, не являющихся человеком, из клеток животных. Так, например, молекулы, составляющие систему CRISPR-Cas3 согласно изобретению, вводят в ооцит. Полученный ооцит затем трансплантируют в матку самки млекопитающего, не являющейся человеком, которая находится в состоянии псевдобеременности. После этого получают помет. Трансплантация может быть проведена в оплодотворенной яйцеклетке 1-клеточного эмбриона, 2-клеточного эмбриона, 4-клеточного эмбриона, 16-клеточного эмбриона или эмбриона на стадии морула. При желании, ооцит можно культивировать до трансплантации в подходящих условиях. Трансплантация и культивирование ооцитов могут быть проведены общеизвестным методом (Nagy A. et al., Manipulation Mouse Embryo, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 2003). Из индивидуума, не являющегося человеком, также можно получить клоны или потомство, у которых была отредактирована желаемая ДНК.

Кроме того, давно известно, что соматические клетки растений обладают тотипотентностью к дифференцировке, и были разработаны способы регенерации растений из клеток различных полученных растений. Следовательно, может быть получено растение, в котором нужная ДНК была подвергнута нокину путем введения молекул, составляющих систему CRISPR-Cas3 согласно изобретению, в клетки растения и последующей регенерации растений из полученных клеток растений. Также могут быть получены потомство, клоны или материалы для размножения, в которых была отредактирована желаемая ДНК. В качестве метода редифференцировки ткани растения путем культивирования ткани для выращивания особи может быть применен метод, хорошо известный в данной области техники (Protocols for Plant Transformation, edited by Tabei Yutaka, Kagaku-Dojin, pp. 340-347 (2012)).

2. Набор, используемый в системе CRISPR-Cas3.

Набор, используемый в системе CRISPR-Cas3 согласно изобретению, включает следующие (А) и (В):

А) белок Cas3, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид, и

В) белок Cascade, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид.

Набор может дополнительно содержать sg-PHK, полинуклеотид, кодирующий sg-PHK, или экспрессионный вектор, содержащий этот полинуклеотид.

Составные элементы набора согласно изобретению могут представлять собой вариант, в котором все или некоторые из этих элементов являются смешанными, или могут представлять собой вариант, в котором каждый из них является независимым.

Набор согласно изобретению может быть использован в таких областях, как приготовление фармацевтических препаратов, пищевая промышленность, животноводство, рыболовство, различные отрасли промышленности, биотехнология и биологические исследования.

Здесь и далее набор согласно изобретению описан с точки зрения приготовления фармацевтических препаратов (лекарственных средств). Следует отметить, что в случае использования вышеописанного набора в таких областях, как животноводство, биотехнология и биологические исследования, такой набор может быть использован после ознакомления с нижеследующим объяснением на основе общих технических знаний в этих областях.

Фармацевтический препарат для редактирования ДНК клеток животных, включая человека, может быть приготовлен обычными методами с использованием системы CRISPR-Cas3 согласно изобретению.

Более конкретно фармацевтический препарат может быть приготовлен путем получения молекул, составляющих систему CRISPR-Cas3 согласно изобретению, например, с использованием добавок к фармацевтическому препарату.

Используемый здесь термин "добавка к фармацевтическому препарату" означает вещество, отличающееся от активных ингредиентов, содержащихся в фармацевтическом препарате. Добавка к фармацевтическому препарату представляет собой вещество, содержащееся в фармацевтическом препарате и вводимое для облегчения приготовления, стабилизации качества, повышения эффективности и т.п. Примеры добавок к фармацевтическому препарату, описанных выше, могут включать наполнители, связующие вещества, дезинтеграторы, лубриканты, псевдооживители (твердые антистатики), красители, покрытия для капсул, агенты для нанесения покрытий, пластификаторы, вкусовые добавки, подсластители, ароматизаторы, растворители, агенты, стимулирующие растворение, эмульгаторы, суспендирующие агенты (чувствительные к давлению адгезивы), загустители, регуляторы pH (подкислители, подщелачивающие агенты и буферы), увлажнители (солюбилизаторы), антибактериальные консерванты, хелатообразующие агенты, основы для суппозиторий, основы для мази, отвердители, мягчители, медицинские вода, пропелленты, стабилизаторы и консерванты. Эти добавки к фармацевтическим препаратам могут быть легко выбраны специалистами в данной области в соответствии с предполагаемой лекарственной формой и способом введения, а также в соответствии со стандартной фармацевтической практикой.

Кроме того, фармацевтический препарат для редактирования ДНК клеток животных с использованием системы CRISPR-Cas3 согласно изобретению может содержать дополнительные активные ингредиенты. Дополнительные активные ингредиенты не имеют конкретных ограничений и могут быть соответствующим образом разработаны специалистами в данной области.

Конкретные примеры активных ингредиентов и добавок к фармацевтическим препаратам, описанные выше, могут быть изучены в соответствии со стандартами, установленными, например, Управлением по контролю и качеству продуктов и лекарственных средств США (FDA), Европейским агентством по лекарственным средствам (EMA), Министерством здравоохранения, труда и благосостояния Японии.

Примеры способов доставки фармацевтического препарата к желаемым клеткам включают способы с использованием вирусных векторов, нацеленных на клетки (аденовирусных векторов, аденоассоциированных вирусных векторов, лентивирусных векторов, векторов на основе вируса Сендай и т.п.) или антител, специфически распознающих клетки. Фармацевтический препарат может быть приготовлен в любой лекарственной форме в зависимости от цели. Кроме того, вышеуказанный фармацевтический препарат, как и положено, назначается врачом или медицинским работником.

Набор согласно изобретению предпочтительно включает также инструкцию по его применению.

Примеры

Настоящее изобретение более подробно описано ниже со ссылкой на примеры, но оно не ограничивается только этими нижеследующими примерами.

А. Создание системы CRISPR-Cas3 в эукариотической клетке.

Материал и метод.

1. Получение репортерных векторов, содержащих последовательности-мишени.

Последовательности-мишени представляют собой последовательность, происходящую от человеческого гена CCR5 человека (SEQ ID NO: 19), и спейсерную последовательность CRISPR E.coli (SEQ ID NO: 22).

Для встраивания последовательностей-мишеней в векторы были получены синтетический полинуклеотид (SEQ ID NO: 20), содержащий последовательность-мишень, происходящую от человеческого гена CCR5 (SEQ ID NO: 19), и синтетический полинуклеотид (SEQ ID NO: 21), содержащий последовательность, комплементарную вышеуказанной последовательности-мишени (SEQ ID NO: 19). Аналогичным образом были получены синтетический полинуклеотид (SEQ ID NO: 23), содержащий последовательность-мишень, происходящую от спейсерной последовательности CRISPR E.coli (SEQ ID NO: 22), и синтетический полинуклеотид (SEQ ID NO: 24), содержащий последовательность, комплементарную вышеуказанной последовательности-мишени (SEQ ID NO: 22). Все вышеперечисленные синтетические полинуклеотиды были получены от Hokkaido System Science Co., Ltd.

Вышеуказанные полинуклеотиды были встроены в репортерные векторы методом, описанным в литературе Sakuma T. et al. (2013) Efficient TALEN construction and evaluation methods for human cell and animal applications, Genes to Cells, Vol. 18 (Issue 4), pp. 315-326. Описание этого метода приводится ниже. Сначала полинуклеотиды, имеющие последовательности, комплементарные друг другу (полинуклеотид SEQ ID NO: 20 и полинуклеотид SEQ ID NO: 21; полинуклеотид SEQ ID NO: 23 и полинуклеотид SEQ ID NO: 24), нагревали при 95°C в течение 5 мин, а затем охлаждали до комнатной температуры и гибридизовали. Для вышеуказанной стадии использовали блочный инкубатор (BI-515A, Astec). Затем полинуклеотид, гибридизованный с образованием двухцепочечной структуры, встраивали в базовый вектор с получением репортерного вектора.

Последовательности полученных репортерных векторов представлены в SEQ ID NO: 31 (репортерный вектор, содержащий последовательность-мишень, происходящую от человеческого гена CCR5) и SEQ ID NO: 32 (репортерный вектор, содержащий последовательность-мишень, происходящую от спей-

серной последовательности CRISPR *E. coli*). Кроме того, структура репортерного вектора показана на фиг. 4(d).

2. Получение экспрессионных векторов Cse1 (Cas8), Cse2 (Cas11), Cas5, Cas6, Cas7 и cr-PHK.

Амплификация и получение вставок.

В данном случае рассматриваются полинуклеотиды, имеющие модифицированные последовательности нуклеиновых кислот, кодирующие Cse1 (Cas8), Cse2 (Cas11), Cas5, Cas6 и Cas7 (с SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, и SEQ ID NO: 6 соответственно). Сначала продуцирование полинуклеотидов, связанных с SEQ ID NO: 2-SEQ ID NO: 3-SEQ ID NO: 6-SEQ ID NO: 4-SEQ ID NO: 5 в порядке следования (полинуклеотидов, имеющих связанные последовательности нуклеиновых кислот для кодирования Cse1 (Cas8)-Cse2 (Cas11)-Cas7-Cas5-Cas6 в этом порядке), было передано Корпорации GenScript как заказ на приобретение. Последовательности нуклеиновых кислот, кодирующие белки Cse1 (Cas8)-Cse2 (Cas11)-Cas7-Cas5-Cas6, были связаны с пептидами 2A (аминокислотная последовательность GSGATNFSLLKQAGDVEENPGP (SEQ ID NO: 58)).

Следует отметить, что последовательности нуклеиновой кислоты, кодирующие пептиды 2A, слегка различались в зависимости от сайтов связывания белка Cas и представляли собой следующие последовательности.

Последовательность между Cse1 (Cas8) и Cse2 (Cas11)

GGAAGCGGAGCAACCAACTTCAGCCTGCTGAAGCAGGCCGGCGATGTGGAGGAGA
ATCCAGGCCCC (SEQ ID NO: 59).

Последовательность между Cse2 (Cas11) и Cas7

GGCTCCGGCGCCACCAATTTTCTCTGCTGAAGCAGGCAGGCGATGTGGAGGAGAA
CCCAGGACCT (SEQ ID NO: 60).

Последовательность между Cas7 и Cas5

GGATCTGGAGCCACCAATTTTCAGCCTGCTGAAGCAAGCAGGCGACGTGGAAGAAAA
CCCAGGACCA (SEQ ID NO: 61).

Последовательность между Cas5 и Cas6

GGATCTGGGGCTACTAATTTTCTCTGCTGAAGCAAGCCGGCGACGTGGAAGAGAA
TCCAGGACCG (SEQ ID NO: 62).

Затем каждый из полинуклеотидов амплифицировали в условиях ПЦР (праймер и период времени), как показано в следующей таблице. Для ПЦР использовали термоячейку 2720 Thermal cycler (прикладные биосистемы).

Таблица 1

Праймер	Последовательность	Время
Cse1 – U	GCAAAGAATTCAGATCTCCACCATGCCTAAGAA GAAGAGAAAAAGTGAACCTGCTGATTGAC (SEQ ID NO: 36)	98°C (10 сек.) → 68°C (1 мин.) 35 циклов
Cse1 – L	TCATCGATGCATCTCGAGTTATCCATTAGAAGG TCCTCCCTGTGGCTTC (SEQ ID NO: 37)	
Cse2 – U	GCAAAGAATTCAGATCTCCACCATGCCCAAGAA GAAGCGGAAGGTGGCCGATGAGATCGAC (SEQ ID NO: 38)	98°C (10 сек.) → 68°C (1 мин.) 35 циклов
Cse2 – L	TCATCGATGCATCTCGAGTTAGGCGTTCTTATTT GTGGTCAGCACGAAG (SEQ ID NO: 39)	
Cas5 – U	GCAAAGAATTCAGATCTCCACCATGCCCAAGAA GAAGCGGAAGGTGTCCAATTTTCATCAAC (SEQ	

	ID NO: 40)	
Cas5 – L	TCATCGATGCATCTCGAGTTATGCCTCTCCATTG TTCCGCACCCAGCTC (SEQ ID NO: 41)	
Cas6 – U	GCAAAGAATTCAGATCTCCACCATGCCCAAGAA GAAGCGGAAAGTGACCTGAGCAAAGTG (SEQ ID NO: 42)	
Cas6 – L	TCATCGATGCATCTCGAGTTACAGAGGTGCCAG TGACAGCAGCCAC (SEQ ID NO: 43)	
Cas7 – U	GCAAAGAATTCAGATCTCCACCATGCCCAAGAA GAAGCGGAAGGTGCGCTCCTACCTGATC (SEQ ID NO: 44)	98°C (10 сек.) → 68°C (1 мин. 40 сек.)
Cas7 – L	TCATCGATGCATCTCGAGTTACTGGCTCACGTC CATTCCTCCCTTGATC (SEQ ID NO: 45)	35 циклов

Полинуклеотиды, имеющие следующие комплементарные последовательности, были получены в виде полинуклеотидов, имеющих последовательности нуклеиновых кислот для экспрессии cr-PHK.

1. Полинуклеотиды для экспрессии cr-PHK, соответствующие последовательности, происходящей от человеческого гена CCR5 (SEQ ID NO: 25 и 26, полученные от Hokkaido System Science Co., Ltd.).

2. Полинуклеотиды для экспрессии cr-PHK, соответствующие спейсерной последовательности CRISPR E.coli (SEQ ID NO: 27 и 28, полученные от Hokkaido System Science Co., Ltd.).

3. Полинуклеотиды для экспрессии cr-PHK, соответствующие последовательности, происходящей от человеческого гена EMX1 (SEQ ID NO: 29 и 30, полученные от Pharmacia).

Лигирование и трансформация.

В качестве субстратной плазмиды использовали pPB-CAG-EBNXN (поставляемый от Sanger Center). В NEB-буфере смешивали 1,6 мкг субстратной плазмиды, 1 мкл рестриктирующего фермента BglIII (New England Biolabs) и 0,5 мкл XhoI (New England Biolabs) и подвергали реакции при 37°C в течение 2 ч. Расщепленные субстратные плазмиды очищали с помощью набора для гель-экстракции (Qiagen).

Полученные таким образом субстратные плазмиды и вышеуказанные вставки лигировали с помощью системы сборки Гибсона. Лигирование осуществляли в соответствии с протоколом системы сборки Гибсона с отношением субстратных плазмид к вставкам, равным 1:1 (при 50°C в течение 25 мин, общий объем реакционного раствора: 8 мкл).

Затем 6 мкл раствора плазмид, полученного, как описано выше (раствора для реакции лигирования), и компетентных клеток (приготовленных в лаборатории Takeda) использовали для проведения трансформации обычным методом.

После этого плазмидные векторы очищали от трансформированной E. coli с применением препаративного щелочного метода. Вкратце, плазмидные векторы выделяли с использованием набора QIAprep Spin Miniprep (Qiagen), и выделенные плазмидные векторы очищали методом преципитации этанолом, а затем доводили до концентрации 1 мкг/мкл в ТЕ-буферном растворе.

Структура каждого плазмидного вектора показана на фиг. 4(а)-4(с). Кроме того, последовательности нуклеиновых кислот экспрессионных векторов pre-cr-PHK показаны в SEQ ID NO: 33 (экспрессионный вектор для экспрессии cr-PHK, соответствующий последовательности, происходящей от человеческого гена CCR5), SEQ ID NO: 34 (экспрессионный вектор для экспрессии cr-PHK, соответствующий спейсерной последовательности CRISPR E. coli) и SEQ ID NO: 35 (экспрессионный вектор для экспрессии cr-PHK, соответствующий последовательности, происходящей от человеческого гена EMX1).

3. Получение экспрессионного вектора Cas3.

Полинуклеотид, имеющий модифицированную последовательность нуклеиновой кислоты, кодирующую Cas3 (SEQ ID NO: 1), был получен от Genscript. В частности, вектор pUC57, включающий полинуклеотид, описанный выше, был получен от Genscript.

Вышеуказанный вектор расщепляли рестриктирующим ферментом NotI. Затем для выравнивания края фрагмента использовали 2 ед. фрагмента Кленова (Takara Bio Inc.) и 1 мкл 2,5 мМ смеси dNTP (Takara Bio Inc.). После этого указанный фрагмент очищали путем гель-экстракции (Qiagen). Очищенный фрагмент дополнительно расщепляли рестриктирующим ферментом XhoI и очищали путем гель-экстракции (Qiagen).

Очищенный фрагмент лигировали с использованием субстратной плазмиды (вектора pTL2-CAG-IRES-NEO, приготовленного в лаборатории Takeda) и набора для лигирования (Mighty Mix, Takara Bio Inc.). После этого трансформацию и очистку проводили так же, как описано в п.2. Выделенный плазмидный вектор получали так, чтобы он имел концентрацию 1 мкг/мкл в ТЕ-буферном растворе.

4. Получение плазмидного вектора, содержащего BPNLS.

Были получены экспрессионные векторы Cas3, Cse1 (Cas8), Cse2 (Cas11), Cas5, Cas6 и Cas7, в которых BPNLS были связаны с 5'-концом и 3'-концом (см. фиг. 7).

Получение вставки для каждой группы белков Cas, содержащих BPNLS на обоих концах, было заказано Thermo Fisher Scientific. Конкретная последовательность вышеуказанной вставки представляет собой (AGATCTTAATACGACTCACTATAGGGAGAGCCGCCACCATGGCC: SEQ ID NO: 56) - (любая из SEQ ID NO: 7-12) - (TAATATCCTCGAG: SEQ ID NO: 57). SEQ ID NO: 56 представляет собой последовательность, снабженную сайтом расщепления ферментов BgIII. SEQ ID NO: 57 представляет собой последовательность, снабженную сайтом расщепления ферментом XhoI.

Вектор pMK, включающий вышеуказанную последовательность, расщепляли рестриктирующими ферментами BgIII и XhoI и очищали путем гель-экстракции (Qiagen). Очищенный фрагмент лигировали с использованием субстратной плазмиды (pPB-CAG-EBNXXN, поставляемой от Sanger Center) и набора для лигирования (Mighty Mix, Takara Bio Inc.). После этого трансформацию и очистку проводили, как описано в п.2. Выделенный плазмидный вектор получали так, чтобы он имел концентрацию 1 мкг/мкл в ТЕ-буферном растворе.

5. Получение плазмидного вектора, содержащего Cascade (2A).

Был получен экспрессионный вектор, в котором последовательность нуклеиновой кислоты имела Cse1 (Cas8), Cse2 (Cas11), Cas7, Cas5 и Cas6, связанные в указанном порядке. Более конкретно был получен экспрессионный вектор, имеющий следующую структуру: (NLS-Cse1 (Cas8): SEQ ID NO: 2)-2A-(NLS-Cse2 (Cas11): SEQ ID NO: 3)-2A-(NLS-Cas7: SEQ ID NO: 6)-2A-(NLS-Cas5: SEQ ID NO: 4)-2A-(NLS-Cas6: SEQ ID NO: 5) (см. фиг. 8). Следует отметить, что аминокислотная последовательность NLS представляет собой PKKKRKV (SEQ ID NO: 52), а последовательность нуклеиновой кислоты представляет собой CCCAAGAAGAAGCGGAAGGTG (SEQ ID NO: 53).

Кроме того, аминокислотная последовательность пептида 2A представляет собой GSGATNFSLLKQAGDVEENPGP (SEQ ID NO: 58) (соответствующие последовательности нуклеиновых кислот представляют собой SEQ ID NO: 59-62).

Полипептид, имеющий вышеуказанную последовательность нуклеиновой кислоты, был получен от GenScript. Вектор pUC57, включающий вышеуказанную последовательность, расщепляли рестриктирующим ферментом EcoRI-HF и очищали путем гель-экстракции (Qiagen). Очищенный фрагмент лигировали с использованием субстратной плазмиды (вектор pTL2-CAG-IRES-Puro, полученный в лаборатории Takeda), и набора для лигирования (Mighty Mix, Takara Bio Inc.). После этого трансформацию и очистку проводили так же, как в п.2. Выделенный плазмидный вектор получали так, чтобы он имел концентрацию 1 мкг/мкл в ТЕ-буферном растворе.

Пример A-1.

Активность расщепления последовательности-мишени экзогенной ДНК оценивали следующим образом. cr-PHK и Cas3, Cse1 (Cas8), Cse2 (Cas11), Cas5, Cas6 и Cas7, присоединенные к сигналам локализации в ядре для получения модифицированной последовательности нуклеиновой кислоты, были экспрессированы в клетках 293Т НЕК (эмбриональных клетках почек человека).

Перед трансфекцией клетки НЕК 293Т культивировали в 10 см-чашке. Культивирование клеток НЕК 293Т проводили в среде EF (GIBCO) при 37°C в атмосфере 5% CO₂. Плотность клеток НЕК 293Т в среде EF довели до 3×10⁴/100 мкл.

Кроме того, 100 нг указанного выше репортерного вектора, 200 нг каждой из плазмиды Cas3, плазмиды Cse1 (Cas8), плазмиды Cse2 (Cas11), плазмиды Cas5, плазмиды Cas6, плазмиды Cas7 и плазмиды cr-PHK; 60 нг вектора pRL-TK (способного экспрессировать люциферазу Renilla, Promega); и 300 нг вектора pBlueScript II KS (+) (Agilent Technologies) смешивали в 25 мкл Opti-MEM (Thermo Fisher Scientific). Условия с использованием репортерного вектора, имеющего последовательность-мишень, происходящую от CCR5 в качестве репортерного вектора, соответствуют условиям 1 на фиг. 1, а условия с использованием репортерного вектора, имеющего спейсерную последовательность CRISPR E. coli, соответствуют условиям 10 на фиг. 1.

Затем 1,5 мкл липофектамина 2000 (Thermo Fisher Scientific) и 25 мкл OptiMEM (Thermo Fisher Scientific) смешивали и инкубировали при комнатной температуре в течение 5 мин. После этого вышеуказанную смесь "плазида+OptiMEM" и смесь "липофектамин 2000+OptiMEM" смешивали и инкубировали при комнатной температуре в течение 20 мин. Полученную смесь смешивали с 1 мл вышеуказанной среды EF, содержащей клетки НЕК 293Т, и высевали в 96-луночный планшет (высевали всего в 12 лунок, по 1 лунке на комбинацию векторов).

После культивирования при 37°C в атмосфере 5% CO₂ в течение 24 ч проводили двойной анализ на люциферазу в соответствии с протоколом двойного анализа люциферазной системы Dual-Glo (Promega). Для оценки люциферазы и люциферазы Renilla использовали Centra XS³ LB 960 (BERTHOLD TECHNOLOGIES).

Такой же эксперимент был проведен в следующих условиях в качестве контрольного эксперимента.

1. Вместо любой плазмиды Cas3, плазмиды Cse1 (Cas8), плазмиды Cse2 (Cas11), плазмиды Cas5, плазмиды Cas6 и плазмиды Cas7, то же самое количество вектора pBlueScript II KS (+) (Agilent Tech-

nologies) смешивали для экспрессии (2-7 на фиг. 1).

2. Вместо плазмиды *сг*-РНК, используемой в вышеуказанной процедуре, были смешаны плазмиды для экспрессии *сг*-РНК, не комплементарной последовательности-мишени. В частности, в целях проведения экспрессии плазмиды для экспрессии *сг*-РНК, соответствующие спейсерной последовательности CRISPR *E. coli*, смешивали для нацеливания на последовательность, происходящую от гена CCR5 (8 на фиг. 1), а плазмиды для экспрессии *сг*-РНК, соответствующей последовательности, происходящей от гена CCR5, смешивали в случае нацеливания на спейсерную последовательность CRISPR *E. coli* (11 на фиг. 1).

3. В качестве негативного контроля экспрессировали только репортерный вектор, имеющий последовательность-мишень, происходящую от CCR5 (9 на фиг. 1), и только репортерный вектор, имеющий спейсерную последовательность CRISPR *E. coli* (12 на фиг. 1).

Результаты.

Результаты двойного анализа на люциферазу представлены на графике на фиг. 1, а условия эксперимента показаны в нижней таблице на фиг. 1. На фиг. 1 "CCR5-мишень" и "спейсер-мишень" означают последовательность-мишень, происходящую от CCR5, и спейсерную последовательность CRISPR *E. coli* соответственно. Кроме того, "CCR5-*сг*-РНК" и "спейсер-*сг*-РНК" представляют собой последовательность, комплементарную CCR5-мишени, и последовательность, комплементарную спейсеру-мишени, соответственно.

На фиг. 1 система, в которую вводили плазмиду *сг*-РНК, комплементарную последовательности-мишени, и все плазмиды Cas3, Cse1 (Cas8), Cse2 (Cas11), Cas5, Cas6 и Cas7 обладали более высокой активностью расщепления, чем другие системы (при сравнении 1 и 2-8, и 10 и 11). Следовательно, было обнаружено, что Cas3, Cse1 (Cas8), Cse2 (Cas11), Cas5, Cas6 и Cas7 могут экспрессироваться в человеческих клетках посредством экспрессионных векторов в соответствии с вариантом осуществления настоящего изобретения.

Кроме того, было высказано предположение, что введение указанных выше экспрессионных векторов в человеческие клетки приводит к образованию комплексов Cas3, Cascade и *сг*-РНК в человеческих клетках и к расщеплению последовательности-мишени.

Кроме того, на фиг. 1 сравнение между 8 и 9 и между 11 и 12 показало, что активность расщепления была эквивалентна активности негативных контролен в системе, экспрессирующей *сг*-РНК, которая не комплементарна последовательности-мишени. Другими словами, было высказано предположение, что система CRISPR-Cas3 согласно изобретению может специфически расщеплять последовательности, комплементарные кРРНК, в клетках млекопитающих.

Пример А-2.

Был проведен эксперимент для того, чтобы определить, возможно ли расщепление эндогенной ДНК человеческих клеток с помощью систем CRISPR-Cas типа I с применением метода, описанного в примере А-1.

В частности, последовательность нуклеиновой кислоты в человеческих клетках была модифицирована для экспрессии *пре-сг*-РНК, и Cas3, Cse1 (Cas8), Cse2 (Cas11), Cas5, Cas6 и Cas7, присоединенных к сигналам локализации в ядре, и был проведен анализ на расщепление последовательности эндогенного гена CCR5 клеток.

Те же самые клетки НЕК 239Т, как и в примере А-1, высевали в 24-луночный планшет с плотностью 1×10^5 клеток/лунку и культивировали в течение 24 ч.

50 мкл Opti-MEM (Thermo Fisher Scientific) смешивали с 1 мкг плазмиды Cas3, 1,3 мкг плазмиды Cse1 (Cas8), 1,3 мкг плазмиды Cse2 (Cas11), 1,1 мкг плазмиды Cas5, 0,8 мкг плазмиды Cas6, 0,3 мкг плазмиды Cas7 и 1 мкг плазмиды *сг*-РНК. Затем к вышеуказанной смеси ДНК добавляли смесь 5 мкл липофектамина (зарегистрированный товарный знак) 2000 (Thermo Fisher Scientific), 50 мкл Opti-MEM (Thermo Fisher Scientific) и 1 мл среды EF. После этого 1 мл полученной смеси добавляли в вышеуказанный 24-луночный планшет.

После культивирования при 37°C в атмосфере 5% CO₂ в течение 24 ч среду заменяли на 1 мл среды EF. Через 48 ч после трансфекции (через 24 ч после замены среды) клетки собирали и доводили до концентрации 1×10^4 клеток/5 мкл в PBS.

Вышеуказанные клетки нагревали при 95°C в течение 10 мин. Затем добавляли 10 мг протеиназы К и инкубировали при 55°C в течение 70 мин. Кроме того, продукт, подвергнутый термообработке при 95°C в течение 10 мин, использовали в качестве матрицы для ПЦР.

10 мкл вышеуказанной матрицы амплифицировали путем проведения 35 циклов 2-стадийной ПЦР. В данном случае праймеры, имеющие последовательности SEQ ID NO: 47 и 48, были использованы в качестве праймеров для ПЦР. Кроме того, в качестве ДНК-полимеразы использовали KOD FX (Toyobo Co., Ltd.), а процедуру 2-стадийной ПЦР проводили в соответствии с протоколом, прилагаемым к KOD FX. Продукт, амплифицированный с помощью ПЦР, очищали с использованием набора для ПЦР-очистки QIAquick (QIAGEN). Конкретная процедура была проведена в соответствии с протоколом, прилагаемым к вышеуказанному набору.

dA присоединяли к 3'-концу очищенной ДНК, полученной с использованием ДНК-полимеразы Taq (Toyobo Co., Ltd.). Очищенную ДНК подвергали электрофорезу в 2% агарозном геле и вырезали полосу размером приблизительно от 500 до 700 п.о. Затем ДНК экстрагировали из разрезанного геля и очищали с использованием набора для гель-экстракции (QIAGEN). После этого проводили клонирование TA с использованием векторной системы Easy rGEM-T (Promega), и вышеуказанную ДНК клонировали. Наконец, ДНК, клонированную щелочным препаративным методом, экстрагировали и анализировали с помощью последовательности Сэнгера. Для анализа использовали набор для секвенирования, который включал терминирующий краситель BigDye (зарегистрированный товарный знак) и реагент для проведения цикла 3.1 (Thermo Fisher Scientific), и анализатор ДНК Applied Biosystems 3730 (Thermo Fisher Scientific).

Описание эндогенной последовательности гена CCR5, которая является мишенью для системы CRISPR-Cas в этом примере, приводится на фиг. 2. На фиг. 2 экзоны указаны заглавными буквами, а интроны - строчными.

В этом примере мишень представляет собой последовательность в гене CCR5, локализованном в области 21 третьего хромосомного короткого плеча (P) (фиг. 2; вся длина последовательности нуклеиновой кислоты CCR5 показана в SEQ ID NO: 46). В частности, последовательность в экзоне 3 гена CCR5 была использована в качестве последовательности-мишени. В качестве контроля использовали последовательность-мишень Cas9, которая также находилась приблизительно в том же положении. Точнее говоря, вся подчеркнутая последовательность представляет собой последовательность-мишень для системы CRISPR-Cas типа I (AAG и следующие 32 основания), а последовательность с двойным подчеркиванием представляет собой последовательность-мишень для Cas9 (CGG и предыдущие 20 оснований). Последовательность sg-PHK была сконструирована так, чтобы она была нацелена на последовательность-мишень для системы CRISPR-Cas типа I (AAG и следующие 32 основания).

Результаты.

Результаты вышеописанного эксперимента показали, что были получены клон 1 с делетированными 401 п.о., клон 2 с делетированными 341 п.о., клон 3 с делетированными 268 п.о. и клон 4 с делетированными 344 п.о. по сравнению с исходными последовательностями нуклеиновых кислот (фиг. 3A-3D). Это указывает на то, что эндогенная ДНК человеческих клеток может быть делетирована посредством системы CRISPR-Cas3 согласно изобретению. В частности, было высказано предположение, что вышеуказанная система CRISPR-Cas позволяет редактировать ДНК человеческих клеток.

В этом примере наблюдали клоны с делециями пар оснований. Этот факт подтверждает, что расщепление ДНК посредством системы CRISPR-Cas3 согласно изобретению происходит в нескольких сайтах.

ДНК из нескольких сотен пар оснований (от 268 до 401 п.о.) была делетирована посредством системы CRISPR-Cas3 согласно изобретению. Эта делеция была более крупной, чем делеция, полученная посредством системы CRISPR-Cas с использованием Cas9 (обычно расщепляющей только в одном сайте ДНК).

Пример A-3.

Был проведен эксперимент для того, чтобы определить, может ли эндогенная ДНК человеческих клеток расщепляться системой CRISPR-Cas3 с применением метода, описанного в примере A-1.

В частности, последовательность нуклеиновой кислоты в человеческих клетках была модифицирована для экспрессии pre-sg-PHK и Cas3, Cse1 (Cas8), Cse2 (Cas11), Cas5, Cas6 и Cas7, присоединенных к сигналам локализации в ядре, и был проведен анализ на расщепление последовательности эндогенного гена EMX1 клеток.

Те же самые клетки HEK 239T, как и в примере A-1, высевали в 24-луночный планшет с плотностью 1×10^5 клеток/лунку и культивировали в течение 24 ч.

50 мкл Opti-MEM (Thermo Fisher Scientific) смешивали с 500 нг плазмиды Cas3, 500 нг плазмиды Cse1 (Cas8), 1 мкг плазмиды Cse2 (Cas11), 1 мкг плазмиды Cas5, 1 мкг плазмиды Cas6, 3 мкг плазмиды Cas7 и 500 мкг плазмиды sg-PHK. Затем к вышеуказанной смеси добавляли 4 мкл липофектамина (зарегистрированный товарный знак) 2000 (Thermo Fisher Scientific) и 50 мкл Opti-MEM (Thermo Fisher Scientific) и смешивали. Полученную смесь инкубировали при комнатной температуре в течение 20 мин, а затем добавляли к клеткам HEK 293T.

В данном случае структура экспрессионного вектора группы белков Cas, используемой в примере A-3, показана на фиг. 7. Как показано на фиг. 7, вышеупомянутый экспрессионный вектор получают путем объединения последовательности, кодирующей группу белков Cas, с BP-NLS (двухкомпонентными NLS) (см. Suzuki K. et al. (2016) In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration, Nature, Vol. 540 (Issue 7631), pp. 144-149). Аминокислотная последовательность BP-NLS представляет собой KRTADGSEFESPKKKRKVE (SEQ ID NO: 54), а последовательность нуклеиновой кислоты представляет собой

AAGCGGACTGCTGATGGCAGTGAATTTGAGTCCCCAAAGAAGAAGAGAAAGGTGG

AA (SEQ ID NO: 55).

После культивирования вышеуказанных клеток HEK 293T при 37°C в атмосфере 5% CO₂ в течение

24 ч, среду заменяли на 1 мл среды EF (1 мл на 1 лунку). Через 48 ч после трансфекции (через 24 ч после замены среды) клетки собирали и доводили до концентрации 1×10^4 клеток/5 мкл в PBS.

Вышеописанные клетки нагревали при 95°C в течение 10 мин. Затем добавляли 10 мг протеиназы К и инкубировали при 55°C в течение 70 мин. Кроме того, продукт, подвергнутый термообработке при 95°C в течение 10 мин, использовали в качестве матрицы для ПЦР.

10 мкл вышеуказанной матрицы амплифицировали путем проведения 40 циклов 3-стадийной ПЦР. В данном случае праймеры, имеющие последовательности SEQ ID NO: 50 и 51, были использованы в качестве праймеров для ПЦР. Кроме того, в качестве ДНК-полимеразы использовали "метку горячей точки" (QIAGEN), и проводили процедуру трехстадийной ПЦР в соответствии с протоколом по использованию "метки горячей точки". Продукт, амплифицированный с помощью ПЦР, подвергали электрофорезу в 2% агарозном геле и вырезали полосу от около 900 до 1100 п.о. Затем ДНК экстрагировали из разрезанного геля и очищали с использованием набора для гель-экстракции (QIAGEN). Конкретную процедуру осуществляли в соответствии с протоколом, прилагаемым к вышеуказанному набору.

После этого проводили клонирование ТА с использованием векторной системы Easy pGEM-T (Promega), и вышеуказанную ДНК клонировали. И наконец, ДНК, клонированную щелочным препаративным методом, экстрагировали и анализировали с помощью последовательности Сэнгера. Для анализа использовали набор для секвенирования, который включал терминирующий краситель BigDye (зарегистрированный товарный знак) и реагент для проведения цикла 3.1 (Thermo Fisher Scientific), и анализатор ДНК Applied Biosystems 3730 (Thermo Fisher Scientific).

Описание эндогенной последовательности гена EMX1, которая является мишенью для системы CRISPR-Cas в примере A-3, приводится на фиг. 5. На фиг. 5 экзоны указаны заглавными буквами, а интроны - строчными.

В примере A-3 мишень представляет собой последовательность в гене EMX1, локализованном в области 13 третьего хромосомного короткого плеча (P) (фиг. 5; вся длина последовательности нуклеиновой кислоты EMX1 показана в SEQ ID NO: 49). В частности, последовательность в экзоне 3 гена EMX1 была использована в качестве последовательности-мишени. В качестве контроля использовали последовательность-мишень Cas9, которая также находилась приблизительно в том же положении. Точнее говоря, вся подчеркнутая расположенная выше последовательность представляет собой последовательность-мишень для системы CRISPR-Cas типа I (AAG и следующие 32 основания), а подчеркнутая нижерасположенная последовательность представляет собой последовательность-мишень для Cas9 (TGG и предыдущие 20 оснований). Последовательность cr-PHK, используемая в примере A-3, была сконструирована так, чтобы она была нацелена на последовательность-мишень для системы CRISPR-Cas (AAG и следующие 32 основания).

Результаты.

Результаты вышеописанного эксперимента показали, что были получены клон 1 с двумя делегированными сайтами 513 и.о. и 363 и.о. и клон 2 с делегированными 694 и.о., по сравнению с исходными последовательностями нуклеиновых кислот (фиг. 6А и 6В). Эти результаты эксперимента также показали, что эндогенная ДНК человеческих клеток может быть делетирована посредством системы CRISPR-Cas3 согласно изобретению. В частности, было высказано предположение, что вышеуказанная система CRISPR-Cas позволяет редактировать ДНК человеческих клеток.

Кроме того, аналогично примеру A-2 расщепление происходило в двух или более сайтах двухцепочечной ДНК, и ДНК из нескольких сотен пар оснований была делетирована. Следовательно, результаты примера A-3 более убедительно подтверждают предположения, полученные в примере A-2.

Пример A-4.

Активность расщепления последовательности-мишени экзогенной ДНК оценивали следующим образом. Клетки НЕК 293Т получали так, чтобы они экспрессировали систему CRISPR-Cas3, в которой последовательности нуклеиновой кислоты были модифицированы, а последовательности нуклеиновой кислоты, кодирующие белки Cascade, были связаны.

В примере A-4 100 нг репортерного вектора; 200 нг каждой из плазмид Cas3, Cascade (2A) и cr-PHK; 60 нг вектора pRL-TK (способного экспрессировать люциферазу Renilla, Promega) и 300 нг вектора pBlueScript II KS (+) (Agilent Technologies) смешивали в 25 мкл Opti-MEM (Thermo Fisher Scientific). Условия использования репортерного вектора, имеющего последовательность-мишень, происходящую от CCR5, в качестве репортерного вектора, соответствуют условиям 1 в (b) на фиг. 9, а условия использования репортерного вектора, имеющего спейсерную последовательность CRISPR E. coli, соответствуют условиям 6 в (b) на фиг. 9.

В данном случае в качестве вышеописанных репортерных векторов были использованы два вида репортерных векторов, полученных в п.1 препаративного примера (т.е. векторы, имеющие структуру, показанную на фиг. 4(d)). Кроме того, в качестве вышеуказанной плазмиды Cascade (2A) были использованы экспрессионные векторы, полученные в п.4 препаративного примера (т.е. вектор, имеющий структуру, показанную на фиг. 8).

Двойной люциферазный анализ проводили методом, описанным в примере A-1, за исключением того, что использовали вышеуказанные экспрессионные векторы.

Более того, такой же эксперимент был проведен при следующих условиях в качестве контрольного эксперимента.

1. Вместо любой из плазмид Cas3 и Cascade (2A) то же самое количество вектора pBluescript II KS (+) (Agilent Technologies) смешивали для экспрессии (2 и 3 на фиг. 9).

2. Вместо плазмиды *cr*-PHK, используемой в вышеуказанной процедуре, были смешаны плазмиды для экспрессии *cr*-PHK, не комплементарной последовательности-мишени. В частности, в целях проведения экспрессии плазмиды для экспрессии *cr*-PHK, соответствующие спейсерной последовательности CRISPR *E. coli*, смешивали для нацеливания на последовательность, происходящую от гена CCR5 (4 на фиг. 9), а плазмиды для экспрессии *p*PHK, соответствующей последовательности, происходящей от гена CCR5, смешивали в случае нацеливания на спейсерную последовательность CRISPR *E. coli* (7 на фиг. 9).

3. В качестве негативного контроля экспрессировали только репортерный вектор, имеющий последовательность-мишень, происходящую от CCR5 (5 на фиг. 9), и только репортерный вектор, имеющий спейсерную последовательность CRISPR *E. coli* (8 на фиг. 9).

Результаты.

Результаты двойного анализа на люциферазу представлены на графике фиг. 9, а условия эксперимента показаны в нижней таблице на фиг. 9. На фиг. 9 "CCR5-мишень" и "спейсер-мишень" означают последовательность-мишень, происходящую от CCR5, и спейсерную последовательность CRISPR *E. coli*, соответственно. Кроме того, "CCR5-*cr*-PHK" и "спейсер-*cr*-PHK" представляют собой последовательность, комплементарную CCR5-мишени, и последовательность, комплементарную спейсеру-мишени, соответственно.

Как показано на фиг. 9, система, в которую вводили плазмиду *cr*-PHK, комплементарную последовательности-мишени, и обе плазмиды Cas3 и Cascade (2A), обладали значительно более высокой активностью расщепления, чем другие системы (при сравнении 1 и 2-5 и 6, 7 и 8). Следовательно, было обнаружено, что даже в системе, в которой последовательности нуклеиновых кислот, кодирующие Cascade, были связаны для экспрессии, последовательности, комплементарные *cr*-PHK в клетках млекопитающих, могут специфически расщепляться посредством системы CRISPR-Cas в соответствии с вариантом осуществления настоящего изобретения.

В. Исследование факторов и т.п., влияющих на редактирование генома системой CRISPR-Cas3 в эукариотической клетке.

Материал и метод.

1. Конфигурация гена Cas и *cr*-PHK.

Гены, составляющие Cas3 и Cascade (Cse1, Cse2, Cas5, Cas6 и Cas7), происходящие от штамма *E. coli* K-12, к которым были присоединены bpNLS с 5'-стороны и с 3'-стороны, были сконструированы и клонированы путем оптимизации кодонов для клеток млекопитающих с последующим синтезом генов. Эти гены были субклонированы ниже промотора CAG pPB-CAG. Плазмида EBNXN была подарена институтом Сэнгера. Мутанты Cas3, такие как H74A (инактивированная нуклеаза; dn), K320N (инактивированная геликаза; dh), и двойные мутанты S483A и T485A (инактивированная геликаза, вариант 2; dh2) получали путем аутолигирования ПЦР-продуктов Prime STAR MAX. Что касается плазмиды, экспрессирующей *cr*-PHK, то была синтезирована последовательность *cr*-PHK, имеющая два сайта рестриктирующих ферментов BbsI в положении спейсера под контролем промотора U6. Все плазмиды, экспрессирующие *cr*-PHK, получали путем встраивания двухцепочечных олигонуклеотидов последовательности-мишени с 32 парами оснований в сайты рестриктирующего фермента BbsI.

Экспрессионная плазмида Cas9-*o*crPHK pX330-U6-Chimeric_BB-CBh-hSpCas9 была получена от Addgene. При конструировании *p*PHK с использованием пакета web-программ для CRISPR, программ для конструирования CRISPR и/или CRISPRdirect были предсказаны уникальные сайты-мишени в геноме человека. Последовательность-мишень клонировали в каркас *o*crPHK pX330 в соответствии с протоколом лаборатории Feng Zhang.

Репортерная плазмида SSA, содержащая два сайта рестриктирующего фермента BsaI, была подарена профессором YAMAMOTO Takashi из Университета Хиросимы. Последовательность-мишень геномной области была встроена в сайты BsaI. В качестве люциферазного вектора Renilla был получен pRL-TK (Promega). Все плазмиды получали методом Midiprep или Maxiprep с использованием набора для очистки плазмид PureLink HiPure (Thermo Fisher).

2. Оценка активности расщепления ДНК в клетках HEK 293T.

Анализ SSA проводили, как в примере А, для детектирования активности расщепления ДНК в клетках млекопитающих. Клетки HEK 293T культивировали при 37°C в 5% CO₂ с модифицированной по способу Дульбекко средой Игла с высоким содержанием глюкозы, в которую была добавлена 10% фетальная бычья сыворотка (Thermo fisher). В лунки 96-луночного планшета высевали 0,5×10⁴ клеток. Через 24 ч плазмиды, экспрессирующие Cas3, Cse1, Cse2, Cas7, Cas5, Cas6 и *cr*-PHK (каждая по 100 нг), репортерные векторы SSA (100 нг) и векторы люциферазы Renilla (60 нг) переносили в клетки HEK 293T с использованием липофектамина 2000 и OptiMEM (Life Technologies) в соответствии с несколько измененным протоколом. Через двадцать четыре часа после трансфекции проводили двойной анализ люциферазной системы Dual-Glo (Promega) в соответствии с протоколом.

3. Детектирование INDEL в клетках HEK 293T.

В лунки 24-луночного планшета высевали $6,5 \times 10^4$ клеток. Через 24 ч плазмиды, экспрессирующие Cas3, Cse1, Cse2, Cas7, Cas5, Cas6 и cr-PHK (каждая, 250 нг), переносили в клетки HEK 293T с использованием липофектамина 2000 и OptiMEM (Life Technologies) в соответствии с несколько измененным протоколом. Через два дня после трансфекции общую ДНК экстрагировали из собранных клеток с использованием набора Tissue XS (Takara-bio Inc.) в соответствии с протоколом. Лocus-мишень амплифицировали с использованием Gflex (Takara bio Inc.) или Quick Taq HS DyeMix (TOYOBO Co., Ltd.) с последующим электрофорезом в агарозном геле. Для детектирования небольших инсерционных/делеционных мутаций в продуктах ПЦР использовали набор для детектирования мутации SURVEYOR (Integrated DNA Technologies) в соответствии с протоколом. Для клонирования TA использовали плазмидный вектор pCR4Blunt-TOPO (Life Technologies) в соответствии с протоколом. Для анализа последовательности использовали набор для секвенирования путем проведения циклов с терминацией красителем BigDye Terminator и генетический анализатор ABI PRISM 3130 (Life Technologies).

Для выявления различных необычных мутаций была получена библиотека ДНК продуктов ПЦР-амплификации с использованием набора для получения библиотеки ДНК TruSeq Nano DNA Library Prep Kit (Illumina), а секвенирование ампликона осуществляли с использованием MiSeq (2×50 п.о.) в соответствии со стандартной процедурой Macrogen. Необработанные риды образцов картировали на человеческом геноме hg38 с помощью BWA-MEM. Данные охвата были визуализированы с помощью программы визуализации Integrative Genomics Viewer (IGV), и была построена гистограмма для области-мишени.

Репортерные клетки HEK 293T, имеющие mCherry-P2A-EGFP c321C>G для детекции SNP-KI (SNIP-нокина) в клетках млекопитающих, были подарены профессором NAKADA Shin-ichiro. Репортерные клетки культивировали с 1 мкг/мл пуромицина. Однопочечную ДНК или 500 нг донорной плазмиды вводили совместно с CRISPR-Cas3 способом, описанным выше. Все клетки собирали через 5 дней после трансфекции и проводили FACS-анализ с использованием AriaIIIu (BD). GFP-позитивные клетки отсортировывали, и общую ДНК экстрагировали способом, описанным выше. Обмен SNP в геноме детектировали с помощью ПЦР-амплификации с использованием ДНК-полимеразы HiDi (myPOLS Biotec).

4. Детектирование сайтов-кандидатов, не являющихся мишенями.

Кандидаты CRISPR типа I-E, не являющиеся мишенями, детектировали в геноме человека hg38 с использованием GGenome в соответствии с двумя различными процедурами. В качестве последовательностей-кандидатов PAM были выбраны AAG, ATG, AGG, GAG, TAG и AAC в соответствии с имеющимися отчетами (Leenay, R.T, et al. Mol. Cell 62, 137-147 (2016), Jung, et al. Mol. Cell. 2017 Jung et al., Cell 170, 35-47 (2017)). Положения с меньшим количеством ошибочных спариваний были выбраны в первом подходе для 32 пар оснований последовательности-мишени, исключая положения, кратные 6, поскольку сообщалось, что такие положения не распознаются как сайты-мишени. В следующем подходе области, полностью совпадающие с 5'-концом стороны PAM последовательности-мишени, были детектированы и перечислены в порядке убывания.

5. Глубокое секвенирование с помощью анализа сайтов, не являющихся мишенями.

При секвенировании всего генома геномную ДНК экстрагировали из трансфицированных клеток HEK 293T и расщепляли с использованием ультразвукового устройства Covaris. Библиотеку ДНК получали с использованием набора для получения библиотеки ДНК без ПЦР TruSeq DNA PC-Free LT Prep Kit (Illumina), и секвенирование генома проводили с использованием HiSeq X (2×150 п.о.) в соответствии со стандартной процедурой Takara Bio Inc. Исходные риды образцов картировали на человеческом геноме hg38 BWA-MEM и идентифицировали с помощью программы Trimmomatic. Несоответствующие пары ридов и расщепленные риды были исключены с помощью программ samtools и Lumpy-sv, соответственно. В целях обнаружения только крупных делеций в одной и той же хромосоме, пары ридов, картированные на различных хромосомах, удаляли с помощью программы BadMateFilter, входящей в пакет программ для анализа генома. Общее количество несоответствующих пар ридов или расщепленных ридов в области 100 т.п.о. подсчитывали с помощью Bedtools для вычисления частоты встречаемости ошибок с негативным контролем. Специально изготовленные ДНК-зонды SureSelectXT были сконструированы с использованием SureDesign в условиях умеренной жесткости и получены методами Agilent для обогащения кандидатов, не являющихся мишенями, перед секвенированием. Области-мишени отбирали следующим образом. Зонды, находящиеся поблизости от областей-мишеней, охватывали 800 т.п.о. выше PAM и 200 т.п.о. ниже PAM. В непосредственной близости от областей CRISPR-Cas3, не являющихся мишенями, присутствовали кандидаты PAM, расположенные на 9 т.п.о. выше и на 1 т.п.о. ниже. В непосредственной близости от областей CRISPR-Cas9, не являющихся мишенями, присутствовали кандидаты PAM, расположенные на 1 т.п.о. выше и на 1 т.п.о. ниже. После получения библиотеки ДНК с использованием набора реагентов SureSelectXT и специального набора зондов, секвенирование генома проводили с помощью HiSeq 2500 (2×150 п.о.) в соответствии со стандартной процедурой Takara Bio Inc. Несоответствующие пары ридов и расщепленные риды на одной и той же хромосоме исключали методом, описанным выше. Общее количество несоответствующих пар ридов или расщепленных ридов в области 10 т.п.о. подсчитывали с помощью Bedtools для вычисления частоты встречаемости ошибок с негативным

контролем.

Пример В-1.

Влияние типов *сг*-РНК и сигнала локализации в ядре на активность расщепления ДНК.

В примере А редактирование генома в эукариотических клетках случайно было достигнуто благодаря использованию системы CRISPR-Cas3, содержащей *pre-сг*-РНК (LRSR; лидерная последовательность-повторяющаяся последовательность-спейсерная последовательность-повторяющаяся последовательность) в качестве *сг*-РНК. В данном случае авторы настоящего изобретения предположили, что причина, по которой редактирование генома в эукариотических клетках с использованием системы CRISPR-Cas3 не было успешным в течение многих лет, связана с тем, что в качестве *сг*-РНК была использована зрелая *сг*-РНК. Исходя из вышесказанного, помимо *pre-сг*-РНК (LRSR) в качестве *сг*-РНК были получены *pre-сг*-РНК (RSR; повторяющаяся последовательность-спейсерная последовательность-повторяющаяся последовательность) и зрелая *сг*-РНК (5'-фланкирующая последовательность-спейсерная последовательность-3'-фланкирующая последовательность), и эффективность редактирования генома оценивали с использованием репортерной системы примера А (фиг. 10А и 10В). Последовательности нуклеиновых кислот *pre-сг*-РНК (LRSR), *pre-сг*-РНК (RSR) и зрелой *сг*-РНК показаны в SEQ ID NO: 63, 64 и 65 соответственно.

Следовательно, если в системе CRISPR-Cas3 использовалась зрелая *сг*-РНК, то активности расщепления ДНК-мишени не наблюдалось. С другой стороны, удивительно то, что в случае использования *pre-сг*-РНК (LRSR и RSR) наблюдалась очень высокая активность расщепления ДНК-мишени. Эти результаты для системы CRISPR-Cas3 контрастируют с результатами для системы CRISPR-Cas9, которая обладала высокой активностью расщепления ДНК при использовании зрелой *сг*-РНК. Кроме того, этот факт говорит о том, что использование зрелой *сг*-РНК является одной из причин, по которой редактирование генома с использованием системы CRISPR-Cas3 в эукариотических клетках не удавалось.

Кроме того, исследование также проводили с использованием сигнала локализации в ядре SV40 и двухкомпонентного сигнала транслокации в ядро в качестве сигналов локализации в ядре, присоединенных к Cas3 (фиг. 11). В результате, при использовании двухкомпонентного сигнала транслокации в ядро наблюдалась более высокая активность расщепления ДНК-мишени.

Поэтому в следующих экспериментах *pre-сг*-РНК (LRSR) использовали в качестве *сг*-РНК, а двухкомпонентный сигнал транслокации в ядро использовали в качестве сигнала локализации в ядре.

Пример В-2.

Влияние последовательности РАМ на активность расщепления ДНК.

В целях подтверждения специфичности системы CRISPR-Cas3 к мишени было исследовано влияние различных последовательностей РАМ на активность расщепления ДНК (фиг. 12). В анализе SSA активность расщепления ДНК давала различные результаты для различных последовательностей РАМ. Наивысшая активность наблюдалась для 5'-AAG РАМ, и заметную активность также проявляли AGG, GAG, TAC, ATG и TAG.

Пример В-3.

Влияние ошибочного спаривания последовательности *сг*-РНК и спейсерной последовательности на активность расщепления ДНК.

Ранее проводимые исследования кристаллической структуры Cascade E. coli показали, что между *сг*-РНК и спейсерной ДНК образуется гетеродуплекс из 5 основных частей. Это происходит из-за нарушения спаривания оснований в каждом шестом положении посредством элемента SAM эффектора Cas7 (фиг. 13). Было оценено влияние ошибочного спаривания последовательности *сг*-РНК и спейсерной последовательности на активность расщепления ДНК. Активность расщепления резко снижалась при любом одном ошибочном спаривании в иницирующей области (положения 1-8), за исключением оснований, не распознаваемых в качестве мишени (положение 6).

Пример В-4.

Исследование необходимости участия доменов Cas3 в активности расщепления ДНК.

Характеризация *in vitro* каталитических свойств белка Cas3 показала, что N-концевой нуклеазный домен HD расщепляет одноцепочечную область субстрата ДНК, а затем геликазный домен SF2 у C-конца раскручивает ДНК-мишень АТФ-зависимым образом в направлении 3'-5'. Были получены три мутанта Cas3, а именно мутант домена HD H74A (dnCas3), мутант K320N с мотивом домена SF2 1 (dhCas3) и двойной мутант S483A/T485A с мотивом домена SF2 3 (dh2Cas3) для того, чтобы определить, необходим ли домен Cas3 для расщепления ДНК (фиг. 14). В результате, активность расщепления ДНК полностью исчезала в случае всех трех мутантов белка Cas3, что свидетельствует о том, что Cas3 может расщеплять ДНК-мишень посредством нуклеазного домена HD и геликазного домена SF2.

Пример В-5.

Исследование активности расщепления ДНК системами CRISPR-Cas3 различных типов.

Системы CRISPR-Cas3 типа 1 имеют высокую степень вариабельности (от А до G типа 1, всего семь типов). В приведенных выше примерах оценена активность расщепления ДНК в эукариотических клетках системой CRISPR-Cas3 типа I-E. С другой стороны, в этом примере оценивали активность расщепления ДНК системами CRISPR-Cas3 других типов 1 (тип I-F и тип I-G). В частности, Cas3 и Cas5-7

Shewanella putrefaciens типа I-F и Cas5-8 из *Pyrococcus furiosus* типа I-G были оптимизированы по кодам и клонированы (фиг. 15). В результате, активность расщепления ДНК была также обнаружена для этих систем CRISPR-Cas3 типа 1 в анализе SSA с использованием клеток 293Т, хотя в эффективности активности расщепления ДНК наблюдались некоторые различия.

Пример В-6.

Исследование мутаций, вводимых в эндогенные гены системой CRISPR-Cas3.

Мутации, введенные в эндогенные гены системой CRISPR-Cas3, были исследованы с использованием системы типа I-E. Ген EMX1 и ген CCR5 были выбраны в качестве генов-мишеней для получения плазмид *pre-cr-PHK* (LRSR). Клетки 293Т подвергали липофекции плазмидами, кодирующими *pre-cr-PHK* и шесть эффекторов Cas (3, 5-8 и 11). В результате, оценка CRISPR-Cas3 показала, что делеция от нескольких сотен до нескольких тысяч пар оснований происходила, главным образом, в направлении выше 5'-РАМ спейсерной последовательности области-мишени (фиг. 16). Была подтверждена микрогомология от 5 до 10 пар оснований в репарированном стыке, что, возможно, было вызвано отжигом комплементарных цепей по пути репарации, зависящего от отжига. Следует отметить, что в зрелых плазмидах *cr-PHK*, редактирование генома в областях EMX1 и CCR5 не наблюдалось.

Было отобрано девять клонов ТА и их последовательности сравнивали с последовательностями EMX1 дикого типа путем секвенирования в целях дополнительной характеристики редактирования генома посредством Cas3 путем секвенирования по Сэнгеру и клонирования ТА продуктов ПЦР (фиг. 17). В 24 из 49 клонов наблюдалась делеция как минимум 596 пар оснований, максимум 1447 пар оснований, а в среднем 985 пар оснований, что подтверждает вставку последовательности (эффективность 46,3%). Половина клонов (n=12) имела крупные делеции, включая последовательности РАМ и спейсерные последовательности, а другая половина была делетирована перед РАМ.

Дальнейшую характеристику Cas3 осуществляли путем секвенирования следующего поколения с помощью продуктов амплификации ПНП с набором праймеров в более широких областях, таких как 3,8 т.п.о. гена EMX1 и 9,7 т.п.о. CCR5. Также оценивали множество сайтов РАМ (AAG, ATG и TTT) для нацеливания с использованием CRISPR типа I-E. При секвенировании ампликона, AAG составлял 38,2%, а ATG - 56,4%. По сравнению с 86,4% TTT и 86,4% Cas9, нацеленными на EMX1, уровень охвата в широкой геномной области перед сайтом РАМ был значительно снижен. Снижение охвата было аналогичным при нацеливании на область CCR 5. В отличие от этого Cas9 индуцировал небольшие инсерции и небольшие делеции (INDEL) в сайтах-мишенях, тогда как Cas3 не имел небольших мутаций INDEL в РАМ или в сайте-мишени. Эти результаты свидетельствуют о том, что система CRISPR-Cas3 вызывает делеции в широком диапазоне перед сайтом-мишенью в клетках человека.

Принимая во внимание ограничения анализа ПЦР, такие как амплификация менее 10 т.п.о. и сильное смещение в пользу более коротких ПЦР-фрагментов, была использована последовательность для захвата на основе микромассивов 1000 т.п.о. или более вокруг локусов-мишеней EMX1 и CCR5 (фиг. 18А и 18В). Также наблюдалась делеция до 24 т.п.о. для локуса EMX1 и до 43 т.п.о. для локуса CCR5. Однако, 90% мутаций в EMX1 и 95% мутаций в CCR5 составляли менее, чем 10 т.п.о. Эти результаты позволяют предположить, что система CRISPR-Cas3 может обладать мощной нуклеазной и геликазной активностями в геноме эукариотов.

Следует отметить, что вероятность индуцирования нежелательных мутаций в геномных областях, не являющихся мишенями, создает серьезные проблемы, особенно для клинических применений, как было продемонстрировано для системы CRISPR-Cas9. Но, тем не менее, в случае использования системы CRISPR-Cas3 каких-либо значимых нежелательных эффектов не наблюдалось.

Промышленное применение

Система CRISPR-Cas3 согласно изобретению позволяет редактировать ДНК эукариотических клеток и, следовательно, может широко применяться в областях, требующих редактирования генома, таких как медицина, сельское хозяйство, лесное хозяйство и рыболовство, различные отрасли промышленности, биологические науки, биотехнология и генотерапия.

ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ получения эукариотической клетки с отредактированной эндогенной ДНК, включающий введение системы CRISPR-Cas3 в эукариотическую клетку, где система CRISPR-Cas3 включает следующие (А)-(С):

(А) белок Cas3, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид,

(В) белок Cascade, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид, и

(С) *pre-cr-PHK*, которая нацелена на эндогенную ДНК, полинуклеотид, кодирующий *pre-cr-PHK*, или экспрессионный вектор, содержащий этот полинуклеотид.

2. Способ создания животного (за исключением человека) или растения с отредактированной эндогенной ДНК, включающий

введение системы CRISPR-Cas3 животному (за исключением человека) или в растение, где система CRISPR-Cas3 включает следующее (А)-(С):

(А) белок Cas3, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид,

(В) белок Cascade, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид, и

(С) пре-cr-РНК, которая нацелена на эндогенную ДНК, полинуклеотид, кодирующий пре-cr-РНК, или экспрессионный вектор, содержащий этот полинуклеотид.

3. Способ по п.1 или 2, дополнительно включающий расщепление пре-cr-РНК, которая нацелена на эндогенную ДНК, белком, составляющим белок Cascade, после введения системы CRISPR-Cas3 в эукариотическую клетку.

4. Способ по любому из пп.1-3, где сигнал локализации в ядре присоединяют к белку Cas3 и/или к белку Cascade.

5. Способ по п.4, где сигнал локализации в ядре представляет собой двухкомпонентный сигнал локализации в ядре.

6. Применение набора в способе по любому из пп.1-5, где набор включает следующие (А)-(С):

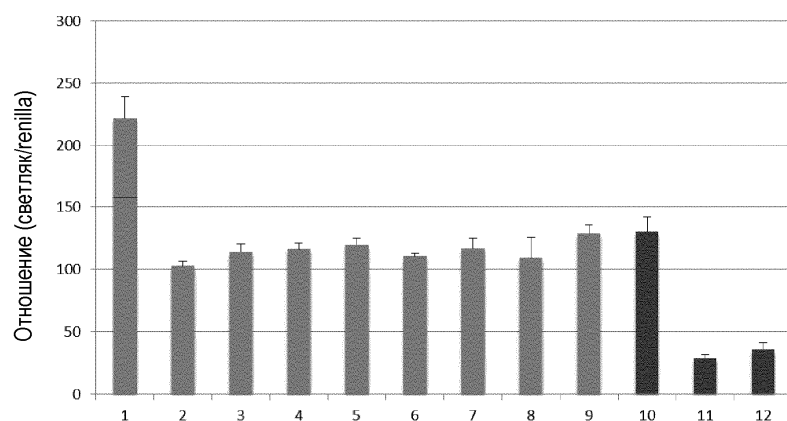
(А) белок Cas3, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид,

(В) белок Cascade, полинуклеотид, кодирующий этот белок, или экспрессионный вектор, содержащий этот полинуклеотид, и

(С) пре-cr-РНК, которая нацелена на эндогенную ДНК, полинуклеотид, кодирующий пре-cr-РНК, или экспрессионный вектор, содержащий этот полинуклеотид.

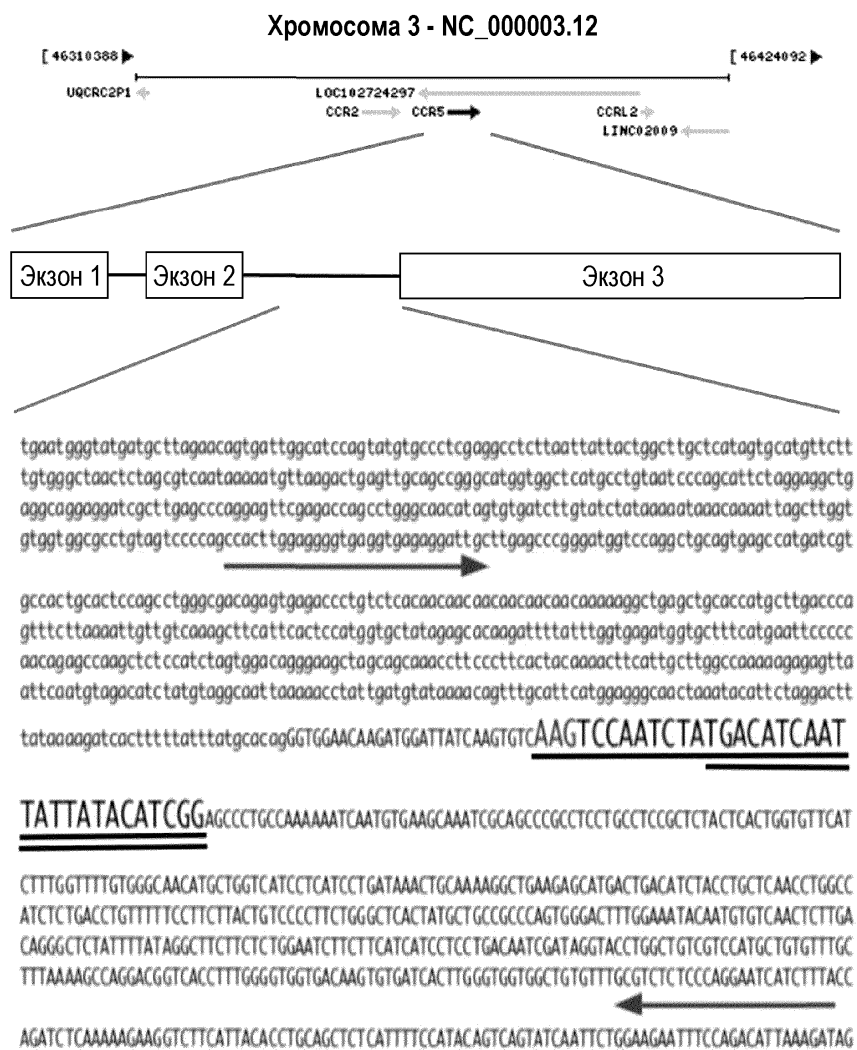
7. Применение по п.6, где сигнал локализации в ядре присоединяют к белку Cas3 и/или к белку Cascade.

8. Применение по п.7, где сигнал локализации в ядре представляет собой двухкомпонентный сигнал локализации в ядре.



Cas3	+	-	+	+	+	+	+	+	-	+	+	-
Cse1(Cas8)	+	+	-	+	+	+	+	+	-	+	+	-
Cse2(Cas11)	+	+	+	-	+	+	+	+	-	+	+	-
Cas7	+	+	+	+	-	+	+	+	-	+	+	-
Cas5	+	+	+	+	+	-	+	+	-	+	+	-
Cas6	+	+	+	+	+	+	-	+	-	+	+	-
CCR5-cr-РНК	+	+	+	+	+	+	+	-	-	-	+	-
CCR5-мишень	+	+	+	+	+	+	+	+	-	-	-	-
спейсер-cr-РНК	-	-	-	-	-	-	-	+	-	+	-	-
спейсер-мишень	-	-	-	-	-	-	-	-	-	+	+	-

Фиг. 1



Фиг. 2

Клон 1

ccacttgagggggtgaggtgagaggattgcttgagcccggtggtccaggctgcagtgagccatgatcgtg
 ccaactgcactccagcctggcgacagagtgagaccctgtctcacaacaacaacaacaacaaaaaggct
 gagctgcaccatgcttgaccagtttcttaaaattgtgtcaaagcttcattcactccatggtgctatagagca
 caagattttatttggtgagatggtgctttcatgaattcccccaacagagccaagctctccatctagtggacag
 ggaagctagcagcaaaccttcccttactacaaaacttcattgcttgccaaaaagagagttaattcaatgt
 agacatctatgtagcaattaaaaacattgatgtataaaacagtttgattcatggagggaactaaata
 cattctaggactttataaaagatcactttttatttatgcacagGGTGGAACAAGATGGATTAT
 CAAGTGTCAAGTCCAATCTATGACATCAATTATTATACATCGGAG
 CCCTGCCAAAAAATCAATGTGAAGCAAATCGCAGCCCGCCTCCTGCCTC
 CGCTCTACTCACTGGTGTTCATCTTTGGTTTTGTGGGCAACATGCTGGT
 CATCCTCATCCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATC
 TACCTGTCTAACCTGGCCATCTCTGACCTGTTTTTCTTCTTACTGTCCC
 CTTCTGGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAAATACAATG
 TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCT
 TCTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTGTCTGTCATGC
 TGTGTTTTGCTTTAAAAGCCAGGACGGTCACCTTTGGGGTGGTGACAAG
 TGTGATCACTTGGGTGGTGGCTGTGTTTGCGTCTCTCCAGGAATCATC
 TTTA

- : Делеция в 401 п.о. (для последних СА, которые расположены
 непосредственно перед делецией)

— : РАМ + последовательность-мишень

Фиг. 3А

Клон 2

ccacttgagggggtgaggtgagaggattgcttgagcccggtggtccaggctgcagtgagccatgatcgtg
 ccaactgcactccagcctggcgacagagtgagaccctgtctcacaacaacaacaacaacaaaaaggct
 gagctgcaccatgcttgaccagtttcttaaaattgtgtcaaagcttcattcactccatggtgctatagagca
 caagattttatttggtgagatggtgctttcatgaattcccccaacagagccaagctctccatctagtggacag
 ggaagctagcagcaaaccttcccttactacaaaacttcattgcttgccaaaaagagagttaattcaatgt
 agacatctatgtagcaattaaaaacattgatgtataaaacagtttgattcatggagggaactaaata
 cattctaggactttataaaagatcactttttatttatgcacagGGTGGAACAAGATGGATTAT
 CAAGTGTCAAGTCCAATCTATGACATCAATTATTATACATCGGAG
 CCCTGCCAAAAAATCAATGTGAAGCAAATCGCAGCCCGCCTCCTGCCTC
 CGCTCTACTCACTGGTGTTCATCTTTGGTTTTGTGGGCAACATGCTGGT
 CATCCTCATCCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATC
 TACCTGTCTAACCTGGCCATCTCTGACCTGTTTTTCTTCTTACTGTCCC
 CTTCTGGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAAATACAATG
 TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCT
 TCTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTGTCTGTCATGC
 TGTGTTTTGCTTTAAAAGCCAGGACGGTCACCTTTGGGGTGGTGACAAG
 TGTGATCACTTGGGTGGTGGCTGTGTTTGCGTCTCTCCAGGAATCATC
 TTTA

- : Делеция в 341 п.о. (для последних АС, которые расположены
 непосредственно перед делецией)

— : РАМ + последовательность-мишень

Фиг. 3В

Клон 3

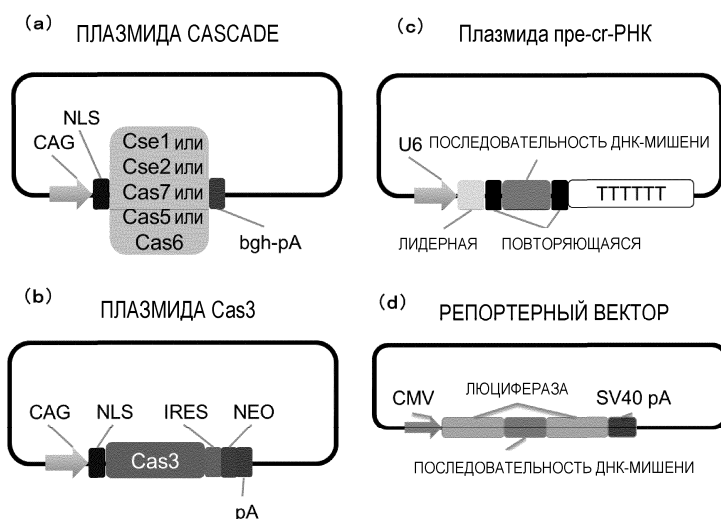
ccacttgagggtgaggtgagaggattgcttgagccgggatgggtccaggctgcagtgagccatgatcgtg
 ccactgcactccagcctggcgacagagtgagaccctgtctcacaacaacaacaacaacaaaaaggct
 gagctgcaccatgcttgaccagtttcttaaaattgttgcaaaagcttcattcactccatggtgctatagagca
 caagattttatttggtgagatgggtgctttcatgaattccccaacagagccaagctctccatctagtgacag
 ggaagctagcagcaaaccttcccttactacaaaacttcattgcttgccaaaaagagagttaattcaatgt
 agacatctatgtaggaattaaaaacattgatgtataaaacagtttgattcatggagggaactaaata
 cattctaggactttataaaagatcactttttattatgcacagGGTGGAACAAGATGGATTAT
 CAAGTGTCCAAGTCCAATCTATGACATCAATTATTATACATCGGAG
 CCCTGCTAAAAAAATCAATGTGAAGCAAATCGCAGCCCGCCTCCTGCCT
 CCGTCTACTCACTGGTGTTCATCTTTGGTTTTGTGGGCAACATGCTGG
 TCATCCTCATCCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACAT
 CTACCTGCTCAACCTGGCCATCTCTGACCTGTTTTTCCTTCTTACTGTCC
 CCTTCTGGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAAATACAAT
 GTGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATC
 TTCTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTGTCTGCCATG
 CTGTGTTTGCTTTAAAAGCCAGGACGGTCACCTTTGGGGTGGTGACAA
 GTGTGATCACTTGGGTGGTGGCTGTGTTTGCGTCTCTCCAGGAATCAT
 CTTTA

- : Делеция в 268 п.о. (отсутствие гомологии до и после делеции)
 — : PAM + последовательность-мишень
 = : C>TA: Деблокированная форма импульса SEQ
 Фиг. 3С

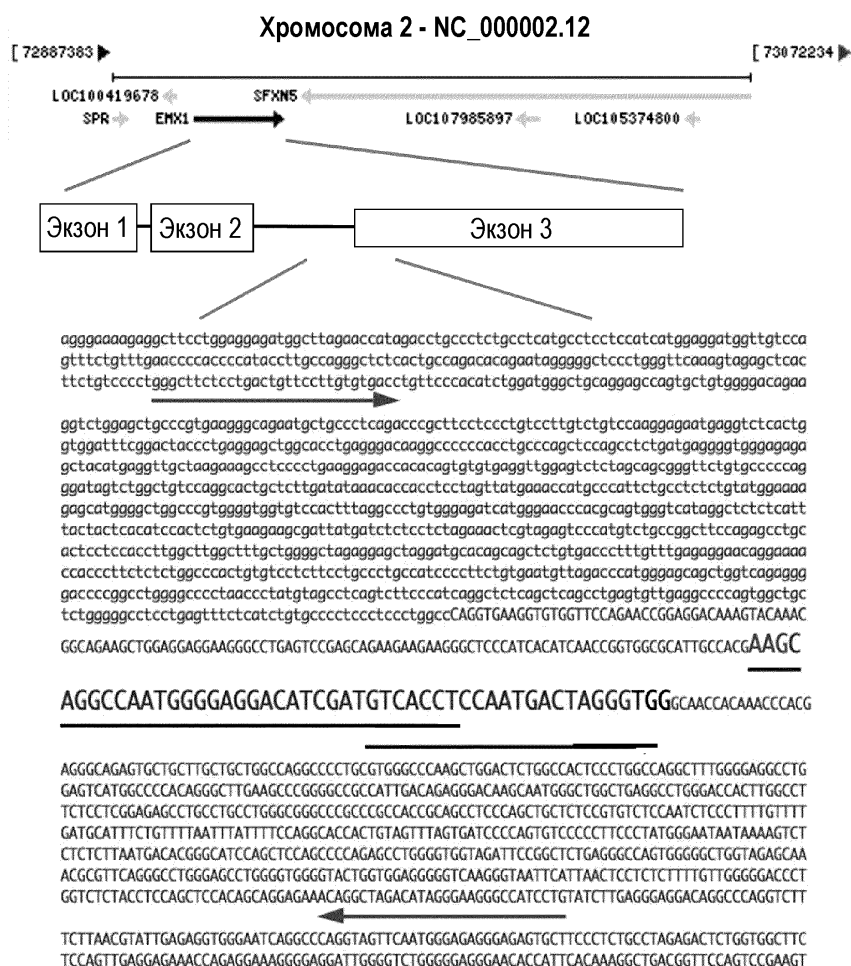
Клон 4

ccacttgagggtgaggtgagaggattgcttgagccgggatgggtccaggctgcagtgagccatgatcgtg
 ccactgcactccagcctggcgacagagtgagaccctgtctcacaacaacaacaacaacaaaaaggct
 gagctgcaccatgcttgaccagtttcttaaaattgttgcaaaagcttcattcactccatggtgctatagagca
 caagattttatttggtgagatgggtgctttcatgaattccccaacagagccaagctctccatctagtgacag
 ggaagctagcagcaaaccttcccttactacaaaacttcattgcttgccaaaaagagagttaattcaatgt
 agacatctatgtaggaattaaaaacattgatgtataaaacagtttgattcatggagggaactaaata
 cattctaggactttataaaagatcactttttattatgcacagGGTGGAACAAGATGGATTAT
 CAAGTGTCCAAGTCCAATCTATGACATCAATTATTATACATCGGAG
 CCCTGCCAAAAATCAATGTGAAGCAAATCGCAGCCCGCCTCCTGCCTC
 CGTCTACTCACTGGTGTTCATCTTTGGTTTTGTGGGCAACATGCTGGT
 CATCCTCATCCTGATAAACTGCAAAAGGCTGAAGAGCATGACTGACATC
 TACCTGCTCAACCTGGCCATCTCTGACCTGTTTTTCCTTCTTACTGTCCC
 STTCTGGGCTCACTATGCTGCCGCCAGTGGGACTTTGGAAATACAATG
 TGTCAACTCTTGACAGGGCTCTATTTTATAGGCTTCTTCTCTGGAATCT
 TCTTCATCATCCTCCTGACAATCGATAGGTACCTGGCTGTCTGCCATGC
 TGTGTTTGCTTTAAAAGCCAGGACGGTCACCTTTGGGGTGGTGACAAG
 TGTGATCACTTGGGTGGTGGCTGTGTTTGCGTCTCTCCAGGAATCATC
 TTTA

- : Делеция в 344 п.о. (отсутствие гомологии до и после делеции)
 Делетированная последовательность была заменена CATCTACAT
 — C>TA: Деблокированная форма импульса SEQ
 Фиг. 3D



Фиг. 4



Фиг. 5

Клон 1

GGGCTTCTCCTGACTGTTCCCTTGTGTGACCTGTTCCACATCTGGATGGGCTGCAGGA
 GCCAGTGCTGTGGGACAGAAGGTCTGGAGCTGCCCCTGAAGGGCAGAATGCTGCCC
 TCAGACCCGCTTCTCCCTGTCTTGTCTGTCCAAGGAGAATGAGGTCTCACTGGTGG
 ATTTGGGACTACCTGAGGAGCTGGCACCTGAGGGACAAGGCCCCCACTGCCCAGC
 TCCAGCCTCTGATGAGGGGTGGGAGAGAGCTACATGAGGTTGCTAAGAAAGCCTCCCC
 TGAAGGAGACCACACAGTGTGTGAGGTTGGAGTCTCTAGCAGCGGGTTCTGTGCCCCC
 AGGGATAGTCTGGCTGTCCAGGCACTGCTCTGATATAAACACCACCTCCTAGTTATGA
 AACCATGCCCATTCTGCCCTCTGTATGAAAAAGAGCATGGGGCTGGCCCGTGGGGTG
 GTGTCCACTTTAGGCCCTGTGGGAGATCATGGGAACCCACGCAGTGGGTCTAGGCTC
 TCTCATTTACTACTCACATCCACTCTGTGAAGAAGCGATTATGATCTCTCCTCTAGAAA
 CTCGTAGAGTCCCATGTCTGCCGGCTTCCAGAGCCTGCACTCCTCCACCTTGGCTTGG
 CTTTGCTGGGGCTAGAGGAGCTAGGATGCACAGCAGCTCTGTGACCCTTTGTTTGA
 GGAACAGGAAAAACACCCCTTCTCTGGGCCACTGTGTCCTCTTCTGCCCCTGCCATCC
 CCTTCTGTGAATGTTAGACCCATGGGAGCAGCTGGTCAGAGGGGACCCCGGCTGGG
 GCCCCTAACCTATGTAGCCTCAGTCTTCCATCAGGCTCTCAGCTCAGCCTGAGTGT
 GAGGCCCCAGTGGCTGCTCTGGGGCCTCCTGAGTTTCTCATCTGTGCCCTCCCTCC
 CTGGCCCAAGGTGAAGGTGTGGTTCCAGAACCGGAGGACAAAGTACAAAAGGCAGAAG
 CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGGCTCCCATCACATCAAC
 CGGTGGCGCATTGCCACGAAGCAGGCCAATGGGGAGGACATCGATGTCACCTCCAATG
 ACTAGGGTGGGCAACCACAAACCCACGAGGGCAGAGTGCTGCTTGCTGCTGGCCAGG
 CCCCCTGCGTGGGCCCCAAGCTGGACTCTGGCCACTCCCTGGCCAGGCTTTGGGGAGGCC
 TGGAGTCATGGCCCCACAGGGCTTGAAGCCCGGGGCCGCCATTGACAGAGGGACAAG
 CAATGGGCTGGCTGAGGCCCTGGGACCACCTTGGCCTTCTCCTCGGAGAGCCTGCTGCCC
 TGGGCGGGCCCGCCCGCCACCGCAGCCTCCAGCTGCTCTCCGTGTCTCCAATCTCCC
 TTCTGTCTTGATGCATTTCTGTTTAAATTATTTTCCAGGCACCACTGTAGTTTAGTGAT
 CCCCAGTGTCCTCCCTTCCCTATGGGAATAATAAAAGTCTCTCTCTTAATGACACGGGCA
 TCCAGCTCCAGCCCCAGAGCCTGGGGTGGTAGATTCCGGCTCTGAGGGCCAGTGGGG
 GCTGGTAGAGCAAAACCGCTTCAGGGCCTGGGAGCCTGGGGTGGGGTACTGGTGGAGG
 GGGTCAAGGGTAATTCAATTAACCTCTCTCTTTTGTGGGGGACCCTGGTCTCTACCTCC
 AGCTCCACAGCAGGAGAAAACAGGCTAGACATAGGGAAGGGCCATCCTG

□ : Делеция 513 п.о., 363 п.о., микрогомология СТС, Т

— : РАМ + последовательность-мишень

Фиг. 6А

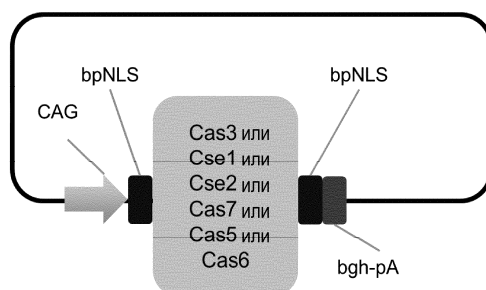
Клон 2

GGGCTTCTCCTGACTGTTCTTGTGTGACCTGTTCCACATCTGGATGGGCTGCAGGA
 GCCAGTGTGTGGGACAGAAGGTCTGGAGCTGCCCCTGAAGGGCAGAATGCTGCCC
 TCAGACCCGCTTCCTCCTGTCTTGTCTGTCCAAGGAGAATGAGGTCTCACTGGTGG
 ATTTGGACTACCTGAGGAGCTGGCACCTGAGGGACAAGGCCCCCACCTGCCCAGC
 TCCAGCCTCTGATGAGGGGTGGGAGAGAGCTACATGAGGTTGCTAAGAAAGCCTCCCC
 TGAAGCAGACCCACACAGTGTGTGAGGTTGAGTCTCTAGCAGCGGCTTCTGTGCCCCC
 AGGGATAGTCTGGCTGTCCAGGCACCTGCTCTTGATATAAACACACCTCTAGTTATGA
 AACCATGCCATTCTGCTCTCTGTATGAAAAGAGCATGGGGCTGGCCCGTGGGGTG
 GTGTCCACTTTAGGCCCTGTGGGAGATCATGGGAACCCACGCAGTGGGTATAGGCTC
 TCTCATTTACTACTACATCCACTCTGTGAAGAAGCGATTATGATCTCTCCTCTAGAAA
 CTCGTAGAGTCCCATGTCTGCGGCTTCCAGAGCCTGCACCTCTCCACCTTGGCTTGG
 CTTTGTGGGGCTAGAGGAGCTAGGATGCACAGCAGCTCTGTGACCTTTGTTTGAGA
 GGAACAGGAAAAACACCTTCTCTGTGCCCACTGTGTCTCTCTCTGCCCCGCCATCC
 CCTTCTGTGAATGTTAGACCCATGGGAGCAGCTGGTCAGAGGGGACCCCGGCTGGG
 GCCCCTAACCCCTATGTAGCCTCAGTCTTCCCATCAGGCTCTCAGCTCAGCCTGAGTGTT
 GAGCCCCCAGTGGCTGCTCTGGGGGCCCTCTGAGTTTCTCATCTGTGCCCTCCCTCC
 CTGGCCSAGCTGAAGGTGTGGTTCCAGAACCAGGAGGACAAAGTACAAACGCGAGAAG
 CTGGAGGAGGAAGGGCCTGAGTCCGAGCAGAAGAAGAAGGCTCCCATCACATCAAC
 CGGTGGCGCATTTGCCACGAAGCAGGCCAATGGGGAGGACATCGATGTCACCTCCAATG
 ACTAGGGTGGGCAACCACAAACCCACGAGGGCAGAGTGTGCTTGTGCTGCTGGCCAGG
 CCCCTGCGTGGGCCCCAAGCTGGACTCTGGCCACTCCCTGGCCAGGCTTTGGGGAGGCC
 TGAGTCAATGGCCACAGGGCTTGAAGCCCGGGGCCCGCCATTGACAGAGGGACAAG
 CAATGGGCTGGCTGAGGGCTGGGACCCTTGGCTTCTCTCGAGAGCCTGCTGCC
 TGGGCGGGCCCGCCGCCACCGCAGCCTCCAGCTGCTCTCGGTGTCTCCAATCTCCC
 TTTTGTTTTATGCAATTTCTGTTTAAATTTATTTTCCAGGCACCACTGTAGTTTAGTGAT
 CCCCAGTGTCTCCCTTCCCTATGGGAATAATAAAAGTCTCTCTCTTAATGACACGGGCA
 TCCAGCTCCAGCCCCAGAGCCTGGGGTGGTAGATTCCGGCTCTGAGGGCCAGTGGGG
 GCTGCTAGAGCAACCGCTTCAGGGCCTGGGAGCCTGGGGTGGGGTACTGGTGGAGG
 GGGTCAAGGTAATTCATTAACTCTCTCTTTTGTGGGGACCTGGTCTCTACCTCC
 AGCTCCACAGCAGGAGAAACAGGCTAGACATAGGGAAGGGCCATCCTG

- : Делеция 694 п.о., микрогомология
 — : РАМ + последовательность-мишень

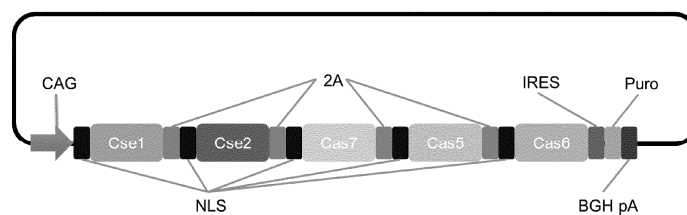
Фиг. 6В

ПЛАЗМИДА CASCADE bpNLS

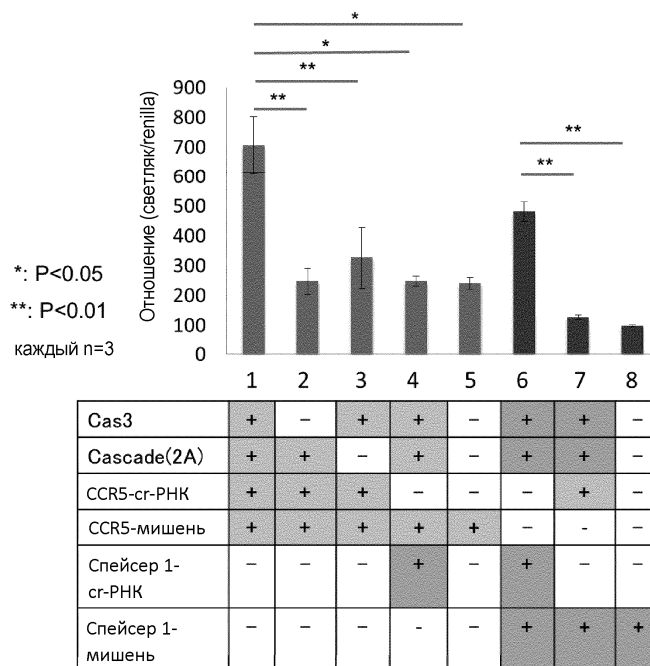


Фиг. 7

Плазмида Cascade (2A)



Фиг. 8



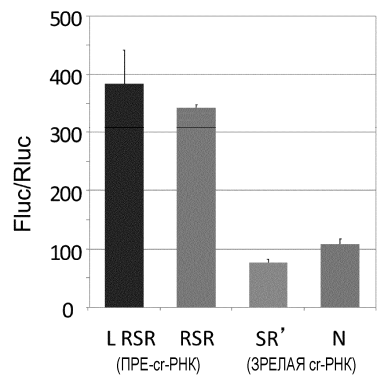
Фиг. 9

ПРЕ-cr-PHK (LRSR) : TGGATGTGTTGTTTGTGTGATACTATAAAGTTGGT
AGATTGTGACTGGCTTAAAAAATCATTAATTAATAATAGGTTATGTTTAG
A|GTGTTCCCCGCGCCAGCGGGGATAAACCG|CAGGCCAATGGGGAGG
ACATCGATGTCACCTC|GTGTTCCCCGCGCCAGCGGGGATAAACCG

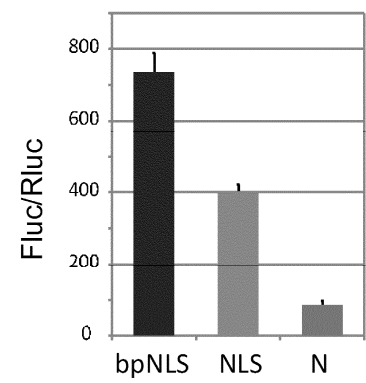
ПРЕ-cr-PHK (RSR) : GTGTTCCCCGCGCCAGCGGGGATAAACCG|CAGGC
CAATGGGGAGGACATCGATGTCACCTC|GTGTTCCCCGCGCCAGCGGG
GATAAACCG

ЗРЕЛАЯ cr-PHK : ATAAACCG|CAGGCCAATGGGGAGGACATCGATGTCACCT
C|GTGTTCCCCGCGCCAGCGGGG

Фиг. 10А

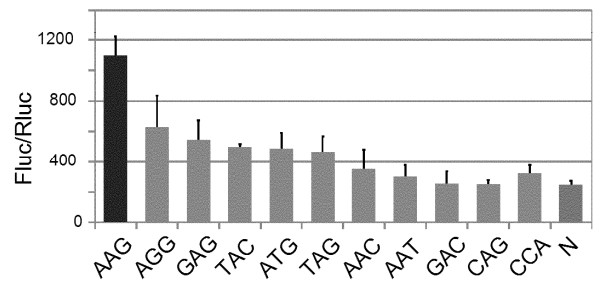


Фиг. 10В

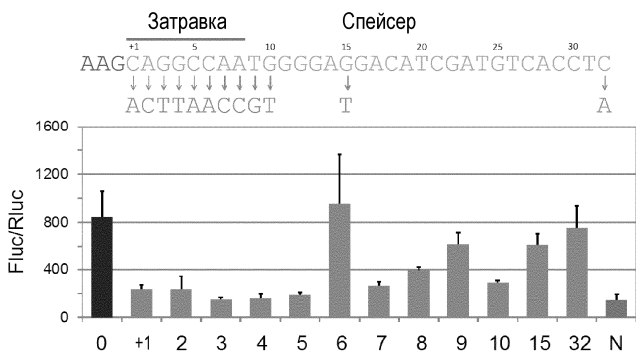


Фиг. 11

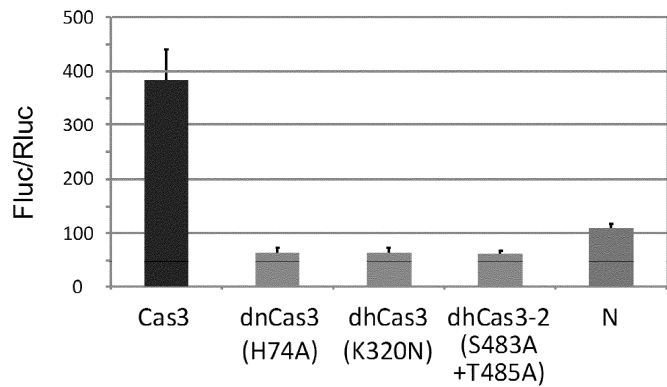
PAM
AAGCAGGCCAATGGGGAGGACATCGATGTCACCTC



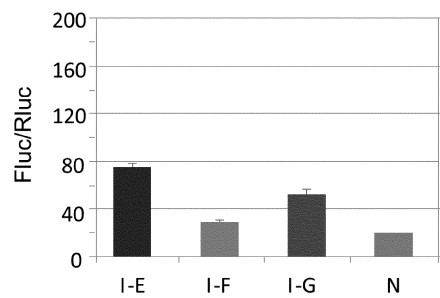
Фиг. 12



Фиг. 13



Фиг. 14

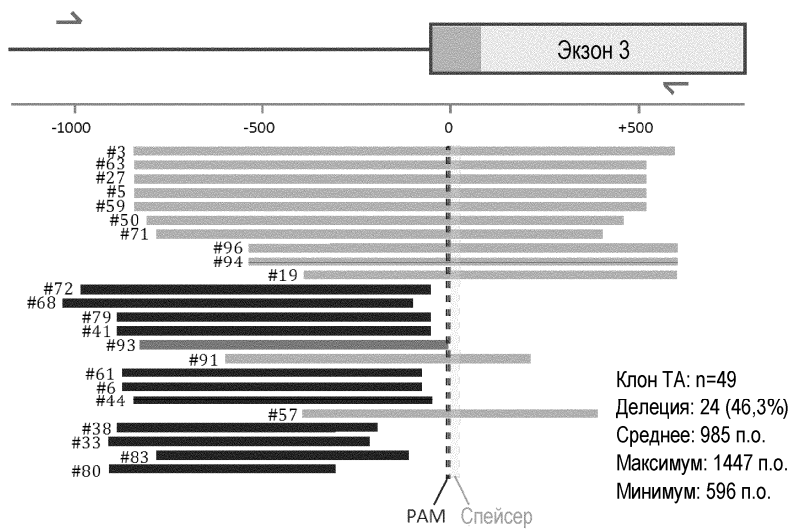


Фиг. 15

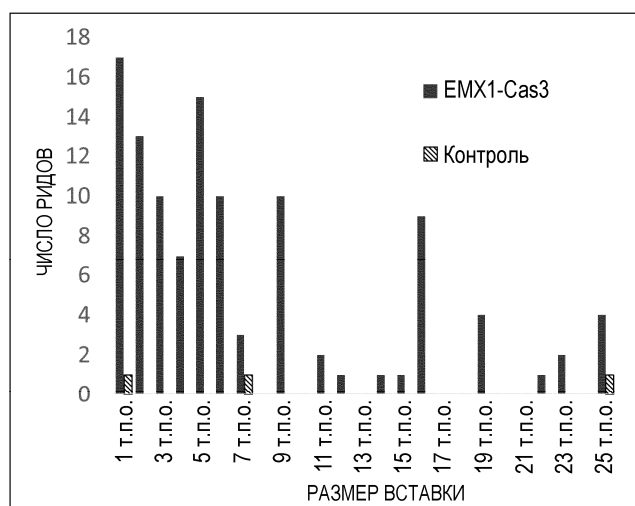
Emx1 1,7 т.п.о.

WT	TGC//TCACATCAACCGGTGGCGCATTTGCCACG AAGCAG//CTT
#3	TGC//CTGCC --- (-1447 п.о.) --- CTACC//CTT
#5	TGC//GCCCA --- (-1370 п.о.) --- CCTGG//CTT
#6	TGC//GAGCT --- (-797 п.о.) --- GAGGA//CGAAGCAG//CTT
#19	TGC//ACCCT --- (-994 п.о.) --- CAGCT//CTT
#27	TGC//GCCCA --- (-1370 п.о.) --- CCTGG//CTT
#33	TGC//TGAGG --- (-690 п.о.) --- TCTCA//CGAAGCAG//CTT
#38	TGC//GGACT --- (-690 п.о.) --- AGGCC//CGAAGCAG//CTT
#41	TGC//GGACT --- (-834 п.о.) --- AAGAA//CGAAGCAG//CTT
#93	TGC//TGATG --- (-820 п.о.) --- CAGGCCAATGGG//CTT

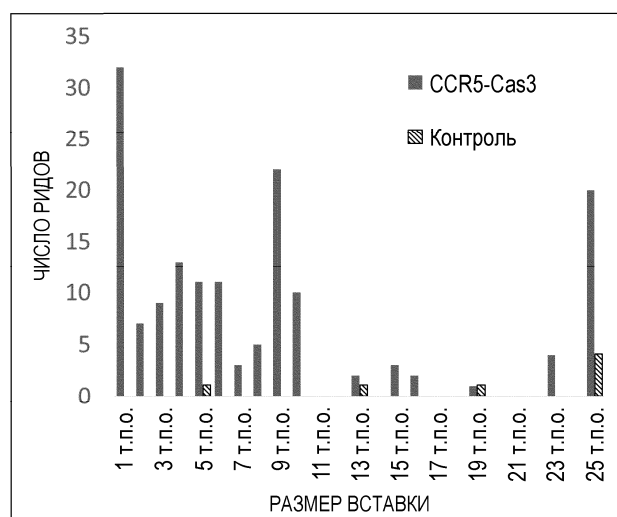
Фиг. 16



Фиг. 17



Фиг. 18А



Фиг. 18В

