

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0044164

(43) 공개일자 2021년04월22일

(51) 국제특허분류(Int. Cl.)

C12N 9/22 (2006.01) C12N 15/10 (2017.01)

C12N 15/113 (2010.01) C12Q 1/34 (2006.01)

(52) CPC특허분류

C12N 9/22 (2013.01)

C12N 15/102 (2013.01)

(21) 출원번호 10-2020-0131811

(22) 출원일자 2020년10월13일

심사청구일자 2020년10월13일

(30) 우선권주장

1020190127304 2019년10월14일 대한민국(KR)

(71) 출원인

연세대학교 산학협력단

서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)

(72) 발명자

김형범

서울특별시 마포구 토정로18길 11, 107동 1702호 (현석동, 래미안웰스트림)

김희권

서울특별시 서대문구 연희로10가길 47, 303호(연희동)

김나혜

경상남도 통영시 원문로 52, 701호(북신동, 광우화목맨션)

(74) 대리인

리엔목특허법인

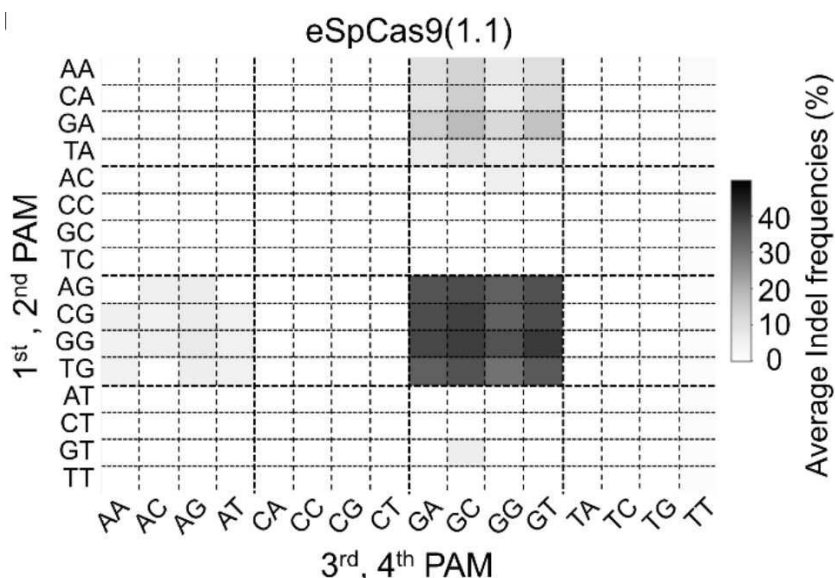
전체 청구항 수 : 총 16 항

(54) 발명의 명칭 신규 프로토스페이스 인접 모티프 서열 및 이를 이용한 세포의 유전체에서 표적 핵산을 변형시키는 방법

(57) 요약

신규한 PAM 서열을 이용하는 세포의 유전체에서 표적 핵산을 변형시키는 방법 및 이에 의해 유전체의 표적 핵산이 변형된 세포를 제공한다. 이에 의하면, 기존에 표적으로 하지 못한 위치를 유전체 편집의 표적으로 하여 유전체 편집을 수행할 수 있고, 이에 의해 유전체 편집의 활용 범위를 확대시킬 수 있다.

대표도 - 도4



(52) CPC특허분류

C12N 15/113 (2013.01)

C12Q 1/34 (2013.01)

C12N 2310/20 (2017.05)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711083110
과제번호	2017R1A2B3004198
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(과기정통부)(R&D)
연구과제명	크리스퍼 유전자가위의 활성화에 영향을 미치는 인자 규명 및
대량산출(high-throughput)방법을 이용한 유전학 연구 기초 기술 개발	
기 여 율	40/100
과제수행기관명	연세대학교
연구기간	2019.03.01 ~ 2020.02.29

이 발명을 지원한 국가연구개발사업

과제고유번호	1711082364
과제번호	2017M3A9B4062403
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	바이오.의료기술개발(R&D)
연구과제명	생체 내 유전자 교정을 통한 근육 및 안 질환 치료 기술 개발
기 여 율	30/100
과제수행기관명	연세대학교
연구기간	2019.01.30 ~ 2019.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1465027621
과제번호	HI17C0676000019
부처명	보건복지부
과제관리(전문)기관명	한국보건산업진흥원
연구사업명	질환극복기술개발(R&D)
연구과제명	효율적인 생체 내 유전자 수술 방법 개발을 통한 유전성 간질환 치료법 발굴
기 여 율	20/100
과제수행기관명	연세대학교 산학협력단
연구기간	2019.01.01 ~ 2019.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711091674
과제번호	2018R1A5A2025079
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	집단연구지원(R&D)
연구과제명	만성난치질환 시스템의학 연구센터
기 여 율	5/100
과제수행기관명	연세대학교
연구기간	2019.03.01 ~ 2020.02.29

이 발명을 지원한 국가연구개발사업

과제고유번호	1465029139
과제번호	HI16C1012010019
부처명	보건복지부
과제관리(전문)기관명	한국보건산업진흥원
연구사업명	임상연구인프라조성(R&D)
연구과제명	유전성 대사성 뇌백질 질환 치료용 정확성과 효율성 높은 유전자가위 제작 및 전달
법 개발	
기 여 율	5/100
과제수행기관명	연세대학교 산학협력단
연구기간	2019.01.01 ~ 2019.12.31

명세서

청구범위

청구항 1

세포의 유전체에서 표적 핵산을 변형시키는 방법으로서, 상기 방법은

표적 핵산을 포함한 세포, Cas(CRISPR associated) 뉴클레아제 또는 이의 변이체를 암호화하는 폴리뉴클레오티드, 및 가이드 RNA를 인큐베이션하는 단계를 포함하고,

상기 표적 핵산은 프로토스페이스-인접 모티프(protospacer adjacent motif: PAM), 및 상기 가이드 RNA에 상보적인 표적 서열을 포함하고,

상기 PAM은 5'-NAGN-3', 5'-ACGG-3', 5'-AGAS-3', 5'-WGGN-3', 5'-SGRN-3', 5'-TGAD-3', 5'-GTGC-3', 5'-RAGN-3', 5'-CAGH-3', 5'-TAGB-3', 5'-VGAG-3', 5'-NGGN-3', 5'-AAGH-3', 5'-CAGY-3', 5'-GAGN-3', 5'-VAAG-3', 5'-WGRN-3', 5'-RGY-3', 5'-SGAD-3', 5'-SGGN-3', 5'-TGTG-3', 5'-YGCG-3', 5'-NYGG-3', 5'-NGRN-3', 5'-SGCC-3', 5'-MTGC-3', 5'-GTGH-3', 5'-RCAG-3', 5'-VTAG-3', 5'-WGCC-3', 5'-SACG-3', 5'-NGCD-3', 5'-NARG-3', 5'-WAGA-3', 5'-GAGT-3', 5'-NGTK-3', 5'-NGCN-3', 5'-NACG-3', 5'-AGGY-3', 5'-NGDG-3', 5'-GAGG-3', 5'-GGGT-3', 5'-VYAG-3', 5'-GAAT-3', 5'-NGVN-3', 5'-VACG-3', 5'-AYGA-3', 5'-GAGH-3', 5'-TGTN-3', 5'-AGTB-3', 5'-SGTK-3', 5'-GATG-3', 5'-NGAN-3', 5'-NAAG-3', 5'-NGCG-3', 5'-NGGR-3', 5'-DGCG-3', 5'-NGGT-3', 5'-NGTG-3', 및 5'-GGAG-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 것이고,

상기 표적 핵산은 상기 Cas 뉴클레아제 또는 이의 변이체, 및 상기 가이드 RNA의 복합체에 의해 인식될 수 있고,

상기 Cas 뉴클레아제 또는 이의 변이체, 및 가이드 RNA의 복합체는 표적 핵산을 서열-특이적으로 변형시키는 것인 방법.

청구항 2

청구항 1에 있어서, 상기 Cas 뉴클레아제 변이체는 eSpCas9(1.1), SpCas9-HF1, HypaCas9, evoCas9, Sniper-Cas9, SpCas9 VQR 변이체, SpCas9 VRER 변이체, SpCas9 VRQR 변이체, SpCas9 VRQR-HF1 변이체, 및 SpCas9 QQR1 변이체로 이루어진 군으로부터 선택된 세균으로부터 유래된 뉴클레아제인 것인 방법.

청구항 3

청구항 1에 있어서, 상기 Cas 뉴클레아제 변이체가 eSpCas9(1.1)인 경우, 상기 PAM은 5'-NAGN-3', 5'-ACGG-3', 5'-AGAS-3', 5'-WGGN-3', 5'-SGRN-3', 5'-TGAD-3', 및 5'-GTGC-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드인 것인 방법.

청구항 4

청구항 1에 있어서, 상기 Cas 뉴클레아제 변이체가 SpCas9-HF1인 경우, 상기 PAM은 5'-RAGN-3', 5'-CAGH-3', 5'-TAGB-3', 5'-VGAG-3', 및 5'-NGGN-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드인 것인 방법.

청구항 5

청구항 1에 있어서, 상기 Cas 뉴클레아제 변이체가 HypaCas9인 경우, 상기 PAM은 5'-AAGH-3', 5'-CAGY-3', 5'-GAGN-3', 및 5'-NGGN-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드인 것인 방법.

청구항 6

청구항 1에 있어서, 상기 Cas 뉴클레아제 변이체가 evoCas9인 경우, 상기 PAM은 5'-VAAG-3', 5'-WGRN-3', 5'-RGY-3', 5'-SGAD-3', 5'-SGGN-3', 5'-TGTG-3', 및 5'-YGCG-3'으로 이루어진 군으로부터 선택된 뉴클레오티드

서열로 이루어진 폴리뉴클레오티드인 것인 방법.

청구항 7

청구항 1에 있어서, 상기 Cas 뉴클레아제 변이체가 Sniper-Cas9인 경우, 상기 PAM은 5'-NAGN-3', 5'-NYGG-3', 5'-NGRN-3', 5'-SGCC-3', 5'-MTGC-3', 및 5'-GTGH-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드인 것인 방법.

청구항 8

청구항 1에 있어서, 상기 Cas 뉴클레아제 변이체가 SpCas9 VQR 변이체인 경우, 상기 PAM은 5'-RCAG-3', 5'-VTAG-3', 5'-WGCC-3', 5'-SACG-3', 5'-NGCD-3', 5'-NGRN-3', 5'-NARG-3', 5'-WAGA-3', 5'-GAGT-3', 및 5'-NGTK-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드인 것인 방법.

청구항 9

청구항 1에 있어서, 상기 Cas 뉴클레아제 변이체가 SpCas9 VRER 변이체인 경우, 상기 PAM은 5'-NGCN-3', 5'-NACG-3', 5'-AGGY-3', 5'-NGDG-3', 5'-GAGG-3', 및 5'-GGGT-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드인 것인 방법.

청구항 10

청구항 1에 있어서, 상기 Cas 뉴클레아제 변이체가 SpCas9 VRQR 변이체인 경우, 상기 PAM은 5'-VYAG-3', 5'-GAAT-3', 5'-NGVN-3', 5'-VACG-3', 5'-WAGA-3', 5'-AYGA-3', 5'-GAGH-3', 5'-NARG-3', 5'-TGTN-3', 5'-AGTB-3', 5'-SGTK-3', 및 5'-GATG-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드인 것인 방법.

청구항 11

청구항 1에 있어서, 상기 Cas 뉴클레아제 변이체가 SpCas9 VRQR-HF1 변이체인 경우, 상기 PAM은 5'-NGAN-3', 5'-NAAG-3', 5'-RCAG-3', 5'-VTAG-3', 5'-NGCG-3', 5'-NGGR-3', 5'-DGGC-3', 5'-GAGG-3', 5'-NGGT-3', 및 5'-NGTG-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드인 것인 방법.

청구항 12

청구항 1에 있어서, 상기 Cas 뉴클레아제 변이체가 SpCas9 QQR1 변이체인 경우, 상기 PAM은 5'-NAAG-3' 및 5'-GGAG-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드인 것인 방법.

청구항 13

청구항 1에 있어서, 상기 가이드 RNA는 상기 표적 핵산에서 PAM의 5' 방향 또는 3' 방향으로 연속적인 2 내지 24 뉴클레오티드 서열과 상보적인 폴리뉴클레오티드인 것인 방법.

청구항 14

청구항 1에 있어서, 상기 가이드 RNA의 길이는 17 뉴클레오티드 내지 24 뉴클레오티드인 것인 방법.

청구항 15

청구항 1에 있어서, 상기 변형은 절단, 삽입, 연결, 탈아미노화, 또는 이들의 조합인 것인 방법.

청구항 16

청구항 1의 방법으로 유전체의 표적 핵산이 변형된 세포.

발명의 설명

기술 분야

신규한 프로토스페이스 인접 모티프 서열을 이용한 세포의 유전체 편집 방법에 관한 것이다.

[0001]

배경 기술

- [0002] CRISPR-Cas9 시스템은, 세균 및 고세균에서 발견되는 적응적 면역 시스템으로, 다양한 생물 및 인간 세포를 포함한 다양한 세포의 유전체 편집에 이용되고 있다. 그러나, 표적 DNA 영역에 이용가능한 프로토스페이서 인접 모티프(proto-spacer adjacent motif: PAM)이 없거나, 불충분한 온-타겟(on-target) 활성, 또는 오프-타겟(off-target) 효과 등으로 인해, CRISPR-Cas9의 적용이 흔히 제한된다. 가장 흔히 사용되는 CRISPR 뉴클레아제인, 스트렙토코커스 피오케네스 Cas9(*Streptococcus pyogenes* Cas9: SpCas9)의 PAM 서열은 5'-NGG-3'이다. 포유동물 세포에서 효율적인 활성을 나타내는 다른 CRISPR 뉴클레아제는 NGG 이외에 다른 PAM도 인식한다. PAM 서열이 없는 서열을 표적으로 하는 CRISPR 뉴클레아제 변이체가 개발되었지만, 표적이 되지 않는 서열들이 여전히 있다.
- [0003] 최근, SpCas9, xCas9 및 SpCas9-NG의 신규한 PAM 서열들이 확인되었다(한국 공개번호 10-2020-0026164 A1(2020.03.10)). 그러나, 다른 Cas9 뉴클레아제 변이체의 표적 서열에 대한 광범위한 조사는 이루어진 바가 없다.
- [0004] 따라서, CRISPR-Cas9 시스템의 적용 범위를 넓히기 위해, Cas9 뉴클레아제 변이체의 표적 서열에 대한 활성을 조사하여, NGG가 아닌 PAM 서열을 발굴할 필요가 있다.

발명의 내용

해결하려는 과제

- [0005] 신규한 PAM 서열을 이용하여 세포의 유전체에서 표적 핵산을 변형시키는 방법을 제공한다.
- [0006] 신규한 PAM 서열을 이용하여 세포의 유전체에서 표적 핵산을 변형시키는 방법에 의해 유전체의 표적 핵산이 변형된 세포를 제공한다.

과제의 해결 수단

- [0007] 일 양상은 세포의 유전체에서 표적 핵산을 변형시키는 방법으로서, 표적 핵산, Cas(clustered regularly interspaced palindromic repeats-associated: CRISPR associated) 뉴클레아제 또는 이의 변이체, 및 가이드 RNA를 인큐베이션하는 단계를 포함하는 방법을 제공한다.
- [0008] 상기 세포는 생체 내 세포 또는 분리된 세포일 수 있다. 상기 세포가 생체 내 세포인 경우, 상기 방법은 생체 내(in vivo)에서 수행될 수 있다. 상기 세포가 분리된 세포인 경우, 상기 방법은 생체 외(ex vivo) 또는 시험관 내(in vitro)에서 수행될 수 있다.
- [0009] 상기 세포는 체세포, 생식세포, 줄기세포, 암세포, 또는 세포주일 수 있다. 상기 세포는 암세포, 줄기세포, 혈관내피 세포, 백혈구, 면역 세포, 상피 세포, 생식 세포, 섬유아세포, 근육 세포, 골수 세포, 표피 세포, 골아 세포, 및 신경세포로 이루어진 군으로부터 선택될 수 있다. 상기 세포는 인간을 포함한 동물, 식물, 세균, 또는 균류의 세포일 수 있다.
- [0010] 상기 유전체(genome)은 생명체 또는 세포가 갖는 모든 유전정보를 말한다.
- [0011] 상기 표적 핵산은 변형시키고자 하는 핵산을 의미한다.
- [0012] 세포의 유전체에서 표적 핵산을 변형시키는 방법은 유전자 교정 또는 유전체 편집(genome editing)으로도 불릴 수 있다.
- [0013] 상기 방법은 표적 핵산을 포함한 세포, Cas(clustered regularly interspaced palindromic repeats-associated 또는 CRISPR associated: Cas) 뉴클레아제 또는 이의 변이체를 암호화하는 폴리뉴클레오티드, 및 가이드 RNA를 인큐베이션하는 단계를 포함한다.
- [0014] 상기 표적 핵산은 프로토스페이서-인접 모티프(proto-spacer adjacent motif: PAM), 및 상기 가이드 RNA에 상보적인 표적 서열을 포함할 수 있다.
- [0015] 상기 프로토스페이서-인접 모티프(PAM)는 표적 핵산에서 Cas 뉴클레아제가 인식하는 뉴클레오티드 서열일 수 있다. 상기 PAM은 5'-NAGN-3', 5'-ACGG-3', 5'-AGAS-3', 5'-WGGN-3', 5'-SGRN-3', 5'-TGAD-3', 5'-GTGC-3', 5'-RAGN-3', 5'-CAGH-3', 5'-TAGB-3', 5'-VGAG-3', 5'-NGGN-3', 5'-AAGH-3', 5'-CAGY-3', 5'-GAGN-3', 5'-VAAG-3', 5'-WGRN-3', 5'-RGYG-3', 5'-SGAD-3', 5'-SGGN-3', 5'-TGTG-3', 5'-YGCG-3', 5'-NYGG-3', 5'-NGRN-3',

5'-SGCC-3', 5'-MTGC-3', 5'-GTGH-3', 5'-RCAG-3', 5'-VTAG-3', 5'-WGCC-3', 5'-SACG-3', 5'-NGCD-3', 5'-NARG-3', 5'-WAGA-3', 5'-GAGT-3', 5'-NGTK-3', 5'-NGCN-3', 5'-NACG-3', 5'-AGGY-3', 5'-NGDG-3', 5'-GAGG-3', 5'-GGGT-3', 5'-VYAG-3', 5'-GAAT-3', 5'-NGVN-3', 5'-VACG-3', 5'-AYGA-3', 5'-GAGH-3', 5'-TGTT-3', 5'-AGTB-3', 5'-SGTK-3', 5'-GATG-3', 5'-NGAN-3', 5'-NAAG-3', 5'-NGCG-3', 5'-NGGR-3', 5'-DGGC-3', 5'-NGGT-3', 5'-NGTG-3', 및 5'-GGAG-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 것일 수 있다. 상기 뉴클레오티드 서열에서, 'A'는 아데닌(adenine: A)을 의미하고, 'G'는 구아닌(guanine: G)을 의미하고, 'C'는 시토신(cytosine: C)을 의미하고, 및 'T'는 티민(thymine: T)을 의미한다. 'N'은 아데닌(A), 구아닌(G), 시토신(C), 티민(T), 우라실(uracil: U), 불명, 또는 기타의 핵산을 의미한다. 'R'은 퓨린(purine)을 의미하고, 아데닌(A) 또는 구아닌(G)일 수 있다. 'Y'는 피리미딘(pyrimidine)을 의미하고, 시토신(C), 티민(T), 또는 우라실(U)일 수 있다. 'M'은 아데닌(A) 또는 시토신(C)일 수 있다. 'K'는 구아닌(G), 티민(T), 또는 우라실(U)일 수 있다. 'S'는 구아닌(G) 또는 시토신(C)일 수 있다. 'W'는 아데닌(A), 티민(T) 또는 우라실(U)일 수 있다. 'B'는 아데닌(A)이 아닌 염기를 의미하고, 구아닌(G), 시토신(C), 티민(T), 또는 우라실(U)일 수 있다. 'D'는 시토신(C)이 아닌 염기를 의미하고, 아데닌(A), 구아닌(G), 티민(T), 또는 우라실(U)일 수 있다. 'H'는 구아닌(G)이 아닌 염기를 의미하고, 아데닌(A), 시토신(C), 티민(T), 또는 우라실(U)일 수 있다. 'V'는 티민(T) 및 우라실(U)이 아닌 염기를 의미하고, 아데닌(A), 구아닌(G), 또는 시토신(C)일 수 있다. 상기 PAM은 표적 서열의 3'-말단 또는 5'-말단에 바로 인접 또는 연속한 것일 수 있다.

- [0016] 상기 Cas 뉴클레아제는 핵산 이중가닥을 절단하는 엔도뉴클레아제(endonuclease)일 수 있다. 상기 Cas 뉴클레아제는 RNA-가이드 DNA 엔도뉴클레아제(RNA guided DNA endonuclease)일 수 있다. 상기 Cas 뉴클레아제는 스트렙토코커스 속(*Streptococcus* sp.), 캄필로박터 속(*Campylobacter* sp.), 레지오넬라 속(*Legionella* sp.), 네이세리아 속(*Neisseria* sp.), 파스테우렐라 속(*Pasteurella* sp.), 프란시셀라(*Francisella* sp.), 및 프레보텔라 속(*Prevotella* sp.)으로 이루어진 군으로부터 선택된 세균으로부터 유래된 뉴클레아제일 수 있다. 상기 Cas 뉴클레아제는 스트렙토코커스 피오게네스(*Streptococcus pyogenes*), 스트렙토코커스 써모필러스(*Streptococcus thermophilus*), 스트렙토코커스 아우레우스(*Streptococcus aureus*), 캄필로박터 제주니(*Campylobacter jejuni*), 레지오넬라 뉴모필라(*Legionella pneumophila*), 네이세리아 메닝기티디스(*Neisseria meningitidis*), 파스테우렐라 물토시다(*Pasteurella multocida*), 프란시셀라 노비시다(*Francisella novicida*), 및 프레보텔라 디시엔스(*Prevotella disiens*)로 이루어진 군으로부터 선택된 세균으로부터 유래된 뉴클레아제일 수 있다. 상기 Cas 뉴클레아제는 Cas9, Cpf1, C2c1, C2c2, C2c3, Cas3, Cas5, Cas7, Cas8, 또는 Cas10일 수 있다. 상기 Cas9은 예를 들어, 스트렙토코커스 피오게네스 유래 Cas9(*Streptococcus pyogenes* Cas9: SpCas9)이다.
- [0017] 상기 Cas 뉴클레아제의 변이체는 Cas 뉴클레아제의 기능을 보유한 변이체이다. 상기 Cas 뉴클레아제의 변이체는 eSpCas9(1.1), SpCas9-HF1, HypaCas9, evoCas9, Sniper-Cas9, SpCas9 VQR 변이체, SpCas9 VRER 변이체, SpCas9 VRQR 변이체, SpCas9 VRQR-HF1 변이체, 및 SpCas9 QQR1 변이체로 이루어진 군으로부터 선택될 수 있다.
- [0018] 상기 eSpCas9(1.1)은 특이도가 강화되도록 돌연변이시킨 SpCas9로서, eSpCas9 1.1 또는 eSpCas9로도 불릴 수 있다. eSpCas9(1.1)은 야생형 SpCas9에 대해 K848A, K1003A, 및 R1060A의 돌연변이를 갖는 것일 수 있다(Slaymaker, I. M. et al., Science. 2016 Jan 1; 351(6268): 84-88).
- [0019] 상기 SpCas9-HF1은 고-충실도(high-fidelity)를 갖도록 돌연변이시킨 SpCas9일 수 있다. 상기 SpCas9-HF1은 야생형 SpCas9에 대해 N497A, R661A, Q695A, 및 Q926A의 돌연변이를 갖는 것일 수 있다(Kleinstiver, B. P. et al. Nature. 2016 Jan 28; 529(7587): 490-495).
- [0020] 상기 HypaCas9는 높은 정확도(hyper-accuracy)를 갖도록 돌연변이시킨 SpCas9일 수 있다. 상기 HypaCas9은 야생형 SpCas9에 대해 N692A, M694A, Q695A, H698A의 돌연변이를 갖는 것일 수 있다(Janice S. Chen et al., Nature. 2017 Oct 19; 550(7676): 407-410).
- [0021] 상기 evoCas9은 고-충실도(high-fidelity)를 갖도록 돌연변이시킨 SpCas9일 수 있다. 상기 evoCas9은 야생형 SpCas9에 대해 M495V, Y515N, K526E, 및 R661Q(VNEQ)의 돌연변이를 갖는 것일 수 있다(Antonio Casini et al., Nature Biotechnology, 2018 Mar;36(3):265-271).
- [0022] 상기 Sniper-Cas9은 염기 수정(base editor)의 특이도(specificity)가 향상되도록 돌연변이시킨 SpCas9일 수 있다. 상기 Sniper-Cas9은 야생형 SpCas9에 대해 F539S, M763I, 및 K890N의 돌연변이를 갖는 것일 수 있다(Jungjoon K. Lee et al., Nature Communications volume 9, Article number: 3048 (2018)).
- [0023] 상기 SpCas9 VQR 변이체는 NGG PAM 서열과 상이한 PAM 서열 특이도를 갖도록 돌연변이시킨 SpCas9일 수 있다.

상기 SpCas9 VQR 변이체는 야생형 SpCas9에 대해 D1135V, R1335Q, 및 T1337R의 돌연변이를 갖는 것일 수 있다 (Benjamin P. Kleinstiver et al., Nature. 2015 Jul 23; 523(7561): 481-485).

- [0024] 상기 SpCas9 VRER 변이체는 상이한 PAM 서열 특이도를 갖도록 돌연변이시킨 SpCas9일 수 있다. 상기 SpCas9 VRER 변이체는 야생형 SpCas9에 대해 D1135V, G1218R, R1335E, 및 T1337R의 돌연변이를 갖는 것일 수 있다 (Benjamin P. Kleinstiver et al., Nature. 2015 Jul 23; 523(7561): 481-485).
- [0025] 상기 SpCas9 VRQR 변이체는 고-충실도(high-fidelity)를 갖도록 돌연변이시킨 SpCas9일 수 있다. 상기 SpCas9 VRQR 변이체는 야생형 SpCas9에 대해 D1135V, G1218R, R1335Q, 및 T1337R의 돌연변이를 갖는 것일 수 있다 (Kleinstiver, B. P. et al. Nature. 2016 Jan 28; 529(7587): 490-495).
- [0026] 상기 SpCas9 VRQR-HF1은 상이한 PAM 서열 특이도를 갖도록 돌연변이시킨 SpCas9일 수 있다. 상기 SpCas9 VRQR-HF1은 야생형 SpCas9에 대해 N497A, R661A, Q695A, Q926A, D1135V, G1218R, R1335Q, 및 T1337R의 돌연변이를 갖는 것일 수 있다 (Kleinstiver, B. P. et al. Nature. 2016 Jan 28; 529(7587): 490-495).
- [0027] 상기 SpCas9 QQR1은 를 갖도록 돌연변이시킨 SpCas9일 수 있다. 상기 SpCas9 QQR1은 야생형 SpCas9에 대해 G1218R, N1286Q, I1331F, D1332K, R1333Q, R1335Q, 및 T1337R의 돌연변이를 갖는 것일 수 있다 (Carolin Anders et al., Mol Cell. 2016 Mar 17; 61(6): 895-902).
- [0028] 상기 Cas 뉴클레아제 또는 이의 변이체를 암호화하는 폴리뉴클레오티드는 플라스미드 벡터 또는 바이러스 벡터에 함유된 것일 수 있다.
- [0029] 상기 Cas 뉴클레아제 변이체가 eSpCas9(1.1)인 경우, 상기 PAM은 5'-NAGN-3', 5'-ACGG-3', 5'-AGAS-3', 5'-WGGN-3', 5'-SGRN-3', 5'-TGAD-3', 및 5'-GTGC-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드일 수 있다.
- [0030] 상기 Cas 뉴클레아제 변이체가 SpCas9-HF1인 경우, 상기 PAM은 5'-RAGN-3', 5'-CAGH-3', 5'-TAGB-3', 5'-VGAG-3', 및 5'-NGGN-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드일 수 있다.
- [0031] 상기 Cas 뉴클레아제 변이체가 HypaCas9인 경우, 상기 PAM은 5'-AAGH-3', 5'-CAGY-3', 5'-GAGN-3', 및 5'-NGGN-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드일 수 있다.
- [0032] 상기 Cas 뉴클레아제 변이체가 evoCas9인 경우, 상기 PAM은 5'-VAAG-3', 5'-WGRN-3', 5'-RGY-3', 5'-SGAD-3', 5'-SGGN-3', 5'-TGTG-3', 및 5'-YCGC-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드일 수 있다.
- [0033] 상기 Cas 뉴클레아제 변이체가 Sniper-Cas9인 경우, 상기 PAM은 5'-NAGN-3', 5'-NYGG-3', 5'-NGRN-3', 5'-SGCC-3', 5'-MTGC-3', 및 5'-GTGH-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드일 수 있다.
- [0034] 상기 Cas 뉴클레아제 변이체가 SpCas9 VQR 변이체인 경우, 상기 PAM은 5'-RCAG-3', 5'-VTAG-3', 5'-WGCC-3', 5'-SACG-3', 5'-NGCD-3', 5'-NGRN-3', 5'-NARG-3', 5'-WAGA-3', 5'-GAGT-3', 및 5'-NGTK-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드일 수 있다.
- [0035] 상기 Cas 뉴클레아제 변이체가 SpCas9 VRER 변이체인 경우, 상기 PAM은 5'-NGCN-3', 5'-NACG-3', 5'-AGGY-3', 5'-NGDG-3', 5'-GAGG-3', 및 5'-GGGT-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드일 수 있다.
- [0036] 상기 Cas 뉴클레아제 변이체가 SpCas9 VRQR 변이체인 경우, 상기 PAM은 5'-VYAG-3', 5'-GAAT-3', 5'-NGVN-3', 5'-VACG-3', 5'-WAGA-3', 5'-AYGA-3', 5'-GAGH-3', 5'-NARG-3', 5'-TGTN-3', 5'-AGTB-3', 5'-SGTK-3', 및 5'-GATG-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드일 수 있다.
- [0037] 상기 Cas 뉴클레아제 변이체가 SpCas9 VRQR-HF1 변이체인 경우, 상기 PAM은 5'-NGAN-3', 5'-NAAG-3', 5'-RCAG-3', 5'-VTAG-3', 5'-NGCG-3', 5'-NGGR-3', 5'-DGGC-3', 5'-GAGG-3', 5'-NGGT-3', 및 5'-NGTG-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드일 수 있다.
- [0038] 상기 Cas 뉴클레아제 변이체가 SpCas9 QQR1 변이체인 경우, 상기 PAM은 5'-NAAG-3' 및 5'-GGAG-3'으로 이루어진 군으로부터 선택된 뉴클레오티드 서열로 이루어진 폴리뉴클레오티드일 수 있다.

- [0039] 용어 "가이드(guide) RNA"는 유전체 편집을 통해 세포내에서 표적 핵산을 인식하여 표적 핵산을 절단, 삽입, 또는 연결시키는 폴리뉴클레오타이드를 말한다. 상기 가이드 RNA는 표적 핵산 중 표적 서열에 상보적인 서열을 포함할 수 있다. 상기 가이드 RNA는 상기 표적 핵산에서 PAM의 5' 방향 또는 3' 방향으로 연속적인 2 내지 24 뉴클레오타이드(이하, 'nt'라 함)의 뉴클레오타이드 서열과 상보적인 폴리뉴클레오타이드일 수 있다. 상기 가이드 RNA의 길이는 17 nt, 18 nt, 19 nt, 20 nt, 21 nt, 22 nt, 23 nt, 또는 24 nt일 수 있다. 상기 가이드 RNA는 단일-사슬 가이드 RNA(single-chain guide RNA: sgRNA)일 수 있다. 상기 sgRNA는 표적 핵산 서열에 특이적인 crRNA(CRISPR RNA) 및/또는 Cas 뉴클레아제와 복합체를 형성하는 tracrRNA를 포함할 수 있다. 상기 가이드 RNA는 플라스미드 벡터 또는 바이러스 벡터에 함유된 것일 수 있다.
- [0040] 상기 인큐베이션은 표적 핵산을 포함하는 세포에 Cas 뉴클레아제 또는 이의 변이체를 암호화하는 폴리뉴클레오타이드 및 가이드 RNA를 도입하는 것일 수 있다. 상기 도입은 통합(integration), 형질전환(transformation), 형질도입(transduction), 형질감염(transfection), 또는 이들의 조합을 포함한다. 상기 도입은 일시적(transient) 또는 안정한(stable) 도입일 수 있다.
- [0041] 상기 표적 핵산은 상기 Cas 뉴클레아제 또는 이의 변이체, 및 상기 가이드 RNA의 복합체에 의해 인식될 수 있다.
- [0042] 상기 Cas 뉴클레아제 또는 이의 변이체, 및 가이드 RNA의 복합체는 표적 핵산을 서열-특이적으로 변형될 수 있다. 상기 변형은 삽입, 절단, 삽입, 연결, 탈아미노화, 또는 이들의 조합일 수 있다. 상기 절단은 유전체 DNA의 이중가닥의 절단일 수 있다. 상기 절단은 평활 말단(blunt end) 또는 접착 말단(sticky end)일 수 있다. 상기 변형은 표적 핵산의 절단 및 절단 부위에 외래 폴리뉴클레오타이드의 삽입일 수 있다. 외래 폴리뉴클레오타이드를 유전체의 절단 부위에 삽입하는 것은 상동성 의존적 방법에 의한 것일 수 있다. 상기 상동성 의존적 방법은 상동 재조합(homologous recombination) 또는 상동성-직접 수선(homology-directed repair: HDR)일 수 있다.
- [0043] 상기 방법은 시험관 내(in vitro), 생체 외(ex vivo), 또는 생체 내(in vivo)에서 수행될 수 있다.
- [0044] 다른 양상은 일 양상의 세포의 유전체에서 표적 핵산을 변형시키는 방법으로 유전체의 표적 핵산이 변형된 세포를 제공한다.
- [0045] 상기 세포, 유전체, 표적 핵산, 및 변형은 전술한 바와 같다.

발명의 효과

- [0046] 신규한 PAM 서열을 이용하는 세포의 유전체에서 표적 핵산을 변형시키는 방법 및 이에 의해 유전체의 표적 핵산이 변형된 세포에 의하면, 기존에 표적으로 하지 못한 위치를 유전체 편집의 표적으로 하여 유전체 편집을 수행할 수 있고, 이에 의해 유전체 편집의 활용 범위를 확대시킬 수 있다.

도면의 간단한 설명

- [0047] 도 1은 고정된 프로토스페이서를 이용하여 PAM 서열을 결정하기 위한 실험 과정을 나타낸 개략도이다.
- 도 2a 내지 도 2c는 실시예에 사용된 SpCas9 뉴클레아제 및 그의 변이체의 모식도이다(빨간색 화살표: 야생형 SpCas9과 비교하여 돌연변이가 있는 위치, Arg: arginine-rich bridge helix, PI: PAM-interacting domain).
- 도 3 내지 도 15는 실시예에 사용된 SpCas9 뉴클레아제 및 그의 변이체에 의해 4-nt 후보 PAM 서열에서 유도된 평균 인텔 빈도를 나타내는 히트맵(heat map)이다.

발명을 실시하기 위한 구체적인 내용

- [0048] 이하 실시예를 통하여 보다 상세하게 설명한다. 그러나, 이들 실시예는 하나 이상의 구체예를 예시적으로 설명하기 위한 것으로 본 발명의 범위가 이들 실시예에 한정되는 것은 아니다.
- [0049] **실시예 1. SpCas9 변이체의 PAM 서열의 확인**
- [0050] **1. 가이드 RNA, PAM 서열, 및 표적 서열을 함유한 플라스미드 라이브러리의 준비**
- [0051] SpCas9 PAM 서열인 NGG 외에 다른 PAM 서열 확인을 위해, Twist Bioscience Co.에 의뢰하여 올리고뉴클레오타이드 라이브러리를 제작하였다(도 1).
- [0052] 각 올리고뉴클레오타이드는, 5' 말단으로부터 19-nt의 sgRNA, BsmBI 제한효소 부위, 바코드 1(20-nt 서열), 제2

BsmBI 제한효소 부위, 바코드 2(15-nt 서열), 및 무작위적 서열을 갖는 8-nt PAM 서열을 함유하는 30-nt 표적 서열을 포함하도록 디자인되었다.

[0053] 우선, 30종의 프로토스페이서와 쌍을 이루는 271종의 상이한 6-nt PAM 서열(256개의 NNNNTA + 16개의 AGGTNN - 1개의 겹치는 AGGGTA)을 함유하는 8,130개 (271 x 30)의 표적 서열을 가진 라이브러리를 제작하였다. 38% 내지 42% 범위의 SpCas9-유도 인텔 빈도와 관련하여 완벽히 매칭된 GN₁₉ 프로토스페이서를 선택하였다. 인텔 빈도는 렌티바이러스로 SpCas9를 도입한 후 3일에 NGG PAM 서열을 갖고 렌티바이러스 도입된 7,314개의 표적 서열로 결정하였고, 범위는 0% 내지 99 %의 범위이다. 7,314개의 sgRNA 중 38% 내지 42% 범위의 인텔 빈도를 갖는 30종을 무작위로 선별하였다.

[0054] 다음으로, 미스매치(mismatch)된 표적 서열에서 뉴클레아제 활성을 검증하기 위해 2,940 2,940 sgRNA를 디자인 하였다: NGG PAM을 갖는 2,940=30 sgRNAs x 98개의 표적 [미스매치 없는 1개의 표적 (5개의 상이한 바코드) + 각각 1 염기 미스매치를 갖는 60개의 표적 + 각각 2 염기의 미스매치를 갖는 19개의 표적 + 각각 3 염기의 미스매치를 갖는 18개의 표적], 및 732=4 sgRNAs×61 표적 [미스매치 없는 1개의 표적 (5개의 상이한 바코드) + 각각 1 염기 미스매치를 갖는 60개의 표적] x 3개의 PAM(NGAG, NAAG 및 NGCG).

[0055] sgRNA와 표적 서열을 함유하는 플라스미드 라이브러리는 2단계 클로닝 방법을 이용하여, 올리고뉴클레오타이드 풀(pool)의 PCR 증폭 과정에서 가이드 RNA와 표적 서열간의 언커플링(uncoupling)을 방지하였다.

[0056] 제1 단계는 가이드 RNA 및 표적 서열을 포함하는 초기 플라스미드 라이브러리 제작이다. BsmBI 제한효소(NEB)를 사용하여 Lenti-gRNA-Puro 플라스미드(Addgene, #84752)를 선형화시키고, 선형화된 벡터에 PCR-증폭된 올리고뉴클레오타이드 풀(표적 서열)을 라이게이션시켰다. 반응물을 대장균에 형질전환시키고, 선별된 콜로니로부터 플라스미드를 분리하였다. 올리고뉴클레오타이드 풀 증폭에 사용된 프라이머는 다음과 같다.

[0057] 정방향 프라이머:

[0058] 5'-TTGAAAGTATTTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACC-3' (서열번호 1)

[0059] 역방향 프라이머:

[0060] 5'-GAGTAAGCTGACCGCTGAAGTACAAGTGGTAGAGTAGAGATCTAGTTACGCCAAGCT-3' (서열번호 2)

[0061] 제2 단계는 sgRNA 스캐폴드의 삽입이다. 제1 단계에서 제작된 플라스미드 라이브러리는 BsmBI 제한효소(NEB)로 절단하고 아가로스 전기영동 후 겔에서 핵산 단편을 정제하였다. sgRNA 스캐폴드를 함유한 인서트 단편을 합성하고 TOPO 벡터(T-두가닥(blunt) 벡터, Solgent)에 클로닝하였다. 인서트 단편의 서열을 하기에 나타내었고, 폴리 T 서열을 포함한 sgRNA는 굵은 글씨로 표시하고, BsmBI 제한효소 부위는 밑줄로 표시하였다.

[0062] 5'-CGTCTCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAG

[0063] TGGCACCGAGTCGGTGTCTTTTTGGGAGACG-3' (서열번호 3)

[0064] 인서트 단편을 함유한 TOPO 벡터를 BsmBI 제한효소(Enzynomics)로 절단하여 83-nt 인서트 단편을 분리하였다. 절단된 플라스미드 라이브러리에 인서트 단편을 라이게이션시키고, 대장균에 형질전환시킨 후, 선별된 콜로니로부터 플라스미드 라이브러리를 분리하였다.

[0065] 2. 세포 라이브러리의 준비

[0066] 우선, 렌티바이러스 라이브러리 생산을 위해, 인간 배아 신장 세포인 HEK293T 세포(ATCC)를 준비하였다. 실시예 1.1에서 준비된 플라스미드, psPAX2, 및 pMD2.G를 혼합하고, 리포펙타민 2000 (Invitrogen)을 사용하여 HEK293T 세포에 형질감염시켰다. 형질감염 후 12 시간에 신선한 배지를 세포에 가하고, 형질 감염 후 36 시간에 바이러스를 함유한 상층액을 수득하였다. 수득된 상층액은 Millex-HV 0.45 μ m 저-단백질 결합 멤브레인(Millipore)으로 여과하고, 분액은 사용시까지 -80℃에 보관하였다. 바이러스 수율은 Lenti-X p24 Rapid Titer Kit(Clontech)로 측정하여 검증하였다. 바이러스 역가 산출을 위해, 순차 희석된 바이러스 분액을 8 μ g/ml의 폴리브렌(polybrene)의 존재에서 HEK293T 세포에 형질도입하고, 2 μ g/ml 퓨로마이신 또는 20 μ g/ml 블라스티시딘(blasticidin) S(InvivoGen)의 존재에서 배양하여 산출하였다.

[0067] 준비된 렌티바이러스 라이브러리의 형질도입을 위해, HEK293T 세포를 배양 접시에 밤새 배양하였다. 감염다중도(multiplicity of infection: MOI) 0.3의 렌티바이러스 라이브러리를 8 μ g/ml의 폴리브렌의 존재에서 HEK293T 세포에 형질도입하고, 세포를 밤새 배양하였다. 세포를 2 μ g/ml 퓨로마이신의 존재에서 배양하여 형질도입되지

많은 세포를 제거하고, 1.2×10^7 세포의 양으로 세포 라이브러리를 유지하였다.

[0068] 3. 세포 라이브러리에 Cas9의 전달

[0069] 1.2×10^7 세포의 양의 세포 라이브러리를 준비하고, SpCas9, eSpCas9(1.1), SpCas9-HF1, HypaCas9, evoCas9, Sniper-Cas9, xCas9, SpCas9 VQR 변이체, SpCas9 VRER 변이체, SpCas9 VRQR 변이체, SpCas9 VRQR-HF1 변이체, SpCas9 QQR1 변이체 및 SpCas9-NG를 암호화하는 바이러스를 $8 \mu\text{g}/\text{ml}$ 의 폴리브렌의 존재에서 세포에 형질도입시켰다(도 2a 내지 도 2c). 형질도입은 MOI 5로 수행하였고, $20 \mu\text{g}/\text{ml}$ 블라스티시딘 S의 존재 하에서 세포를 선별하였다.

[0070] 4. 인텔 빈도의 측정

[0071] 실시예 1.3에서 준비된 세포에서 유전체 중 인텔(삽입/결실, insertion/deletion: indel)의 빈도를 측정하기 위해, 고심도(deep) 시퀀싱 및 인텔 빈도의 분석을 수행하였다.

[0072] 고심도 시퀀싱(deep sequencing)을 위해, Wizard Genomic DNA purification kit(Promega)를 사용하여 세포로부터 유전체 DNA를 분리하였다. 삽입된 표적 사열은 고성능 실험을 위해 2X Taq PCR Smart mix(Solgent)를 사용하여 PCR 증폭하였다. 첫번째 PCR에서, 각 세포 라이브러리에 대해 총 $240 \mu\text{g}$ 의 유전체 DNA를 이용하여 라이브러리에 대해 $1000\times$ 이상의 커버리지가 되도록 하였다(10^6 개의 세포 당 약 $10 \mu\text{g}$ 의 유전체 DNA). 각 반응 당 $2.5 \mu\text{g}$ 의 유전체 DNA를 1차 PCR하고, 반응산물을 모두 풀링(pooling)한 후 정제하였다. 정제된 산물 중 50 ng 의 시료를 Illumina 어댑터 및 바코드 서열을 함유한 프라이머를 이용하여 2차 PCR 증폭하였다. 증폭된 산물은 전기영동 후 정제하고, HiSeq 또는 MiniSeq(Illumina)를 이용하여 분석하였다.

[0073] 실험에 사용된 프라이머는 다음과 같다.

[0074] 1차 PCR 반응용 프라이머

[0075] 정방향 프라이머:

[0076] 5'-ACACTCTTCCCTACACGACGCTCTCCGATCTCTTGAAAAAGTGGCACCGAGTCG-3' (정방향 프라이머, 서열번호 4)

[0077] 5'-ACACTCTTCCCTACACGACGCTCTCCGATCTCTTGAAAAAGTGGCACCGAGTCG-3' (서열번호 5)

[0078] 5'-ACACTCTTCCCTACACGACGCTCTCCGATCTCGCTTGAAAAAGTGGCACCGAGTCG-3' (서열번호 6)

[0079] 역방향 프라이머:

[0080] 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTTAAGTCGAGTAAGCTGACCGCTGAAG-3' (서열번호 7)

[0081] 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTATTAAGTCGAGTAAGCTGACCGCTGAAG-3' (서열번호 8)

[0082] 5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTTATTAAGTCGAGTAAGCTGACCGCTGAAG-3' (서열번호 9)

[0083] 2차 PCR 반응용 프라이머

[0084] 정방향 프라이머:

[0085] 5'-AATGATACGGCGACCGAGATCTACAC (index) ACACTCTTCCCTACACGAC-3' (서열번호 10)

[0086] 역방향 프라이머:

[0087] 5'-CAAGCAGAAGACGGCATACGAGAT (index) GTGACTGGAGTTCAGACGTGT-3' (서열번호11)

[0088] 인텔 빈도의 분석을 위해 Python scripts 프로그램을 변형하여 고심도 시퀀싱 데이터를 분석하였다. 15-nt의 바코드와 그의 상류에 위치한 4-nt의 총 19-nt의 서열을 이용하여 각 가이드 RNA와 표적 서열의 쌍을 식별하였다. 인텔이 예상된 절단 부위(즉, 그 부위의 가운데에 위치한 8-nt 영역)에 위치해 있을 경우, 인텔은 Cas9-유도된 돌연변이라고 간주하였다. 어레이 합성 및 PCR 증폭의 결과인 배경 인텔 빈도를 제거하기 위해, 총 리드(read), 인텔 리드, 및 Cas9을 도입하지 않은 경우의 인텔 빈도를 산출하고, 하기 수학적식에 따라 인텔 빈도(%)를 산출하였다.

$$\text{인텔 빈도}(\%) = \frac{\text{인텔 리드} - (\text{총 리드의 수} \times \text{배경 인텔 빈도})}{\text{총 리드} - (\text{총 리드의 수} \times \text{배경 인텔 빈도})} \times 100$$

[0089]

[0090] (수확식 1)

[0091] 고심도 시퀀싱 데이터는 NCBI Sequence Read Archive(SRA; www.ncbi.nlm.nih.gov/sra/)에 Accession No. SRR10215483에 업로드하였다.

[0092] 5. PAM 서열의 결정

[0093] 실시예 1.4에 따라 수득된 고심도 시퀀싱 데이터에서, xCas9, SpCas9-NG, 또는 SpCas9 뉴클레아제에서 높은 인텔 빈도를 갖는 PAM 서열을 선별하였다.

[0094] 사람 세포에서 형질도입 후 6일까지 5% 이상의 평균 인텔 빈도를 갖는 PAM 서열을 선별하고, Cas9 뉴클레아제 및 그의 변이체와 PAM 서열에 따른 인텔 빈도를 분석하였다.

[0095] SpCas9(야생형), eSpCas9(1.1), SpCas9-HF1, HypaCas9, evoCas9, Sniper-Cas9, xCas9, SpCas9 VQR 변이체, SpCas9 VRER 변이체, SpCas9 VRQR 변이체, SpCas9 VRQR-HF1 변이체, SpCas9 QQR1 변이체 및 SpCas9-NG에 의해 4-nt 후보 PAM 서열에서 유도된 평균 인텔 빈도를 나타내는 히트맵(heat map)을 각각 도 3 내지 도 15에 나타내었다(단방향 ANOVA 분석 후 Tukey 사후 검정). 4-nt 후보 PAM 서열(총 256($=4^4$)개의 후보 서열) 중 5% 이상의 평균 인텔 빈도를 갖는 PAM 서열을 음영으로 표시하였고, 5% 미만의 평균 인텔 빈도를 갖는 서열은 흰색으로 나타내었다. 높은 평균 인텔 빈도를 갖는 PAM 서열은 빨간색 상자로 표시하였다. 5'-NNNN-3'의 모든 가능한 4-nt PAM 서열에 대해 실험을 진행하였으며, 각각의 PAM 서열당 22 내지 30개의 표적 서열을 이용하여 분석을 진행하였다.

[0096] 도 3 내지 도 15를 분석하여 확인된 4-nt PAM 서열을 하기 표 1에 나타내었다.

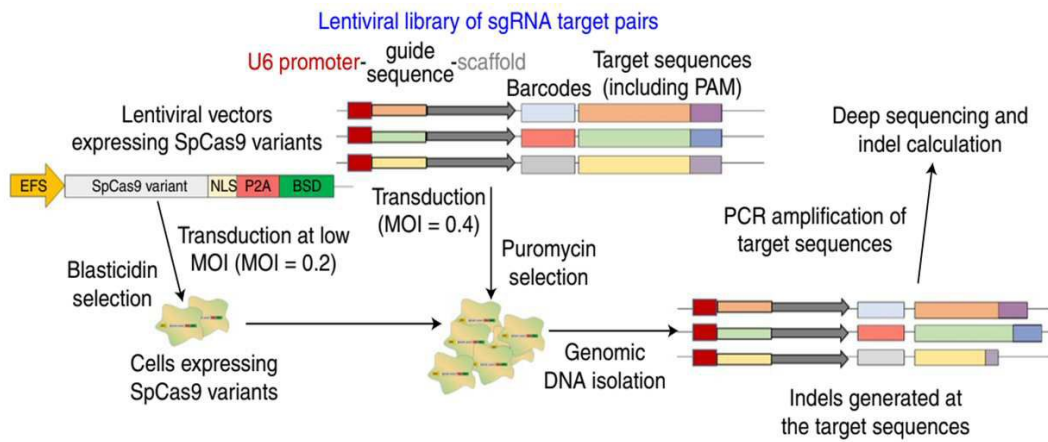
표 1

Cas9 뉴클레아제	PAM 서열
SpCas9	5'-NGRN-3', 5'-NAGN-3', 5'-GGCN-3', 5'-YGCC-3', 5'-GTGH-3', 5'-NYGG-3', 5'-MTGC-3', 5'-GCGC-3'
eSpCas9(1.1)	5'-NAGN-3', 5'-ACGG-3', 5'-AGAS-3', 5'-WGGN-3', 5'-SGRN-3', 5'-TGAD-3', 5'-GTGC-3'
SpCas9-HF1	5'-RAGN-3', 5'-CAGH-3', 5'-TAGB-3', 5'-VGAG-3', 5'-NGGN-3'
HypaCas9	5'-AAGH-3', 5'-CAGY-3', 5'-GAGN-3', 5'-NGGN-3'
evoCas9	5'-VAAG-3', 5'-WGRN-3', 5'-RGY-3', 5'-SGAD-3', 5'-SGGN-3', 5'-TGTG-3', 5'-YGCG-3'
Sniper-Cas9	5'-NAGN-3', 5'-NYGG-3', 5'-NGRN-3', 5'-SGCC-3', 5'-MTGC-3', 5'-GTGH-3'
xCas9	5'-NGDN-3', 5'-VAGN-3', 5'-SAWC-3', 5'-NGCH-3', 5'-VGCG-3', 5'-GACC-3', 5'-TAGS-3', 5'-AYGG-3', 5'-YCGG-3', 5'-GTGC-3'
SpCas9 VQR	5'-RCAG-3', 5'-VTAG-3', 5'-WGCC-3', 5'-SACG-3', 5'-NGCD-3', 5'-NGRN-3', 5'-NARG-3', 5'-WAGA-3', 5'-GAGT-3', 5'-NGTK-3'
SpCas9 VRER	5'-NGCN-3', 5'-NACG-3', 5'-AGGY-3', 5'-NGDG-3', 5'-GAGG-3', 5'-GGGT-3'
SpCas9 VRQR	5'-VYAG-3', 5'-GAAT-3', 5'-NGVN-3', 5'-VACG-3', 5'-WAGA-3', 5'-AYGA-3', 5'-GAGH-3', 5'-NARG-3', 5'-TGTN-3', 5'-AGTB-3', 5'-SGTK-3', 5'-GATG-3'
SpCas9 VRQR-HF1	5'-NGAN-3', 5'-NAAG-3', 5'-RCAG-3', 5'-VTAG-3', 5'-NGCG-3', 5'-NGGR-3', 5'-DGGC-3', 5'-GAGG-3', 5'-NGGT-3', 5'-NGTG-3'
SpCas9 QQR1	5'-NAAG-3', 5'-GGAG-3'
SpCas9-NG	5'-NGNN-3', 5'-NAGN-3', 5'-VAHD-3', 5'-SAMC-3', 5'-VATC-3', 5'-TAHG-3', 5'-RCDG-3', 5'-TACT-3', 5'-TATA-3', 5'-VTNG-3', 5'-GTAW-3', 5'-RCGT-3', 5'-VTGW-3', 5'-RTGC-3', 5'-TTGG-3'

[0098] 따라서, 신규한 PAM 서열을 이용하여 인간 유전체에서 표적가능한 위치를 발굴함하고, PAM 서열에 따른 뉴클레아제의 인텔 유도 활성을 예측함으로써, 유전체 편집의 활용 가능성을 증가시킬 수 있었다.

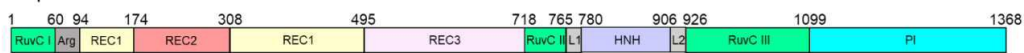
도면

도면1

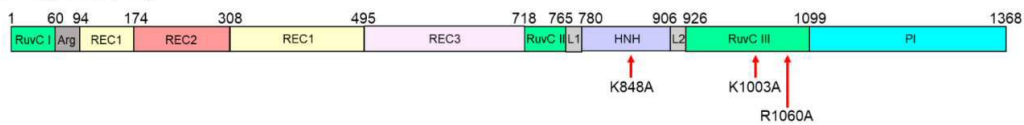


도면2a

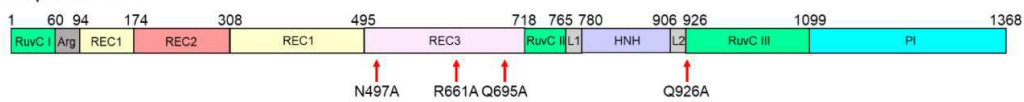
a. SpCas9



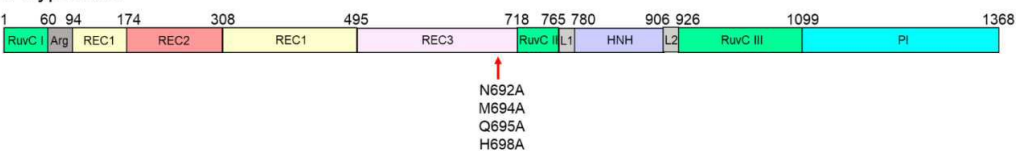
b. eSpCas9(1.1)



c. SpCas9-HF1

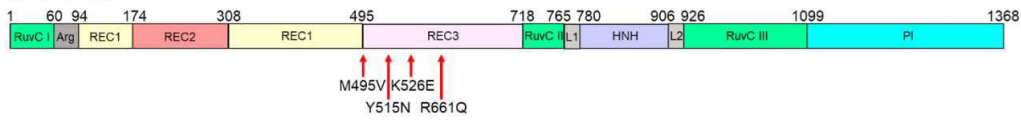


d. HypaCas9

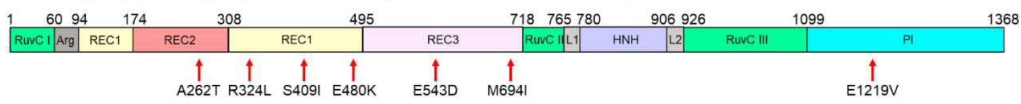


도면2b

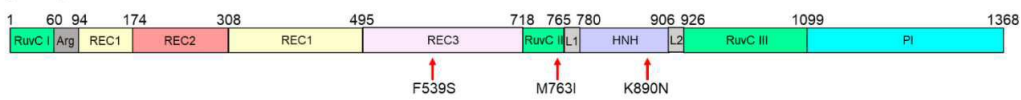
e. evoCas9



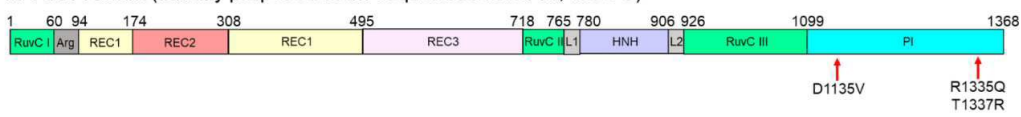
f. xCas9(3.7) (Initially proposed PAM sequences : NG, GAA, GAT)



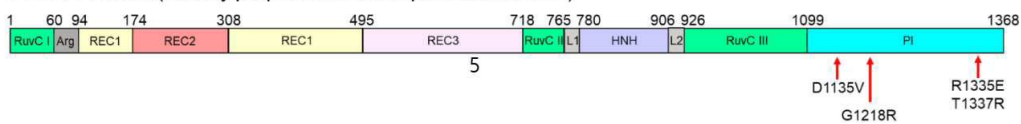
g. Sniper-Cas9



h. VQR variant (Initially proposed PAM sequences : NGAN, NGCG)

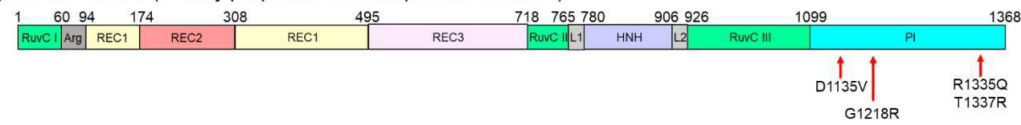


i. VRER variant (Initially proposed PAM sequences : NGCG)

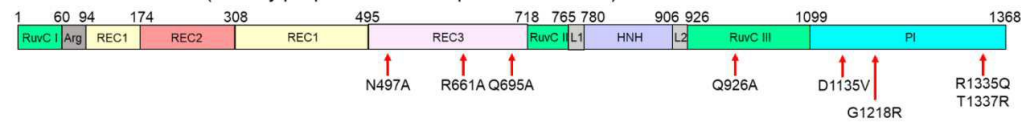


도면2c

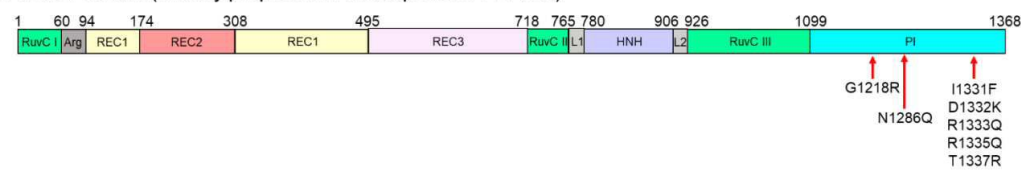
j. VRQR variant (Initially proposed PAM sequences : NGAH)



k. VRQR-HF1 variant (Initially proposed PAM sequences : NGAH)



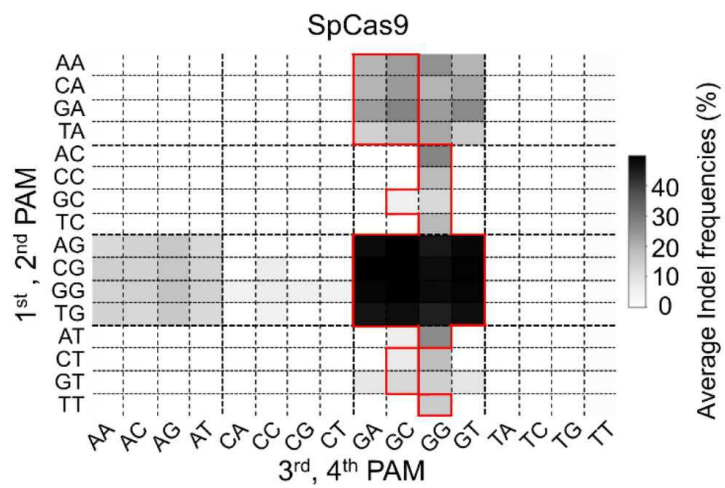
l. QQR1 variant (Initially proposed PAM sequences : NAAG)



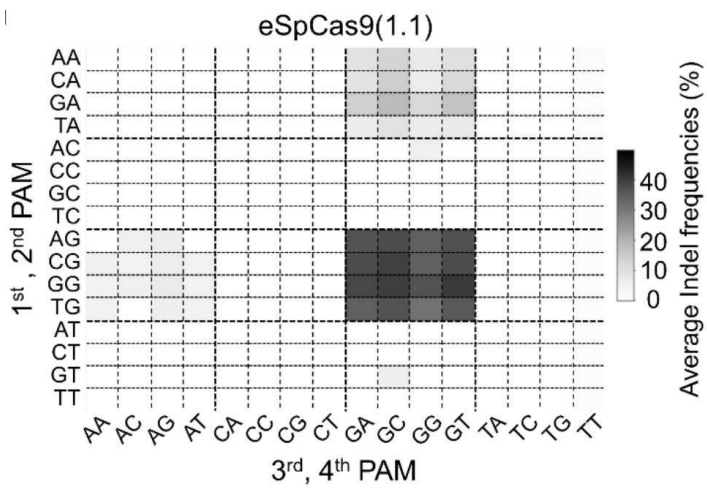
m. SpCas9-NG (Initially proposed PAM sequences : NG)



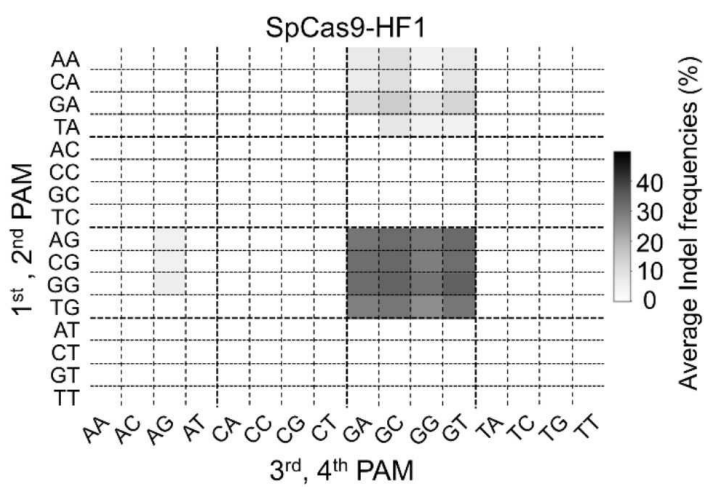
도면3



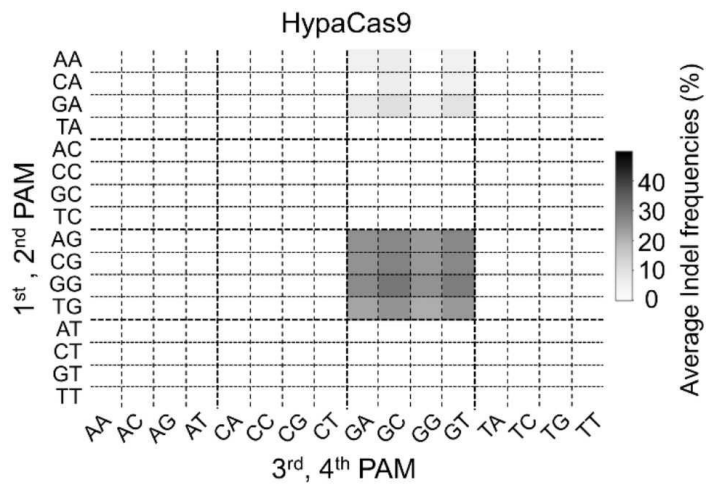
도면4



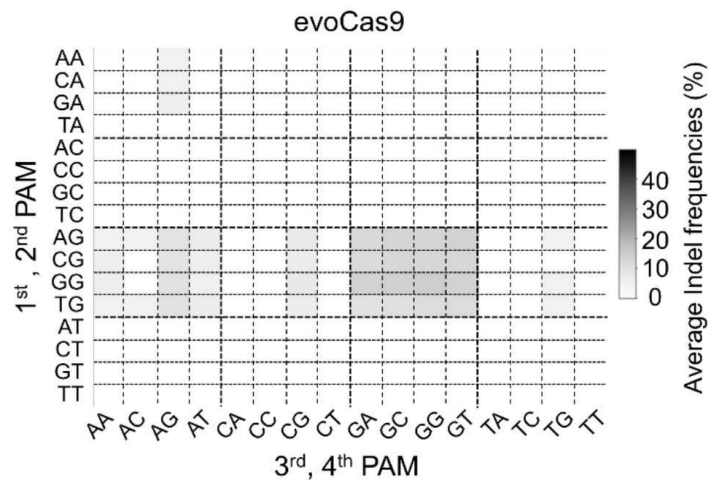
도면5



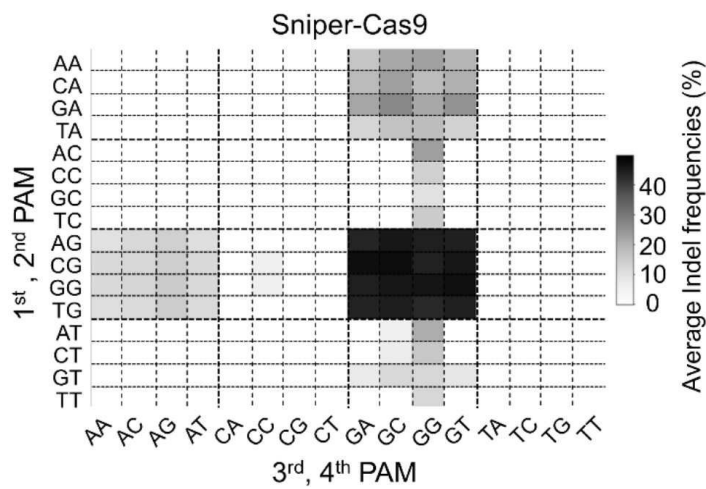
도면6



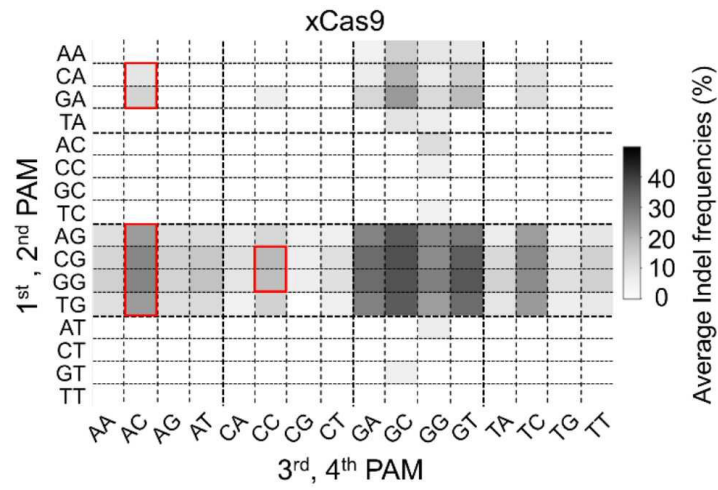
도면7



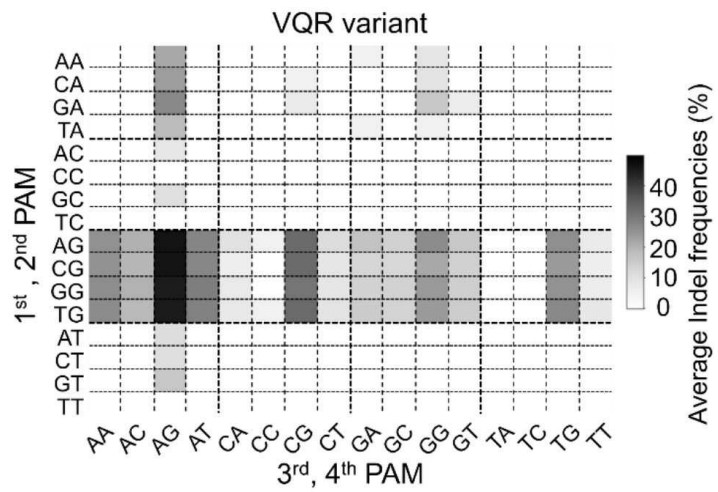
도면8



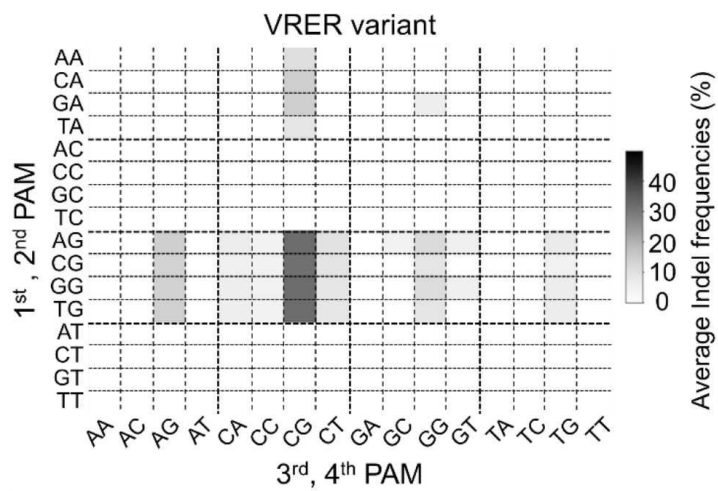
도면9



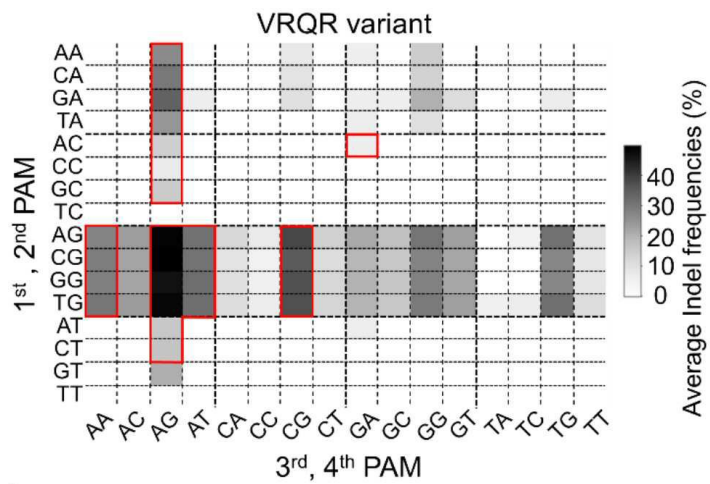
도면10



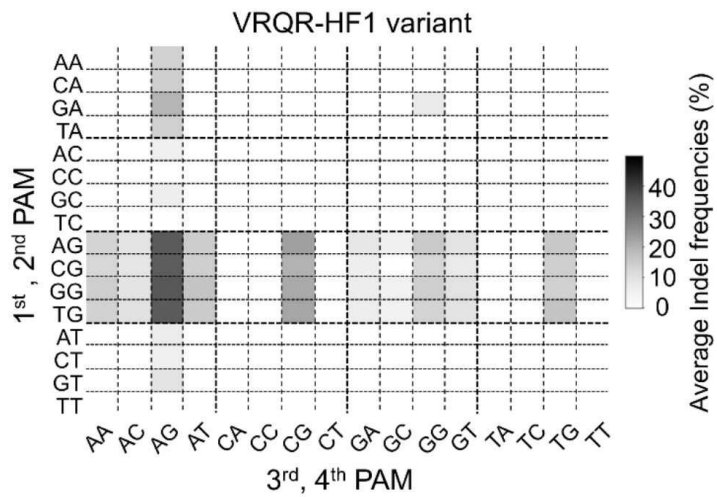
도면11



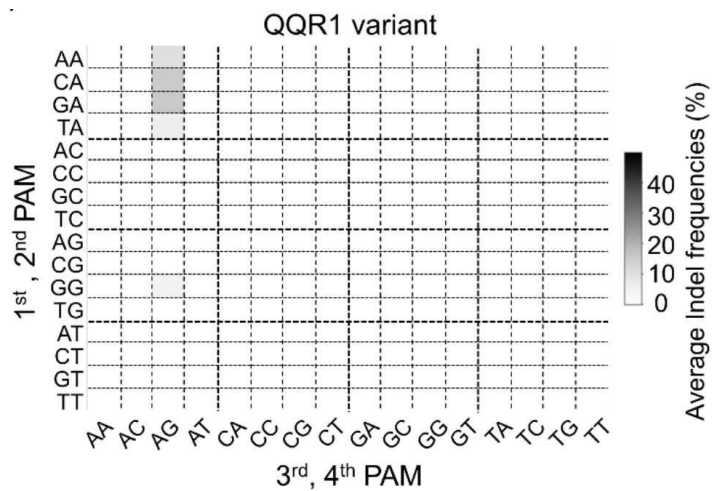
도면12



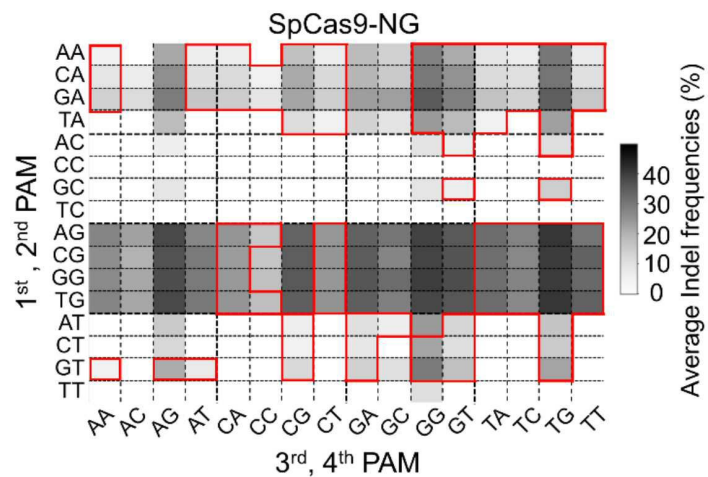
도면13



도면14



도면15



서열 목록

<110> Industry-Academic Cooperation Foundation, Yonsei University

<120> Novel protospacer adjacent motif sequence and method for
modifying a target nucleic acid in the genome of a cell using the
same

<130> PN135253KR

<150> KR 10-2019-0127304

<151> 2019-10-14

<160> 11

<170> KoPatent In 3.0

<210> 1

<211> 55

<212> DNA

<213> Artificial Sequence

<220><223> Forward primer for oligonucleotide pool amplification

<400> 1

ttgaaagtat ttcgatttct tggctttata tatcttgtgg aaaggacgaa acacc

55

<210> 2

<211> 57

<212> DNA

<213> Artificial Sequence

<220><223> Reverse primer for oligonucleotide pool amplification

<400> 2

gagtaagctg accgctgaag tacaagtggt agagtagaga tctagttacg ccaagct

57

<210> 3
 <211> 97
 <212> DNA
 <213> Artificial Sequence
 <220><223> Insert fragment containing the sgRNA scaffold
 <400> 3
 cgtctctgtt ttagagctag aaatagcaag ttaaaataag gctagtcgt tatcaacttg 60
 aaaaagtggc accgagtcgg tgcttttttg ggagacg 97

<210> 4
 <211> 56
 <212> DNA
 <213> Artificial Sequence
 <220><223> Forward primer for 1st PCR reaction
 <400> 4
 acactctttc cctacacgac gctcttccga tctcttgaaa aagtggcacc gagtcg 56

<210> 5
 <211> 57
 <212> DNA
 <213> Artificial Sequence
 <220><223> Forward primer for 1st PCR reaction
 <400> 5
 acactctttc cctacacgac gctcttccga tctcttgaa aaagtggcac cgagtcg 57

<210> 6
 <211> 58
 <212> DNA
 <213> Artificial Sequence
 <220><

223> Forward primer for 1st PCR reaction
 <400> 6
 acactctttc cctacacgac gctcttccga tctcgcttga aaaagtggca ccgagtcg 58

<210> 7
 <211> 61
 <212> DNA
 <213> Artificial Sequence

<220><223> Reverse primer for 1st PCR reaction
 <400> 7
 gtgactggag ttcagacgtg tgctcttccg atctttaagt cgagtaagct gaccgctgaa 60
 g 61
 <210> 8
 <211> 62
 <212> DNA
 <213> Artificial Sequence
 <220><223>
 223> Reverse primer for 1st PCR reaction
 <400> 8
 gtgactggag ttcagacgtg tgctcttccg atctattaag tcgagtaagc tgaccgctga 60
 ag 62
 <210> 9
 <211> 63
 <212> DNA
 <213> Artificial Sequence
 <220><223> Reverse primer for 1st PCR reaction
 <400> 9
 gtgactggag ttcagacgtg tgctcttccg atcttattaa gtcgagtaag ctgaccgctg 60
 aag 63
 <210> 10
 <211> 49
 <212> DNA
 <213> Artificial Sequence
 <220><223> Forward primer for 2nd PCR reaction
 <400> 10
 aatgatacgg cgaccaccga gatctacaca cactctttcc ctacacgac 49
 <210> 11
 <211> 45
 <212> DNA
 <213> Artificial Sequence
 <220><223> Reverse primer for 2nd PCR reaction
 <400> 11

caagcagaag acggcatacg agatgtgact ggagttcaga cgtgt

45