(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0155953
(43) 공개일자 2024년10월29일

(51) 국제특허분류(Int. Cl.)
C12N 15/86 (2006.01) *C07K 14/005* (2006.01)
C12N 15/113 (2010.01) *C12N 9/12* (2006.01)
C12N 9/22 (2006.01)
(52) CPC특허분류
C12N 15/86 (2013.01)
C07K 14/005 (2013.01)
(21) 출원번호 10-2024-7032958
(22) 출원일자(국제) 2023년03월08일
심사청구일자 없음
(85) 번역문제출일자 2024년10월02일
(86) 국제출원번호 PCT/US2023/063954
(87) 국제공개번호 WO 2023/172966
국제공개일자 2023년09월14일
(30) 우선권주장
63/317,720 2022년03월08일 미국(US)

(71) 출원인
웨이크 포리스트 유니버시티 헬스 사이언시즈
미국 노스캐롤라이나 27157 윈스턴 살렘 메디칼 센터 블러바드
(72) 발명자
루, 바이송
미국 27106 노스캐롤라이나 윈스턴-세일럼 브룩스 랜딩 드라이브 169
아탈라, 앤서니
미국 27104 노스캐롤라이나 윈스턴-세일럼 노스 스태포드 로드 345
(74) 대리인
특허법인(유)남아이피그룹

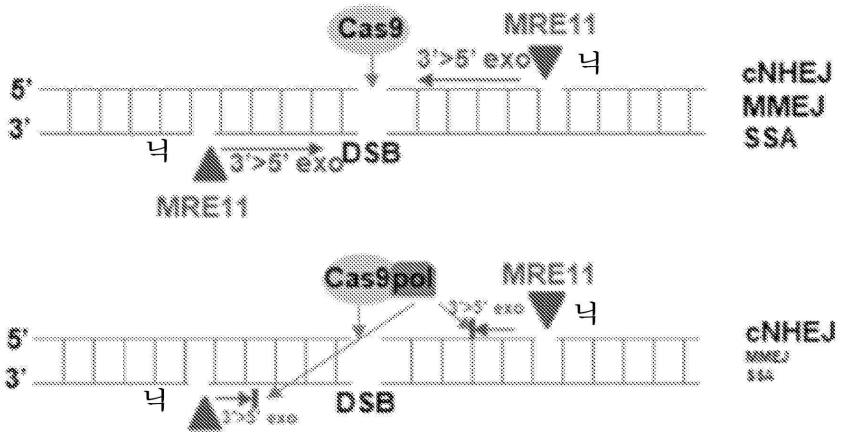
전체 청구항 수 : 총 44 항

(54) 발명의 명칭 **진핵 유전자 편집을 위한 조성물, 시스템 및 방법**

(57) 요 약

CRISPR 성분 코딩 서열, 즉, sgRNA 및 CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 융합 단백질을 세포에 전달하기 위한 포유동물 발현 플라스미드가 본원에 제공된다. 본원에 설명된 임의의 포유동물 발현 플라스미드를 사용하는 방법이 또한 제공된다.

대 표 도 - 도1a



(52) CPC특허분류
C12N 15/113 (2013.01)
C12N 9/1252 (2013.01)
C12N 9/22 (2013.01)
C12Y 207/07007 (2013.01)
C12N 2310/20 (2017.05)
C12N 2740/16043 (2013.01)
C12N 2795/10322 (2013.01)
C12N 2795/18122 (2013.01)

명 세 서

청구범위

청구항 1

비-바이러스 핵산 서열에 작동 가능하게 연결된 진핵 프로모터를 포함하는 포유동물 발현 플라스미드로서, 상기 비-바이러스 핵산 서열이

(i) 융합 단백질을 인코딩하는 핵산 서열로서, 상기 융합 단백질이

(a) DNA 중합효소 도메인을 포함하는 폴리펩티드; 및

(b) CRISPR-관련 엔도뉴클레아제 코딩 서열을 포함하는, 핵산 서열: 및

(ii) 가이드 RNA(gRNA) 코딩 서열로서, 상기 gRNA 코딩 서열이 적어도 하나의 앱타머 코딩 서열을 포함하는, 가이드 RNA(gRNA) 코딩 서열을 포함하는,

포유동물 발현 플라스미드.

청구항 2

제1항에 있어서, CRISPR-관련 엔도뉴클레아제 코딩 서열이 Cas9 단백질을 인코딩하는 포유동물 발현 플라스미드.

청구항 3

제1항에 있어서, DNA 중합효소 도메인을 포함하는 폴리펩티드가 E. 콜리 DNA 중합효소 I(DNA PolI)인 포유동물 발현 플라스미드.

청구항 4

제1항에 있어서, DNA 중합효소 도메인을 포함하는 폴리펩티드가 DNA PolI의 클레노우(Klenow) 단편인 포유동물 발현 플라스미드.

청구항 5

제1항에 있어서, DNA 중합효소 도메인을 포함하는 폴리펩티드가 감소된 중합효소 활성을 갖는 포유동물 발현 플라스미드.

청구항 6

제1항에 있어서, 적어도 하나의 앱타머 코딩 서열이 MS2 코트 단백질, PP7 코트 단백질, 람다 N RNA-결합 도메인, 또는 Com 단백질로 구성된 군으로부터 선택된 ABP에 의해 특이적으로 결합된 앱타머 서열을 인코딩하는 포유동물 발현 플라스미드.

청구항 7

제1항에 있어서, 앱타머가 MS2 앱타머 서열 또는 com 앱타머 서열인 포유동물 발현 플라스미드.

청구항 8

제1항에 있어서, sgRNA 코딩 서열이 sgRNA 코딩 서열의 테트라루프 또는 ST2 루프에 삽입된 적어도 하나의 앱타머를 포함하는 포유동물 발현 플라스미드.

청구항 9

제8항에 있어서, sgRNA 코딩 서열이 gRNA 코딩 서열의 ST2 루프에 삽입된 적어도 하나의 com 앱타머를 포함하는 포유동물 발현 플라스미드.

청구항 10

a) Gag 뉴클레오티드 서열에 작동 가능하게 연결된 진핵 프로모터를 포함하는 패키징 플라스미드로서, 상기 Gag 뉴클레오티드 서열이 뉴클레오캡시드(NC) 코딩 서열 및 매트릭스 단백질(MA) 코딩 서열을 포함하고, 상기 NC 코딩 서열 또는 MA 코딩 서열 중 하나 또는 둘 모두가 적어도 하나의 비-바이러스 앱타머-결합 단백질(ABP) 뉴클레오티드 서열을 포함하고, 상기 패키징 플라스미드가 기능성 인테그라제 단백질을 인코딩하지 않는, 패키징 플라스미드;

b) 제1항 내지 제9항 중 어느 한 항의 적어도 하나의 포유동물 발현 플라스미드; 및

c) 외피 당단백질 코딩 서열을 포함하는 외피 플라스미드를 포함하는,

렌티바이러스 패키징 시스템.

청구항 11

제10항에 있어서, 패키징 플라스미드가 Rev 뉴클레오티드 서열 및 Tat 뉴클레오티드 서열을 추가로 포함하는 렌티바이러스 패키징 시스템.

청구항 12

제10항에 있어서, Rev 뉴클레오티드 서열을 포함하는 제2 패키징 플라스미드를 추가로 포함하는 렌티바이러스 패키징 시스템.

청구항 13

제10항에 있어서, 적어도 하나의 비-바이러스 ABP 뉴클레오티드 서열이 MS2 코트 단백질, PP7 코트 단백질, 람다 N 펩티드, 또는 Com 단백질을 인코딩하는 렌티바이러스 패키징 시스템.

청구항 14

렌티바이러스 입자로서,

A) 뉴클레오캡시드(NC) 단백질 또는 매트릭스(MA) 단백질을 포함하는 융합 단백질로서, 상기 NC 단백질 또는 MA 단백질이 적어도 하나의 비-바이러스 앱타머 결합 단백질(ABP)을 포함하는, 융합 단백질; 및

B) (i) 융합 단백질을 인코딩하는 핵산 서열로서, 상기 융합 단백질이

(a) DNA 중합효소 도메인을 포함하는 폴리펩티드; 및

(b) CRISPR-관련 엔도뉴클레아제 코딩 서열을 포함하는, 핵산 서열; 및

(ii) 가이드 RNA(gRNA) 코딩 서열로서, 상기 gRNA 코딩 서열이 적어도 하나의 앱타머 코딩 서열을 포함하고, 상기 렌티바이러스-유사 입자가 기능성 인테그라제 단백질을 포함하지 않는, 가이드 RNA(gRNA) 코딩 서열을 포함하는, 리보뉴클레오티드 단백질(RNP) 복합체를 포함하는,

렌티바이러스 입자.

청구항 15

제14항에 있어서, CRISPR-관련 엔도뉴클레아제 코딩 서열이 Cas9 단백질을 인코딩하는 렌티바이러스 입자.

청구항 16

제14항에 있어서, DNA 중합효소 도메인을 포함하는 폴리펩티드가 E. 콜리 DNA 중합효소 I(DNA PolI)인 렌티바이러스 입자.

청구항 17

제14항에 있어서, DNA 중합효소 도메인을 포함하는 폴리펩티드가 DNA PolI의 클레노우 단편인 렌티바이러스 입자.

청구항 18

제14항에 있어서, DNA 중합효소 도메인을 포함하는 폴리펩티드가 감소된 중합효소 활성을 갖는 렌티바이러스 입자.

청구항 19

렌티바이러스 입자를 생산하는 방법으로서, 상기 방법이

a) 제10항 내지 제13항 중 어느 한 항의 시스템의 패키징 플라스미드, 적어도 하나의 포유동물 발현 플라스미드, 및 외피 플라스미드로 복수의 진핵 세포를 형질감염시키는 단계; 및

b) 렌티바이러스 입자가 생산되기에 충분한 시간 동안 형질감염된 진핵 세포를 배양하는 단계를 포함하는,

방법.

청구항 20

제19항에 있어서, 렌티바이러스 입자가

(i) 융합 단백질을 인코딩하는 핵산 서열로서, 상기 융합 단백질이

(a) DNA 중합효소 도메인을 포함하는 폴리펩티드; 및

(b) CRISPR-관련 엔도뉴클레아제 코딩 서열을 포함하는, 핵산 서열: 및

(ii) 가이드 RNA를 포함하는, 리보뉴클레오티드 단백질(RNP) 복합체를 포함하는 방법.

청구항 21

제20항에 있어서, 복수의 진핵 세포가 포유동물 세포인 방법.

청구항 22

제19항의 방법에 의해 제조된 렌티바이러스 입자.

청구항 23

세포에서 유전체 표적 서열을 변형시키는 방법으로서, 상기 방법이 복수의 바이러스 입자로 복수의 진핵 세포를 형질도입시키는 단계를 포함하고, 상기 복수의 바이러스 입자가 제14항 내지 제18항 중 어느 한 항에 따른 렌티바이러스-유사 입자를 포함하고, 여기서 RNP 복합체가 세포의 유전체 DNA 내의 유전체 표적 서열에 결합하고, 여기서 CRISPR-관련 엔도뉴클레아제가 유전체 표적 서열을 절단하여 이중-가닥 파손을 생성함으로써 유전체 표적 서열을 변형시키는, 방법.

청구항 24

제23항에 있어서, 비-상동성 말단 연결(NHEJ)이 DNA 중합효소 도메인에 융합되지 않은 CRISPR-관련 엔도뉴클레아제에 의해 변형된 세포와 비교하여 세포에서 증가하는 방법.

청구항 25

제23항에 있어서, NHEJ 대 비-NHEJ의 비가 세포에서 증가하는 방법.

청구항 26

제25항에 있어서, 비-NHEJ 말단 연결이 미세상동성-매개 말단 연결(MMEJ) 및/또는 단일-가닥 어닐링(SSA)인 방법.

청구항 27

제23항에 있어서, 표적-내 결실의 수가 세포에서 감소되는 방법.

청구항 28

제27항에 있어서, 500개 염기쌍 초과의 크기의 표적-내 결실의 수가 세포에서 감소하는 방법.

청구항 29

제23항에 있어서, 표적-내 1 염기쌍(1-bp) 결실 대 1-bp 초과의 표적-내 결실의 비가 세포에서 증가하는 방법.

청구항 30

제29항에 있어서, 표적-내 1 염기쌍(1-bp) 결실 대 500개 염기쌍 초과의 결실의 비가 세포에서 증가하는 방법.

청구항 31

제23항에 있어서, 주형 삽입(templated insertion; TIS)의 수가 세포에서 증가하는 방법.

청구항 32

제31항에 있어서, TIS 대 비-TIS의 비가 세포에서 증가하는 방법.

청구항 33

제23항에 있어서, 복수의 진핵 세포가 포유동물 세포인 방법.

청구항 34

제23항에 있어서, 복수의 진핵 세포가 대상체에 존재하는 세포인 방법.

청구항 35

제34항에 있어서, 대상체가 인간 대상체인 방법.

청구항 36

제35항에 있어서, 대상체가 복수의 바이러스 입자로 주사되는 방법.

청구항 37

제1항의 플라스미드를 함유하는 세포.

청구항 38

제10항의 렌티바이러스 패키징 시스템을 함유하는 세포.

청구항 39

제14항의 렌티바이러스 입자를 함유하는 세포.

청구항 40

제23항의 방법을 사용하여 변형된 세포.

청구항 41

a) 대상체로부터 세포를 획득하는 단계;

b) 제23항 내지 제36항 중 어느 한 항의 방법을 사용하여 대상체의 세포를 변형시키는 단계; 및

c) 변형된 세포를 대상체에게 투여하는 단계를 포함하는,

대상체에서 질환을 치료하기 위한 방법.

청구항 42

제41항에 있어서, 질환이 암인 방법.

청구항 43

제41항에 있어서, 질환이 뒤시엔느 근이영양증인 방법.

청구항 44

제41항에 있어서, 세포가 T 세포인 방법.

발명의 설명

기 술 분 야

[0001] **이전 관련 출원**

[0002] [0001] 본 출원은 2022년 3월 8일에 출원된 미국 가출원 번호 63/317,720호에 대한 이익 및 우선권을 주장하며, 이는 그 전체가 본원에 참조로 포함된다.

[0003] **서열 목록**

[0004] [0002] 본 출원은 XML 포맷의 서열 목록을 포함한다. 095199-1375708.xml로 명명된 서열 목록은 2023년 3월 7일에 작성되었으며, 크기는 175 킬로바이트이며, 그 전체가 본원에 참조로 포함된다.

[0005] **분야**

[0006] [0003] 본 발명의 개시는 진핵 유전자 편집을 위한 조성물 및 이를 사용하는 방법을 설명한다.

배 경 기 술

[0007] [0001] CRISPR-기반 유전체 편집은 예측 불가능하게 표적-내 결실, 예를 들어, 세포의 유전체에서 긴 DNA 결실(예를 들어, >500 bp)을 생성할 수 있다. 이러한 DNA 결실을 생성할 가능성은 체세포 유전체 편집에 안전성 위험을 초래하고, 유전체 편집의 결과를 덜 예측가능하게 만든다. 따라서, 표적-내 결실을 감소시키면서 세포 유전체를 편집하기 위한 조성물 및 방법이 안전하고 효율적인 유전체 편집을 위해 필요하다.

발명의 내용

[0008] [0002] 비-바이러스 핵산 서열에 작동 가능하게 연결된 진핵 프로모터를 포함하는 포유동물 발현 플라스미드가 본원에 제공되며, 여기서 비-바이러스 핵산 서열은 (i) 융합 단백질을 인코딩하는 핵산 서열로서, 상기 융합 단백질이 (a) DNA 중합효소 도메인을 포함하는 폴리펩티드; 및 (b) CRISPR-관련 엔도뉴클레아제 코딩 서열을 포함하는, 핵산 서열; 및 (ii) 가이드 RNA(gRNA) 코딩 서열로서, 상기 gRNA 코딩 서열이 적어도 하나의 앱타머 코딩 서열을 포함하는, 가이드 RNA(gRNA) 코딩 서열을 포함한다.

[0009] [0003] 일부 구현예에서, DNA 중합효소 도메인을 포함하는 폴리펩티드는 E. 콜리(E. coli) DNA 중합효소 I(DNA Pol I) 또는 이의 단편을 포함한다. 일부 구현예에서, DNA 중합효소 도메인을 포함하는 폴리펩티드는 E. 콜리 DNA 중합효소 I(DNA Pol I)의 클레노우(Klenow) 단편을 포함한다.

[0010] [0004] 일부 구현예에서, CRISPR-관련 엔도뉴클레아제 코딩 서열은 Cas9 단백질 또는 Cpf1 단백질을 인코딩한다.

[0011] [0005] 일부 구현예에서, 적어도 하나의 앱타머 코딩 서열은 MS2 코트 단백질, PP7 코트 단백질, 람다 N RNA-결합 도메인, 또는 Com 단백질로 구성된 군으로부터 선택된 ABP에 의해 특이적으로 결합된 앱타머 서열을 인코딩한다. 일부 구현예에서, 앱타머는 MS2 앱타머 서열 또는 com 앱타머 서열이다. 일부 구현예에서, sgRNA 코딩 서열은 sgRNA 코딩 서열의 테트라루프 또는 ST2 루프에 삽입된 적어도 하나의 앱타머를 포함한다. 일부 구현예에서, sgRNA 코딩은 gRNA 코딩 서열의 ST2 루프에 삽입된 적어도 하나의 com 앱타머를 포함한다.

[0012] [0006] a) Gag 뉴클레오티드 서열에 작동 가능하게 연결된 진핵 프로모터를 포함하는 패키징 플라스미드로서, 상기 Gag 뉴클레오티드 서열이 뉴클레오캡시드(NC) 코딩 서열 및 매트릭스 단백질(MA) 코딩 서열을 포함하고, 상기 NC 코딩 서열 또는 MA 코딩 서열 중 하나 또는 둘 모두가 적어도 하나의 비-바이러스 앱타머-결합 단백질(ABP) 뉴클레오티드 서열을 포함하고, 상기 패키징 플라스미드가 기능성 인테그라제 단백질을 인코딩하지 않는, 패키징 플라스미드; b) 본원에 제공된 적어도 하나의 포유동물 발현 플라스미드; 및 c) 외피 당단백질 코딩 서열을 포함하는 외피 플라스미드를 포함하는 렌티바이러스 패키징 시스템이 또한 제공된다.

[0013] [0007] 일부 구현예에서, 패키징 플라스미드는 Rev 뉴클레오티드 서열 및 Tat 뉴클레오티드 서열을 추가로 포함

한다. 일부 구현예에서, 시스템은 Rev 뉴클레오티드 서열을 포함하는 제2 패키징 플라스미드를 추가로 포함한다. 일부 구현예에서, 적어도 하나의 비-바이러스 ABP 뉴클레오티드 서열은 MS2 코트 단백질, PP7 코트 단백질, 람다 N 펩티드, 또는 Com 단백질을 인코딩한다.

[0014] **[0008]** (a) 뉴클레오캡시드(NC) 단백질 또는 매트릭스(MA) 단백질을 포함하는 융합 단백질로서, 상기 NC 단백질 또는 MA 단백질이 적어도 하나의 비-바이러스 앱타머 결합 단백질(ABP)을 포함하는, 융합 단백질; 및 (b) (i) 융합 단백질을 인코딩하는 핵산 서열로서, 상기 융합 단백질이 (a) DNA 중합효소 도메인을 포함하는 폴리펩티드, 및 (b) CRISPR-관련 엔도뉴클레아제 코딩 서열을 포함하는, 핵산 서열; 및 (ii) 가이드 RNA(gRNA) 코딩 서열로서, 상기 gRNA 코딩 서열이 적어도 하나의 앱타머 코딩 서열을 포함하고, 여기서 렌티바이러스-유사 입자가 기능성 인테그라제 단백질을 포함하지 않는, 가이드 RNA(gRNA) 코딩 서열을 포함하는, 리보뉴클레오티드 단백질(RNP) 복합체를 포함하는 렌티바이러스 입자가 추가로 제공된다.

[0015] **[0009]** 일부 구현예에서, CRISPR-관련 엔도뉴클레아제 코딩 서열은 Cas9 단백질 또는 Cpf1 단백질을 인코딩한다.

[0016] **[0010]** 일부 구현예에서, 융합 폴리펩티드는 E. 콜리 DNA Pol I을 포함한다. 일부 구현예에서, 폴리펩티드는 DNA Pol I의 클레노우 단편을 포함한다.

[0017] **[0011]** 또한, 렌티바이러스 입자를 생산하는 방법이 제공되며, 상기 방법은 a) 복수의 진핵 세포를 본원에 설명된 임의의 시스템의 패키징 플라스미드, 적어도 하나의 포유동물 발현 플라스미드, 및 외피 플라스미드로 형질감염시키고; b) 렌티바이러스 입자가 생산되기에 충분한 시간 동안 형질감염된 진핵 세포를 배양하는 것을 포함한다.

[0018] **[0012]** 일부 구현예에서, 렌티바이러스 입자는 (i) 융합 단백질을 인코딩하는 핵산 서열로서, 상기 융합 단백질이 (a) DNA 중합효소 도메인을 포함하는 폴리펩티드 및 (b) CRISPR-관련 엔도뉴클레아제 코딩 서열을 포함하는, 핵산 서열; 및 (ii) 가이드 RNA를 포함하는 리보뉴클레오티드 단백질(RNP) 복합체를 포함한다. 일부 구현예에서, 복수의 진핵 세포는 포유동물 세포이다.

[0019] **[0013]** 또한, 본원에 설명된 임의의 플라스미드, 렌티바이러스 패키징 시스템 또는 렌티바이러스-유사 입자를 포함하는 세포가 제공된다. 본원에 제공된 임의의 방법에 의해 변형된 세포가 또한 제공된다.

[0020] **[0014]** 세포에서 유전체 표적 서열을 변형시키는 방법이 추가로 제공되며, 상기 방법은 복수의 진핵 세포에 복수의 바이러스 입자를 형질도입시키는 것을 포함하고, 상기 복수의 바이러스 입자는 (a) NC 단백질 또는 MA 단백질을 포함하는 융합 단백질로서, 상기 NC 단백질 또는 MA 단백질이 적어도 하나의 비-바이러스 ABP를 포함하는, 융합 단백질; 및 (b) (i) 융합 단백질을 인코딩하는 핵산 서열로서, 상기 융합 단백질이 (a) DNA 중합효소 도메인을 포함하는 폴리펩티드, 및 (b) CRISPR-관련 엔도뉴클레아제 코딩 서열을 포함하는, 핵산 서열; 및 (ii) gRNA 코딩 서열로서, 상기 gRNA 코딩 서열이 적어도 하나의 앱타머 코딩 서열을 포함하고, 여기서 렌티바이러스-유사 입자가 기능성 인테그라제 단백질을 포함하지 않는, gRNA 코딩 서열을 포함하는 RNP 복합체로서, 상기 RNP 복합체가 세포의 유전체 DNA 내의 유전체 표적 서열에 결합하고, 상기 CRISPR-관련 엔도뉴클레아제가 유전체 표적 서열을 절단하여 이중-가닥 절단을 생성하고, 이에 의해 유전체 표적 서열을 변형시키는, RNP 복합체를 포함하는, 본원에 제공된 바와 같은 렌티바이러스-유사 입자를 포함한다.

[0021] **[0015]** 일부 구현예에서, NJEH는 DNA 중합효소 도메인에 융합되지 않은 CRISPR-관련 엔도뉴클레아제에 의해 변형된 세포와 비교하여 세포에서 비-NHEJ 말단 연결과 비교하여 증가한다. 일부 구현예에서, 비-NJEH는 미세상동성-매개 말단 연결(MMEJ) 및/또는 단일-가닥 어닐링(SSA)이다.

[0022] **[0016]** 일부 구현예에서, 500개 염기쌍 초과의 크기의 표적-내 결실의 수는 일반 세포와 비교하여 상기 세포에서 감소한다. 일부 구현예에서, 표적-내 1 염기쌍(1-bp) 결실 대 1-bp 초과의 표적-내 결실의 비는 상기 세포에서 증가한다. 일부 구현예에서, 표적-내 1 염기쌍(1-bp) 결실 대 500개 염기쌍 초과의 결실의 비는 상기 세포에서 증가한다. 일부 구현예에서, 주형 삽입(templated insertions; TIS)의 수는 상기 세포에서 증가한다. 일부 구현예에서, TIS 대 비-TIS의 비는 상기 세포에서 증가한다.

[0023] **[0017]** 일부 구현예에서, 복수의 진핵 세포는 포유동물 세포이다. 일부 구현예에서, 복수의 진핵 세포는 대상체에 존재하는 세포이다. 일부 구현예에서, 대상체는 인간 대상체이다. 일부 구현예에서, 대상체는 복수의 바이러스 입자로 주사된다.

[0024] **[0018]** 또한, a) 대상체로부터 세포를 획득하고; b) 본원에 제공된 바와 같은 렌티바이러스-유사 입자를 사용하

여 본원에 제공된 방법을 사용하여 대상체의 세포를 변형시키고; c) 변형된 세포를 대상체에게 투여하는 것을 포함하는 대상체에서 질환을 치료하는 방법이 제공된다. 일부 구현예에서, 질환은 암이다. 일부 구현예에서, 질환은 뒤시엔느(Duchenne) 근이영양증이다. 일부 구현예에서, 세포는 T 세포이다.

도면의 간단한 설명

[0025] **[0019]** 본 출원은 하기 도면을 포함한다. 도면은 조성물 및 방법의 특정 구현예 및/또는 특징을 예시하고, 조성물 및 방법의 임의의 설명(들)을 보충하기 위한 것이다. 도면은 서면 설명이 그러한 경우를 명시적으로 나타내지 않는 한, 조성물 및 방법의 범위를 제한하지 않는다.

[0020] 도 1a는 Cas9에 융합된 DNA 중합효소 I(DNA Pol I 또는 pol I)에 의한 (예를 들어, MRE11에 의한) 대응하는 DNA 절제를 보여주는 도표이다. 예상되는 결과는 DNA 절제를 필요로 하는 미세상동성-매개 말단 연결(MMEJ) 및 단일 가닥 어닐링(SSA) DNA 복구 경로의 억제이다.

[0021] 도 1b는 DNA 중합효소 1이 5' 오버행의 채움을 통해 하나의 염기쌍(1-bp) 삽입(TIS)을 생성한다는 것을 보여준다. 적색 뉴클레오티드는 DNA 중합효소에 의해 채워진다. *CLCN5* sgRNA((GAGGACAAGTCGTACAATGGTGG)(SEQ ID NO: 111) 및 이의 보체(CTCCTGTTCAGCATGTTACCACC)(SEQ ID NO: 112))의 표적 부위를 예로서 사용하였다. 두 가닥 모두의 절단 부위는 작은 화살표로 표시된 바와 같이 하나의 뉴클레오티드(1-nt) 5' 오버행을 생성한다.

[0022] 도 1c는 Cas9에 대한 DNA pol I의 융합이 HEK293T 세포에서 1-bp 결실의 백분율을 증가시키고 *CLCN5* 유전자를 표적화하는 >1-bp 결실을 감소시켰음을 보여주는 그래프이다. 이원 ANOVA에 이어 본페로니(Bonferroni) 사후검정이 수행되었다. 복제 수는 괄호 안에 나열되어 있다.

[0023] 도 1d는 DNA pol I의 Cas9으로의 융합(Cas9-pol I)이 1-bp TIS 대 1-bp 비-TIS의 비를 증가시켰음을 보여주는 그래프이다(양측 t-검정).

[0024] 도 2a는 실시예에 설명된 연구에서 시험된 Cas9-pol I 및 다양한 돌연변이체 융합 단백질을 보여주는 예시이다. 점선은 결실된 영역을 나타낸다.

[0025] 도 2b는 HEK293T 세포에서 *CLCN5* 유전자를 표적화하는 결실 프로파일에 대한 상이한 DNA pol I 돌연변이체의 효과를 나타내는 그래프이다. 괄호 안의 숫자는 생물학적 복제 수를 나타낸다. 적색 점선 위의 그룹은 선 아래의 그룹과 비교하여 통계적 유의성을 나타냈다.

[0026] 도 2c는 HEK293T 세포에서 *CLCN5* 유전자를 표적화하는 삽입에 대한 상이한 DNA pol I 돌연변이체의 효과를 나타내는 그래프이다. 괄호 안의 숫자는 생물학적 복제 수를 나타낸다. 적색 점선 위의 그룹은 선 아래의 그룹과 비교하여 통계적 유의성을 나타냈다.

[0027] 도 2d는 IMR90 세포에서 *CLCN5* 유전자를 표적화하는 결실 프로파일에 대한 상이한 DNA pol I 돌연변이체의 효과를 나타내는 그래프이다. 복제 수는 모든 그룹에 대해 3이었다. * 및 **는 표시된 그룹과 모든 다른 그룹 사이에서 각각 $p<0.05$ 및 $p<0.01$을 나타낸다(이원 ANOVA 후 본페로니 사후검정).

[0028] 도 2e는 IMR90 세포에서 *CLCN5* 유전자를 표적화하는 삽입에 대한 상이한 DNA pol I 돌연변이체의 효과를 나타내는 그래프이다. 복제 수는 모든 그룹에 대해 3이었다. * 및 **는 표시된 그룹과 모든 다른 그룹 사이에서 각각 $p<0.05$ 및 $p<0.01$을 나타낸다(이원 ANOVA 후 본페로니 사후검정).

[0029] 도 3a는 HEK293T 세포에서 *CLCN5*를 표적화하는 Cas9-유도된 DNA 돌연변이 프로파일에 대한 *RBBP8* 녹다운의 효과를 나타내는 그래프이다. 복제 수는 Cas9의 경우 6 및 나머지 그룹의 경우 3이었다. *, ** 및 ***는 표시된 그룹 사이의 $p<0.05$, $p<0.01$ 및 $p<0.001$을 나타낸다(이원 ANOVA 후 본페로니 사후-검정).

[0030] 도 3b는 HEK293T 세포에서 *CLCN5*를 표적화하는 Cas9-유도된 삽입에 대한 *RBBP8* 녹다운의 효과를 나타내는 그래프이다. 복제 수는 Cas9의 경우 6 및 나머지 그룹의 경우 3이었다. *, ** 및 ***는 표시된 그룹 사이의 $p<0.05$, $p<0.01$ 및 $p<0.001$을 나타낸다(이원 ANOVA 후 본페로니 사후-검정).

[0031] 도 3c는 HEK293T 세포에서 *CLCN5*를 표적화하는 다양한 크기를 갖는 결실에 대한 *RBBP8* 녹다운의 효과를 나타내는 그래프이다. 단일 및 이중 화살표는 각각 절제-의존적 및 독립적인 결실을 나타낸다.

[0032] 도 3d는 HEK293T 세포에서 *CLCN5*를 표적화하는 Cas9에 의해 생성된 가장 빈번하게 관찰된 결실을 보여준다. 부분 야생형 *CLCN5* 서열(SEQ ID NO: 113) 뿐만 아니라 11개의 염기쌍 결실을 포함하는 SEQ ID NO: 113(SEQ ID NO: 114), 상이한 11개의 염기쌍 결실을 포함하는 SEQ ID NO: 113(SEQ ID NO: 115), 8개의 염기쌍 결실을

포함하는 SEQ ID NO: 113(SEQ ID NO: 116), 상이한 8개의 염기쌍 결실을 포함하는 SEQ ID NO: 113(SEQ ID NO: 117), 및 16개의 염기쌍 결실을 포함하는 SEQ ID NO: 113(SEQ ID NO: 118)이 제시된다. 점선은 결실을 나타낸다. 미세상동성은 밑줄이 그어져 있다. 절단 부위로부터 떨어져 있는 미세상동성은 캐럿 기호(^)로 표시된다. 별표는 수직 점선으로 표시되는 예측된 절단 크기에서의 미세상동성을 나타낸다.

[0033] 도 4a는 조혈 세포에서 *HBB* 유전자를 표적으로 하는 억제된 결실의 조사를 보여준다. * 및 ***는 Cas9와 Cas9-클레노우 사이에서 각각 $p<0.05$ 및 $p<0.001$을 나타낸다(이원 ANOVA 후 본페로니 사후검정). 부분 야생형 HBB 서열(SEQ ID NO: 119) 뿐만 아니라 3개의 염기쌍 결실을 포함하는 SEQ ID NO: 119(SEQ ID NO: 120) 및 12개의 염기쌍 결실을 포함하는 SEQ ID NO: 119(SEQ ID NO: 121)이 제시된다. 상단 이미지는 결실의 피크를 나타내고 하단 이미지는 가장 빈번하게 관찰된 결실을 나타낸다. 미세상동성은 밑줄이 그어져 있다. 별표(*)는 수직 점선으로 표시되는 예측된 절단 크기에서의 미세상동성을 나타낸다. 캐럿(^)은 예측된 절단 부위로부터 떨어져 있는 미세상동성을 나타낸다. 각 그룹은 3개의 생물학적 복제를 갖는다.

[0034] 도 4b는 HEK293T 세포에서 *DMD* 엑손 53을 표적화하는 억제된 결실의 조사를 보여준다. * 및 ***는 Cas9와 Cas9-클레노우 사이에서 각각 $p<0.05$ 및 $p<0.001$을 나타낸다(이원 ANOVA 후 본페로니 사후검정). 부분 야생형 DMD 엑손 53 서열(SEQ ID NO: 122) 뿐만 아니라 11개의 염기쌍 결실을 포함하는 SEQ ID NO: 122(SEQ ID NO: 123), 9개의 염기쌍 결실을 포함하는 SEQ ID NO: 122(SEQ ID NO: 124), 및 6개의 염기쌍 결실을 갖는 SEQ ID NO: 122(SEQ ID NO: 125)가 제시된다. 상단 이미지는 결실의 피크를 나타내고 하단 이미지는 가장 빈번하게 관찰된 결실을 나타낸다. 미세상동성은 밑줄이 그어져 있다. 별표(*)는 수직 점선으로 표시되는 예측된 절단 크기에서의 미세상동성을 나타낸다. 캐럿(^)은 예측된 절단 부위로부터 떨어져 있는 미세상동성을 나타낸다. 각 그룹은 3개의 생물학적 복제를 갖는다.

[0035] 도 4c는 IMR90 세포에서 *HBB* 유전자를 표적으로 하는 억제된 결실의 조사를 보여준다. * 및 ***는 Cas9와 Cas9-클레노우 사이에서 각각 $p<0.05$ 및 $p<0.001$을 나타낸다(이원 ANOVA 후 본페로니 사후검정). 부분 야생형 *HBB* 서열(SEQ ID NO: 126) 뿐만 아니라 3개의 염기쌍 결실을 포함하는 SEQ ID NO: 126(SEQ ID NO: 127) 및 5개의 염기쌍 결실을 포함하는 SEQ ID NO: 126(SEQ ID NO: 128)이 제시된다. 상단 이미지는 결실의 피크를 나타내고 하단 이미지는 가장 빈번하게 관찰된 결실을 나타낸다. 미세상동성은 밑줄이 그어져 있다. 별표(*)는 수직 점선으로 표시되는 예측된 절단 크기에서의 미세상동성을 나타낸다. 캐럿(^)은 예측된 절단 부위로부터 떨어져 있는 미세상동성을 나타낸다. 각 그룹은 3개의 생물학적 복제를 갖는다.

[0036] 도 4d는 인간 근모세포에서 *DMD* 엑손 44를 표적으로 하는 억제된 결실의 조사를 보여준다. * 및 ***는 Cas9와 Cas9-클레노우 사이에서 각각 $p<0.05$ 및 $p<0.001$을 나타낸다(이원 ANOVA 후 본페로니 사후검정). 부분 야생형 *DMD44* 서열(SEQ ID NO: 129) 뿐만 아니라 10개의 염기쌍 결실을 포함하는 SEQ ID NO: 129(SEQ ID NO: 130), 7개의 염기쌍 결실을 포함하는 SEQ ID NO: 129(SEQ ID NO: 131), 및 상이한 7개 염기쌍 결실을 포함하는 SEQ ID NO: 129(SEQ ID NO: 132)가 제시된다. 상단 이미지는 결실의 피크를 나타내고 하단 이미지는 가장 빈번하게 관찰된 결실을 나타낸다. 미세상동성은 밑줄이 그어져 있다. 별표(*)는 수직 점선으로 표시되는 예측된 절단 크기에서의 미세상동성을 나타낸다. 캐럿(^)은 예측된 절단 부위로부터 떨어져 있는 미세상동성을 나타낸다. 각 그룹은 3개의 생물학적 복제를 갖는다.

[0037] 도 5a는 HEK293T 세포에서 *CLCN5* 유전자를 표적화하는 Cas9 및 Cas9-클레노우에 의해 생성된 큰 결실을 보여준다. 데이터를 3개의 복제로부터 조합하였다. 434 및 4534로 표지된 영역은 결실 검출을 위한 거리를 계산하는데 사용된 2개의 25-bp 서열을 나타낸다. 2471, 2481, 및 2487로 표지된 영역은 sgRNA 표적을 나타내고, 2494로 표지된 영역은 PAM을 나타낸다. 별표는 2개의 상이한 실험에서 독립적으로 관찰된 동일한 결실을 나타낸다.

[0038] 도 5b는 IMR90 세포에서 *CLCN5* 유전자를 표적화하는 Cas9 및 Cas9-클레노우에 의해 생성된 큰 결실을 보여준다. 데이터를 3개의 복제로부터 조합하였다. 434 및 4534로 표지된 영역은 결실 검출을 위한 거리를 계산하는데 사용된 2개의 25-bp 서열을 나타낸다. 2471, 2489, 및 2488로 표지된 영역은 sgRNA 표적을 나타내고, 2494로 표지된 영역은 PAM을 나타낸다. 별표는 2개의 상이한 실험에서 독립적으로 관찰된 동일한 결실을 나타낸다.

[0039] 도 6은 *CLCN5* sgRNA 및 Cas9-pol로 처리된 GFP-리포터 세포에서 통합된 표적 서열의 차세대 시퀀싱(NGS) 분석이다. SEQ ID NOs: 133-141이 제시된다.

[0040] 도 7은 Cas9 및 다양한 엑소뉴클레아제 융합체가 Cas9와 유사한 돌연변이 프로파일을 가졌음을 보여주는

그래프이다. 괄호 안의 숫자는 복제 수를 나타낸다.

[0041] 도 8a는 2-bp TIS에 대한 상이한 pol I 도메인의 효과를 나타내는 그래프이다. HEK293T 세포에서 *CLCN5*를 표적화하는 exo, 3' exo 및 5' exo 융합 단백질에 대한 데이터를 하나의 그룹으로 푸울링하였다. 각 점은 하나의 기준점을 나타낸다. ** 및 ***는 Cas9 그룹과 비교하여 각각 $p<0.05$ 및 $p<0.001$을 나타낸다. 터키(Tukey)의 다중 비교 검정을 일원 ANOVA 후에 수행하였다. Cas9-pol은 증가 경향을 나타내었지만 큰 그룹내 변동으로 인해 통계적 유의성에 도달하지 않았다.

[0042] 도 8b는 3-bp TIS에 대한 상이한 pol I 도메인의 효과를 나타내는 그래프이다. HEK293T 세포에서 *CLCN5*를 표적화하는 exo, 3' exo 및 5' exo 융합 단백질에 대한 데이터를 하나의 그룹으로 푸울링하였다. 각 점은 하나의 기준점을 나타낸다. ** 및 ***는 Cas9 그룹과 비교하여 각각 $p<0.05$ 및 $p<0.001$을 나타낸다. 터키 다중 비교 시험을 일원 ANOVA 후에 수행하였다. Cas9-pol은 증가 경향을 나타내었지만 큰 그룹내 변동으로 인해 통계적 유의성에 도달하지 않았다.

[0043] 도 9는 전체 INDEL 비율이 상이한 돌연변이 유형의 백분율에 영향을 미치지 않았음을 보여주는 그래프이다.

[0044] 도 10a는 IMR90 세포에서 *CLCN5* 돌연변이 프로파일(모든 INDEL의 %)에 대한 *MRE11* 또는 *RBBP8* 녹다운의 효과를 나타내는 그래프이다. 각 그룹은 3개의 생물학적 복제를 가졌다. 이원 ANOVA에 이어 본페로니 사후검정이 수행되었다. ** 및 ***는 표시된 그룹 사이에서 각각 $p<0.01$ 및 $p<0.001$을 나타낸다.

[0045] 도 10b는 IMR90 세포에서 *CLCN5* 돌연변이 프로파일(모든 삽입의 %)에 대한 *MRE11* 또는 *RBBP8* 녹다운의 효과를 나타내는 그래프이다. 각 그룹은 3개의 생물학적 복제를 가졌다. 이원 ANOVA에 이어 본페로니 사후검정이 수행되었다. ** 및 ***는 표시된 그룹 사이에서 각각 $p<0.01$ 및 $p<0.001$을 나타낸다.

[0046] 도 11a는 HEK293T 세포에서 *HBB*를 표적화할 때 Cas9-클레노우에 의해 감소된 결실을 보여준다. 그래프(상단 이미지)는 결실 크기의 백분율을 보여준다. 하단 이미지는 가장 흔히 관찰되는 결실의 서열을 보여준다. 부분적인 야생형 *HBB* 서열(SEQ ID NO: 142) 뿐만 아니라 3개의 염기쌍 결실을 포함하는 SEQ ID NO: 142(SEQ ID NO: 143), 5개의 염기쌍 결실을 포함하는 SEQ ID NO: 142(SEQ ID NO: 144), 및 10개의 염기쌍 결실을 포함하는 SEQ ID NO: 142(SEQ ID NO: 145)가 제시된다. 녹색 선으로 밑줄이 그어진 영역은 수직 점선으로 표시되는 예측된 절단 부위에서의 미세상동성을 나타낸다. 이원 ANOVA에 이어 본페로니 사후검정이 수행되었다. * 및 ***는 두 그룹 사이에서 각각 $p<0.05$ 및 $P<0.001$을 나타낸다.

[0047] 도 11b는 IMR90 세포에서 유전자간 1을 표적화할 때 Cas9-클레노우에 의해 감소된 결실을 보여준다. 그래프(상단 이미지)는 결실 크기의 백분율을 보여준다. 하단 이미지는 가장 흔히 관찰되는 결실의 서열을 보여준다. 부분 야생형 유전자간 1 서열(SEQ ID NO: 146) 뿐만 아니라 4개의 염기쌍 결실을 포함하는 SEQ ID NO: 146(SEQ ID NO: 147)이 제시된다. sgRNA는 음영처리되어 있다(회색으로 표시됨). PAM은 윗줄로 표시된다.

[0048] 도 11c는 IMR90 세포에서 *CLCN5*를 표적화할 때 Cas9-클레노우에 의해 감소된 결실을 보여준다. 그래프(상단 이미지)는 결실 크기의 백분율을 보여준다. 하단 이미지는 가장 흔히 관찰되는 결실의 서열을 보여준다. 주요 결실 피크가 관찰되지 않았기 때문에 *CLCN5*/IMR90에 대한 서열은 나열될 수 없었다. sgRNA는 음영처리되어 있다(회색으로 표시됨). PAM은 윗줄로 표시된다.

발명을 실시하기 위한 구체적인 내용

[0026] **정의**

[0027] **[0049]** 본 명세서 및 첨부된 청구범위에서 사용되는 단수 형태는 문맥에서 명백히 달리 지시하지 않는 한 복수의 지시대상을 포함한다.

[0028] **[0050]** 본원에 제공된 임의의 및 모든 예 또는 예시적인 언어(예를 들어, "예를 들어")의 사용은 단지 본 발명을 더 잘 예시하기 위한 것이며, 달리 청구되지 않는 한 본 발명의 범위를 제한하지 않는다.

[0029] **[0051]** 용어 "~일 수 있다(may)", "~일 수 있다(may be)", "~할 수 있다(can)", 및 "~될 수 있다(can be)", 및 관련 용어는 관련된 주제가 선택적이고(즉, 주제는 일부 예에 존재하고 다른 예에 존재하지 않음), 문맥이 명백하게 달리 지시하지 않는 한, 주제의 능력 또는 확률에 대한 언급이 아닌 것을 전달하기 위한 것이다.

[0030] **[0052]** 용어 "선택적" 및 "선택적으로"는 후속하여 설명되는 사건, 상황, 또는 물질이 발생하거나 존재할 수 있거나 발생하지 않거나 존재하지 않을 수 있고, 상기 설명이 사건, 상황 또는 물질이 발생하거나 존재하는 경우

뿐만 아니라 발생하지 않거나 존재하지 않는 경우를 포함함을 의미한다.

[0031] **[0053]** 본원에서 용어 "포함하는(including)", "포함하는(comprising)" 또는 "갖는", 및 이들의 변형의 사용은 추가의 요소 뿐만 아니라 이후에 열거된 요소 및 이의 등가물을 포함하는 것을 의미한다. 특정 요소를 "포함하는(including)", "포함하는(comprising)", 또는 "갖는" 것으로 언급된 구현예는 또한 이러한 특정 요소를 "필수적 요소로 하여 구성되는(consisting essentially of)" 및 "구성되는" 것으로 고려된다. 본원에서 사용되는 "및/또는"은 관련된 열거된 항목 중 하나 이상의 임의의 및 모든 가능한 조합 뿐만 아니라 대안으로 해석되는 조합의 부재("또는")를 지칭하고 포함한다.

[0032] **[0054]** 본원에서 사용되는 과도기 어구 "필수적 요소로 하여 구성되는"(및 문법적 변형)은 인용된 물질 또는 단계 및 청구된 발명의 "기본적이고 신규한 특성(들)에 실질적으로 영향을 미치지 않는 것들"을 포함하는 것으로 해석되어야 한다. 문헌[In re Herz, 537 F.2d 549, 551-52, 190 U.S.P.Q. 461, 463 (CCPA 1976)](원문에서 강조)을 참조하고; 또한, MPEP §2111.03을 참조한다. 따라서, 본원에서 사용되는 용어 "필수적 요소로 하여 구성되는"은 "포함하는"과 동등한 것으로 해석되어서는 안 된다.

[0033] **[0055]** 용어 "핵산" 또는 "뉴클레오티드"는 단일- 또는 이중-가닥 형태의 데옥시리보핵산(DNA) 또는 리보핵산(RNA)(예를 들어, mRNA) 및 이들의 중합체를 지칭한다. RNA가 설명될 때, 이의 상응하는 DNA가 또한 설명되며, 여기서 우리딘은 티미딘으로 표시되는 것으로 이해된다. 유사하게, DNA가 설명될 때, 이의 상응하는 RNA가 또한 설명되며, 여기서 티미딘은 우리딘으로 표시된다. 구체적으로 제한되지 않는 한, 용어는 참조 핵산과 유사한 결합 특성을 갖고 천연 발생 뉴클레오티드와 유사한 방식으로 대사되는 천연 뉴클레오티드의 공지된 유사체를 함유하는 핵산을 포함한다. 달리 지시되지 않는 한, 특정 핵산 서열은 또한 이의 보존적으로 변형된 변이체(예를 들어, 축퇴 코돈 치환), 대립유전자, 오솔로그, SNP, 및 상보적 서열 뿐만 아니라 명시적으로 지시된 서열을 암시적으로 포함한다. 구체적으로, 축퇴 코돈 치환은 하나 이상의 선택된(또는 모든) 코돈의 제3 위치가 혼합-염기 및/또는 데옥시이노신 잔기로 치환된 서열을 생성함으로써 달성될 수 있다(Batzer et al., Nucleic Acid Res. 19:5081 (1991); Ohtsuka et al., J. Biol. Chem. 260:2605-2608 (1985); 및 Rossolini et al., Mol. Cell. Probes 8:91-98 (1994)). 본 발명의 폴리뉴클레오티드는 또한 단일-가닥 형태, 이중-가닥 형태, 헤어핀, 스템-앤-루프 구조 등을 포함하나 이에 제한되지 않는 모든 형태의 서열을 포함한다.

[0034] **[0056]** 용어 "유전자"는 폴리펩티드 사슬을 생산하거나 인코딩하는 데 관여하는 DNA의 세그먼트를 지칭할 수 있다. 이는 코딩 영역(리더 및 트레일러)의 앞뒤 영역 뿐만 아니라 개별 코딩 세그먼트(엑손) 사이의 개재 서열(인트론)을 포함할 수 있다. 대안적으로, 용어 "유전자"는 rRNA, tRNA, 가이드 RNA, 또는 마이크로 RNA와 같은 비-번역된 RNA를 생산하거나 인코딩하는 데 관여하는 DNA의 세그먼트를 지칭할 수 있다.

[0035] **[0057]** "치료하는"은 경감; 관해; 증상의 감소 또는 질환 상태를 환자가 더 견딜 수 있게 함; 변성 또는 저하 속도의 둔화; 또는 변성의 최종 지점을 덜 쇠약하게 함과 같은 임의의 객관적 또는 주관적 파라미터를 포함하는 질환, 병태, 또는 장애의 치료 또는 개선 또는 예방에서의 임의의 성공 지표를 지칭한다. 증상의 치료 또는 개선은 의사의 검사 결과를 포함하는 객관적 또는 주관적 파라미터에 기초할 수 있다. 따라서, 용어 "치료하는"은 본원에 설명된 바와 같은 질환, 병태 또는 장애와 관련된 증상 또는 상태의 발달을 예방 또는 지연, 경감, 저지 또는 억제하기 위한 본 발명의 개시의 화합물 또는 제제의 투여를 포함한다. 용어 "치료 효과"는 대상체에서 질환, 질환의 증상, 또는 질환의 부작용의 감소, 제거 또는 예방을 지칭한다. 본 발명의 개시의 방법을 사용하여 "치료하는" 또는 "치료"는 본원에 설명된 바와 같은 질환, 병태 또는 장애와 관련된 질환 또는 장애의 위험이 증가될 수 있으나, 아직 증상을 경험하거나 나타내지 않은 대상체에서 증상의 개시를 예방하고, 질환 또는 장애의 증상을 억제하고(이의 발달을 늦추거나 저지함), 질환의 증상 또는 부작용으로부터 경감을 제공하고(고식적 치료를 포함함), 질환의 증상을 경감(퇴행 유발)하는 것을 포함한다. 치료는 예방적(질환의 발병을 예방 또는 지연시키거나, 이의 임상적 또는 무증상 증상의 징후를 예방하기 위해) 또는 질환 또는 병태의 소견 후 증상의 치료적 억제 또는 완화일 수 있다. 본원에서 사용되는 용어 "치료"는 예방적(예를 들어, 예방적), 치유적 또는 고식적 치료를 포함한다.

[0036] **[0058]** "프로모터"는 핵산의 전사를 지시하는 하나 이상의 핵산 조절 서열로서 정의된다. 본원에서 사용되는 프로모터는 전사의 시작 부위 근처에 필요한 핵산 서열, 예를 들어, 중합효소 II 유형 프로모터의 경우, TATA 요소를 포함한다. 프로모터는 또한 선택적으로 전사의 시작 부위로부터 수천 개의 염기쌍만큼 위치할 수 있는 원위 인핸서 또는 억제인자 요소를 포함한다.

[0037] **[0059]** "폴리펩티드", "펩티드", 및 "단백질"은 아미노산 잔기의 중합체를 지칭하기 위해 본원에서 상호교환적으로 사용된다. 세 가지 용어 모두는 하나 이상의 아미노산 잔기가 상응하는 천연 발생 아미노산의 인공 화학적

모방체인 아미노산 중합체 뿐만 아니라 천연 발생 아미노산 중합체 및 비-천연 발생 아미노산 중합체에 적용된다. 본원에서 사용되는 용어는 전장 단백질, 트렁케이션된 단백질, 및 이의 단편, 및 아미노산 사슬을 포함하며, 여기서 아미노산 잔기는 공유 펩티드 결합에 의해 연결된다. 전체에 걸쳐 사용되는 용어 "융합 폴리펩티드" 또는 "융합 단백질"은 2개 이상의 단백질 또는 이의 단편을 포함하는 폴리펩티드이다. 일부 구현예에서, 약 3 내지 10개의 아미노산을 포함하는 링커는 발현시 단백질의 적절한 폴딩을 용이하게 하기 위해 임의의 2개의 단백질 또는 이의 단편 사이에 위치할 수 있다.

[0038] **[0060]** 본원에 설명된 폴리뉴클레오티드 또는 폴리펩티드 서열의 맥락에서 사용되는 용어 "동일성" 또는 "실질적인 동일성"은 참조 서열에 대해 적어도 60% 서열 동일성을 갖는 서열을 지칭한다. 대안적으로, 동일성 백분율은 60% 내지 100%의 임의의 정수일 수 있다. 예시적인 구현예는 본원에 설명된 프로그램, 바람직하게는 하기 설명되는 바와 같은 표준 파라미터를 사용한 BLAST를 사용하여 참조 서열과 비교하여 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99%를 포함한다. 본원에 기재된 임의의 뉴클레오티드 또는 폴리펩티드 서열, 예를 들어, SEQ ID NO: 1-110 중 어느 하나에 대해 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열이 본원에 제공된 조성물 및 방법에 사용될 수 있음이 이해된다. 핵산 서열은 본원에 설명된 임의의 핵산 서열을 포함하거나, 이로 구성되거나, 이를 필수적 요소로 하여 구성될 수 있음이 이해된다. 유사하게, 폴리펩티드는 본원에 설명된 임의의 폴리펩티드 서열을 포함하거나, 이로 구성되거나, 이를 필수적 요소로 하여 구성될 수 있다. 서열 비교를 위해, 전형적으로 하나의 서열은 시험 서열이 비교되는 참조 서열로서 작용한다. 서열 비교 알고리즘을 사용할 때, 시험 및 참조 서열은 컴퓨터에 입력되고, 필요한 경우 하위서열 좌표가 지정되고, 서열 알고리즘 프로그램 파라미터가 지정된다. 디폴트 프로그램 파라미터가 사용될 수 있거나, 대체 파라미터가 지정될 수 있다. 이어서, 서열 비교 알고리즘은 프로그램 파라미터에 기초하여 참조 서열에 비한 시험 서열에 대한 서열 동일성 백분율을 계산한다.

[0039] **[0061]** 본원에서 사용되는 "비교 윈도우"는 20 내지 600, 약 20 내지 50, 약 20 내지 100, 약 50 내지 약 200 또는 약 100 내지 약 150으로 구성된 군으로부터 선택된 연속 위치의 수 중 어느 하나의 세그먼트에 대한 언급을 포함하며, 여기서 서열은 2개의 서열이 최적으로 정렬된 후 동일한 수의 연속 위치의 참조 서열과 비교될 수 있다. 비교를 위한 서열의 정렬 방법은 당 분야에 널리 공지되어 있다. 비교를 위한 서열의 최적 정렬은 문헌[Smith and Waterman Add. APL. Math. 2:482 (1981)]의 국소 상동성 알고리즘, 문헌[Needleman and Wunsch J. Mol. Biol. 48:443 (1970)의 상동성 정렬 알고리즘, 문헌[Pearson and Lipman Proc. Natl. Acad. Sci. (U.S.A.) 85: 2444 (1988)]의 유사성 방법을 위한 검색, 이들 알고리즘의 전산화된 구현(예를 들어, BLAST), 또는 수동 정렬 및 육안 검사에 의해 수행될 수 있다.

[0040] **[0062]** 서열 동일성 및 서열 유사성 백분율을 결정하기에 적합한 알고리즘은 BLAST 및 BLAST 2.0 알고리즘이고, 이는 문헌[Altschul et al. (1990) J. Mol. Biol. 215: 403-410 및 Altschul et al. (1977) Nucleic Acids Res. 25: 3389-3402]에 각각 설명되어 있다. BLAST 분석을 수행하기 위한 소프트웨어는 미국 국립생물공학정보센터(National Center for Biotechnology Information; NCBI) 웹사이트를 통해 공개적으로 이용 가능하다. 알고리즘은 먼저 질의 서열에서 길이 W의 짧은 단어를 식별함으로써 고 스코어링 서열 쌍(HSP)을 식별하는 것을 포함하며, 이는 데이터베이스 서열에서 동일한 길이의 단어와 정렬될 때 일부 양성-값 임계값 스코어 T와 일치하거나 이를 충족시킨다. T는 이웃 단어 스코어 임계값으로 지칭된다(Altschul et al, 상기 참조). 이러한 초기 이웃 단어 히트(hit)는 이를 함유하는 더 긴 HSP를 찾기 위한 검색을 개시하기 위한 시드로서 작용한다. 이후, 누적 정렬 스코어가 증가될 수 있는 한, 단어 히트는 각 서열을 따라 양방향으로 확장된다. 누적 스코어는 뉴클레오티드 서열에 대해 파라미터 M(일치하는 잔기의 쌍에 대한 보상 스코어; 항상 >0) 및 N(미스매칭 잔기에 대한 벌점 스코어; 항상 <0)을 사용하여 계산된다. 각 방향으로의 단어 히트의 확장은 다음의 경우에 중단된다: 누적 정렬 스코어가 이의 최대 달성 값으로부터 양 X만큼 떨어지거나; 하나 이상의 음성-스코어링 잔기 정렬의 축적으로 인해, 누적 스코어가 0 이하가 되거나; 어느 한 서열의 끝에 도달한다. BLAST 알고리즘 파라미터 W, T, 및 X는 정렬의 감도 및 속도를 결정한다. BLASTN 프로그램(뉴클레오티드 서열의 경우)은 디폴트로서 28의 단어 크기(W), 10의 기대치(E), M=1, N=-2, 및 둘 모두의 가닥의 비교를 사용한다.

[0041] **[0063]** BLAST 알고리즘은 또한 2개의 서열 사이의 유사성의 통계적 분석을 수행한다(예를 들어, 문헌[Karlin & Altschul, Proc. Nat'l. Acad. Sci. USA 90:5873-5787 (1993)] 참조). BLAST 알고리즘에 의해 제공되는 유사성의 한 척도는 최소 합 확률(P(N))이며, 이는 2개의 뉴클레오티드 또는 아미노산 서열 사이의 일치가 우연히 발생할 확률의 지표를 제공한다. 예를 들어, 시험 핵산과 참조 핵산의 비교에서 최소 합 확률이 약 0.01 미만, 보다 바람직하게는 약 10-5 미만, 및 가장 바람직하게는 약 10-20 미만인 경우 핵산은 참조 서열과 유사한 것으로

간주된다.

[0042] [0064] 전체에 걸쳐 사용되는 바와 같이, 대상체는 개인을 의미한다. 예를 들어, 대상체는 영장류와 같은 포유동물, 및 보다 구체적으로 인간이다. 비-인간 영장류도 대상체이다. 용어 대상체는 길들여진 동물, 예를 들어, 고양이, 개 등, 가축(예를 들어, 소, 말, 돼지, 양, 염소 등) 및 실험 동물(예를 들어, 흰 족제비, 친칠라, 마우스, 토끼, 래트, 저빌, 기니피그 등)을 포함한다. 따라서, 수의학적 용도 및 의학적 용도 및 제형이 본원에서 고려된다. 이 용어는 특정 연령 또는 성별을 나타내지 않는다. 따라서, 남성이든 여성이든 성인 및 신생아 대상체가 포함되도록 의도된다. 본원에서 사용되는 바와 같이, 환자 또는 대상체는 상호교환적으로 사용될 수 있고, 질병 또는 장애를 앓고 있는 대상체를 지칭할 수 있다.

[0043] [0065] "발현 카세트"는 숙주 세포에서 특정 폴리뉴클레오티드 서열의 전사를 허용하는 일련의 특정 핵산 요소를 갖는, 재조합적으로 또는 합성적으로 생성된 핵산 작제물이다. 발현 카세트는 플라스미드, 바이러스 유전체, 또는 핵산 단편의 일부일 수 있다. 전형적으로, 발현 카세트는 프로모터에 작동 가능하게 연결된 전사될 폴리뉴클레오티드, 이어서 전사 종료 신호 서열을 포함한다. 발현 카세트는 인간 글로빈 유전자로부터의 5' 또는 3' 비번역 영역과 같은 특정 조절 서열을 포함하거나 포함하지 않을 수 있다.

[0044] [0066] "리포터 유전자"는 효소 활성 또는 화학형광 특징과 같은 생화학적 특징으로 인해 쉽게 검출 가능한 단백질을 인코딩한다. 이러한 리포터 단백질은 선택 가능한 마커로서 사용될 수 있다. 이러한 리포터의 한 가지 특정 예는 녹색 형광 단백질이다. 이 단백질로부터 생성된 형광은 다양한 상업적으로 이용 가능한 형광 검출 시스템으로 검출될 수 있다. 다른 리포터는 염색에 의해 검출될 수 있다. 리포터는 또한 적절한 기질과 접촉될 때 검출 가능한 신호를 생성하는 효소일 수 있다. 리포터는 검출 가능한 생성물의 형성을 촉매하는 효소일 수 있다. 적합한 효소는 프로테아제, 뉴클레아제, 리파제, 포스파타제 및 하이드롤라제를 포함하나, 이에 제한되지 않는다. 리포터는 기질이 진핵 형질막에 대해 실질적으로 불침투성이어서 신호 형성을 엄격하게 제어하는 것을 가능하게 하는 효소를 인코딩할 수 있다. 효소를 인코딩하는 적합한 리포터 유전자의 특정 예는 CAT(클로람페니콜 아세틸 트랜스퍼라제; Alton and Vapnek (1979) Nature 282: 864-869); 루시페라제(lux); β-갈락토시다제; LacZ; β.-글루쿠로니다제; 및 알칼리성 포스파타제(Toh, et al. (1980) Eur. J. Biochem. 182: 231-238; 및 Hall et al. (1983) J. Mol. Appl. Gen. 2: 101)를 포함하나 이에 제한되지 않으며, 이들 각각은 그 전체가 본원에 참조로 포함된다. 다른 적합한 리포터는 에피토프를 특이적으로 인식하는 표지된 항체로 검출될 수 있는 특정 에피토프를 인코딩하는 것들을 포함한다.

[0045] [0067] "CRISPR/Cas" 시스템은 외래 핵산에 대한 방어를 위한 광범위한 부류의 박테리아 시스템을 지칭한다. CRISPR/Cas 시스템은 광범위한 진균 및 고세균 유기체에서 발견된다. CRISPR/Cas 시스템은 유형 I, II, 및 III 서브타입을 포함한다. 문헌[Makarova, et al. (*Nat Rev Microbiol.* 2015 Nov; 13(11):722-36)]에 설명된 바와 같은 CRISPR/Cas 시스템 분류는 공유된 특성 및 진화론적 유사성에 기초하여 5개의 유형 및 16개의 서브타입을 정의한다. 이들은 유전체 DNA를 절단하는 효과기 복합체의 구조에 기초하여 2개의 큰 부류로 그룹화된다. 유형 II CRISPR/Cas 시스템은 유전체 공학에 처음 사용되었으며, 2015년에 유형 V가 그 뒤를 이었다. 야생형 유형 II CRISPR/Cas 시스템은 가이드 RNA와 복합체에서 RNA-매개된 뉴클레아제 Cas 단백질 또는 동족체(본원에서 "CRISPR-관련 엔도뉴클레아제"로 지칭됨)를 이용하여 외래 핵산을 인식하고 절단한다. Cas9 단백질은 또한 활성화 RNA(트랜스활성화 또는 tracr RNA로도 지칭됨)를 사용한다. 함께 사용되는 CRISPR-관련 엔도뉴클레아제의 유형에 따라, 가이드 RNA 또는 가이드 RNA 및 활성화 RNA 둘 모두의 활성을 갖는 가이드 RNA는 또한 당 분야에 공지되어 있다. 일부 경우에, 이러한 이중 활성 가이드 RNA는 단일 가이드 RNA(sgRNA)로 지칭된다. 활성화 RNA 서열을 함유하지 않는 합성 가이드 RNA는 또한 sgRNA로 지칭될 수 있다. 본 발명의 개시에서, 용어 sgRNA 및 gRNA는 CRISPR-관련 엔도뉴클레아제와 복합체를 형성하고 CRISPR-관련 엔도뉴클레아제, 예를 들어, 리보핵단백질 복합체에서 CRISPR-관련 엔도뉴클레아제를 표적 DNA 서열에 국소화시키는 RNA 분자를 지칭하기 위해 상호교환적으로 사용된다.

[0046] [0068] 본원에서 사용되는 바와 같이, sgRNA 활성, 또는 RNP 활성, 즉, (1) gRNA 및 (2) CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 융합 단백질을 포함하는 복합체의 RNP 활성의 맥락에서 "활성"은 표적 유전자 요소에 결합하는 sgRNA의 능력을 지칭한다. 전형적으로, 활성은 또한 세포의 유전체를 편집하는 RNP (즉, CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 융합 단백질과 복합체화된 sgRNA)의 능력을 지칭한다. 본원에서 사용되는 바와 같이, 세포의 유전체의 편집과 관련하여 어구 "편집"은 표적 유전체 영역에서 유전체의 서열에서 구조적 변화를 유도하고, 예를 들어, 절단 부위에서 상동성 지정 복구(HDR)를 통해 세포의 유전체로 공여자 서열을 삽입하거나, 서열을 절단하고 비-상동성 말단 연결(NHEJ)을 통해 복구를 허용하는

것을 지칭한다.

[0047] **[0069]** 본원에서 사용되는 용어 "리보핵단백질 복합체", "RNP" 등은 (1) CRISPR-관련 엔도뉴클레아제(예를 들어, Cas9) 및 DNA 중합효소 도메인을 포함하는 융합 단백질, 및 crRNA(예를 들어, 가이드 RNA 또는 단일 가이드 RNA), (2) CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 융합 단백질; 및 트랜스-활성화 crRNA(tracrRNA), (3) CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 융합 단백질, 및 가이드 RNA, 또는 (4) 이들의 조합(예를 들어, CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 융합 단백질, tracrRNA, 및 crRNA 가이드를 함유하는 복합체) 사이의 복합체를 지칭한다.

[0048] **[0070]** 본원에서 사용되는 "세포"는 임의의 진핵 세포, 예를 들어, 인간 T 세포 또는 T 세포로 분화할 수 있는 세포, 예를 들어, TCR 수용체 분자를 발현하는 T 세포일 수 있다. 이들은 조혈 줄기 세포 및 조혈 줄기 세포로부터 유래된 세포를 포함한다. 세포의 집단, 예를 들어, 바이러스 입자를 포함하는 세포의 집단 또는 본원에 제공된 임의의 유전체 편집 방법에 의해 제조된 유전적으로 변형된 세포가 또한 제공된다.

[0049] **[0071]** 본원에서 사용되는 어구 "조혈 줄기 세포"는 혈액 세포를 발생시킬 수 있는 줄기 세포의 유형을 지칭한다. 조혈 줄기 세포는 골수 또는 림프 계통의 세포, 또는 이들의 조합을 발생시킬 수 있다. 조혈 줄기 세포는 말초 혈액 또는 이의 분획으로부터 분리될 수 있지만, 주로 골수에서 발견된다. 다양한 세포 표면 마커를 사용하여 조혈 줄기 세포를 확인, 분류 또는 정제할 수 있다. 일부 경우에, 조혈 줄기 세포는 c-kit+ 및 lin-으로 확인된다. 일부 경우에, 인간 조혈 줄기 세포는 CD34+, CD59+, Thy1/CD90+, CD38lo/-, C-kit/CD117+, lin-로서 확인된다. 일부 경우에, 인간 조혈 줄기 세포는 CD34-, CD59+, Thy1/CD90+, CD38lo/-, C-kit/CD117+, lin-로서 확인된다. 일부 경우에, 인간 조혈 줄기 세포는 CD133+, CD59+, Thy1/CD90+, CD38lo/-, C-kit/CD117+, lin-로서 확인된다. 일부 경우에, 마우스 조혈 줄기 세포는 CD34lo/-, SCA-1+, Thy1+/lo, CD38+, C-kit +, lin-으로 확인된다. 일부 경우에, 조혈 줄기 세포는 CD150+CD48-CD244-이다.

[0050] **[0072]** 본원에서 사용되는 어구 "조혈 세포"는 조혈 줄기 세포로부터 유래된 세포를 지칭한다. 조혈 세포는 유기체, 시스템, 기관, 또는 조직(예를 들어, 혈액, 또는 이의 분획)으로부터 분리에 의해 획득되거나 제공될 수 있다. 대안적으로, 조혈 줄기 세포는 분리될 수 있고, 조혈 세포는 줄기 세포를 분화시킴으로써 획득되거나 제공될 수 있다. 조혈 세포는 추가 세포 유형으로 분화할 가능성이 제한된 세포를 포함한다. 이러한 조혈 세포는 다능성 전구 세포, 계통-제한 전구 세포, 공통 골수 전구 세포, 과립구-대식세포 전구 세포, 또는 거대핵세포-적혈구 전구 세포를 포함하나, 이에 제한되지 않는다. 조혈 세포는 림프구, 적혈구, 과립구, 단핵구, 및 혈소판과 같은 림프구 및 골수 계통의 세포를 포함한다. 일부 구현예에서, 조혈 세포는 면역 세포, 예를 들어, T 세포, B 세포, 대식세포, 자연 살해(NK) 세포 또는 수지상 세포이다. 일부 구현예에서, 세포는 선천 면역 세포이다.

[0051] **[0073]** 본원에서 사용되는 어구 "T 세포"는 T 세포 수용체 분자를 발현하는 림프 세포를 지칭한다. T 세포는 인간 알파 베타(αβ) T 세포 및 인간 감마 델타(γδ) T 세포를 포함한다. T 세포는 나이브 T 세포, 자극된 T 세포, 일차 T 세포(예를 들어, 배양되지 않은 일차 T 세포), 배양된 T 세포, 무한증식 T 세포, 헬퍼 T 세포, 세포독성 T 세포, 기억 T 세포, 조절 T 세포, 자연 살해 T 세포, 이들의 조합, 또는 이들의 하위집단을 포함하나 이에 제한되지 않는다. T 세포는 CD4+, CD8+, 또는 CD4+ 및 CD8+일 수 있다. T 세포는 또한 CD4-, CD8-, 또는 CD4- 및 CD8-일 수 있다. T 세포는 헬퍼 세포, 예를 들어, 유형 TH1, TH2, TH3, TH9, TH17, 또는 TFH의 헬퍼 세포일 수 있다. T 세포는 세포독성 T 세포일 수 있다. 조절 T 세포는 FOXP3+ 또는 FOXP3-일 수 있다. T 세포는 알파/베타 T 세포 또는 감마/델타 T 세포일 수 있다. 일부 경우에, T 세포는 CD4+CD25hiCD127lo 조절 T 세포이다. 일부 경우에, T 세포는 유형 1 조절(Tr1), TH3, CD8+CD28-, Treg17, 및 Qa-1 제한된 T 세포, 또는 이들의 조합 또는 하위-집단으로 구성된 군으로부터 선택된 조절 T 세포이다. 일부 경우에, T 세포는 FOXP3+ T 세포이다. 일부 경우에, T 세포는 CD4+CD25loCD127hi 효과기 T 세포이다. 일부 경우에, T 세포는 CD4+CD25loCD127hiCD45RAhiCD45RO- 나이브 T 세포이다. T 세포는 유전적으로 조작된 재조합 T 세포일 수 있다.

[0052] **[0074]** 본원에서 사용되는 바와 같이, 일차 세포와 관련하여 어구 "일차"는 형질전환되거나 무한증식화되지 않은 세포이다. 이러한 일차 세포는 제한된 횟수로 배양, 계대배양, 또는 계대될 수 있다(예를 들어, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20회 배양됨). 일부 경우에, 일차 세포는 시험관내 배양 조건에 적응된다. 일부 경우에, 일차 세포는 유기체, 시스템, 기관 또는 조직으로부터 분리되고, 선택적으로 분류되고, 배양 또는 계대배양 없이 직접 이용된다. 일부 경우에, 일차 세포는 자극되거나, 활성화되거나, 분화된다. 예를 들어, 일차 T 세포는 CD3, CD28 효능제, IL-2, IFN-γ, 또는 이들의 조합과의 접촉(예를 들어, 이들의 존재하에서 배양)에 의해 활성화될 수 있다.

[0053] **상세한 설명**

[0054] **[0075]** 하기 설명은 본 발명의 조성물 및 방법의 다양한 양태 및 구현예를 언급한다. 특정 구현예는 조성물 및 방법의 범위를 정의하는 것으로 의도되지 않는다. 오히려, 구현예는 단지 개시된 조성물 및 방법의 범위 내에 적어도 포함되는 다양한 조성물 및 방법의 비제한적인 예를 제공한다. 설명은 당업자의 관점에서 읽혀져야 하고; 따라서, 당업자에게 널리 공지된 정보가 반드시 포함되는 것은 아니다.

[0055] **[0076]** CRISPR/Cas9(일정한 간격을 두고 주기적으로 분포하는 짧은 회문 반복서열 및 CRISPR-관련 단백질 9) 시스템은 단일 가이드 RNA(sgRNA)에 의해 가이드되는 DNA 이중-가닥 파손(DSB)을 만들기 위해 단일 효과기 단백질을 사용하고, 인간 세포에서 특이적 유전적 변화를 제조하는 데 사용되어 왔다. 스트렙토코커스 피오게네스(*Streptococcus pyogenes*) Cas9(SpCas9)는 주로 "NGG" 프로토스페이서 인접 모티프(PAM)의 상류에서 둘 모두의 DNA 가닥 3개의 뉴클레오티드(nt)를 절단함으로써 평활 말단을 생성한다. 이는 또한 PAM의 상류에서 표적화 가닥 3 nt, 및 비-표적화 가닥 4, 5, 또는 6 nt를 절단함으로써 1, 2 또는 3 nt 5' 오버행을 갖는 엇갈린 말단을 생성할 수 있다.

[0056] **[0077]** 인간 세포는 CRISPR/Cas9 뉴클레아제에 의해 생성된 이중 가닥 파손(DSB)을 복구하기 위한 다수의 경로를 갖는다: 상동성-지정 복구(HDR), 표준 비-상동성 말단 연결(cNHEJ 또는 NHEJ), 및 미세상동성-매개 말단 연결(MMEJ) 및 단일 가닥 어닐링(SSA)을 포함하는 대안적인 말단 연결 경로. cNHEJ 경로에서, DSB의 2개의 말단은 절제 또는 주형의 개입 없이 재-라이게이션을 위해 처리된다. Pol X 패밀리 구성원(Pol λ, μ, 및 β 및 말단 트랜스퍼라제)은 5' 오버행을 채우는 기능을 할 수 있다. 이러한 채움 반응은 SpCas9에 의해 유도된 널리 관찰된 예측 가능한 삽입 또는 주형 삽입(TIS)을 설명할 수 있다.

[0057] **[0078]** HDR, MMEJ 및 SSA 복구 경로는 모두 3' 오버행을 생성하기 위해 DNA 절제에 의존한다. DNA 절제는 MRE11-RAD50-NBS1 복합체에 의해 개시되고 CtIP(*RBBP8*에 의해 인코딩됨)에 의해 자극된다. MRE11의 엔도뉴클레아제 활성은 DSB에 대해 3' 닉을 생성하고, 3'-5' 엑소뉴클레아제 활성은 닉으로부터 3'-단일 가닥 DNA 오버행을 생성한다. HDR은 DSB를 충실하게 복구하기 위해 긴 3' 오버행 및 DNA 주형을 사용한다. MMEJ 및 SSA는 2개의 3' 오버행에서 각각 미세상동성(2-5 nt) 및 짧은 상동성(10-15 nt)을 사용하여 상동성 영역 사이의 거리에 따른 크기를 갖는 DNA 결실을 생성하는 DNA 합성을 용이하게 한다. 과도한 DNA 절제로부터의 긴 3' 오버행은 또한 DNA pol α 및 관련 복합체에 의해 채워질 수 있다. 이러한 채움(fill-in) 반응은 Cas9 닉카제에 의해 생성된 3' 오버행을 갖는 DSB에서 관찰된 작은 탠덤 중복을 설명한다.

[0058] **[0079]** CRISPR/Cas9-유도된 돌연변이 프로파일은 어느 정도 예측 가능하고, 작은 삽입 및 결실(INDEL)이 우세한 돌연변이 유형이지만, 예측할 수 없는 표적-내 큰 결실이 널리 보고된다. 현재, 큰 결실의 생성을 억제하는 방법은 부족하다. 염기 편집 및 프라임 편집과 같은 DSB를 생성하지 않는 유전체 편집 방법이 개발되었지만, 인간 세포로부터 통합된 프로바이러스 DNA를 근절하고 질환-유발 DNA 반복부를 제거하는 것을 포함하는 많은 적용은 여전히 DSB 생성을 포함한다. 따라서, 정제된 돌연변이 프로파일을 생성할 수 있는 방법은 안전성을 개선시킬 것이다.

[0059] **[0080]** 유전체 편집을 위한 조성물, 시스템, 제조 방법, 및 방법이 본원에 제공된다. 본원에 설명된 조성물 및 방법을 사용하여, 세포의 유전체는 CRISPR-관련 엔도뉴클레아제를 사용하여 효율적으로 편집될 수 있는 한편, 예를 들어, 세포에서 비-NHEJ(예를 들어, 미세상동성-매개 말단 연결(MMEJ) 또는 단일 가닥 어닐링(SSA))와 비교하여 세포에서 비-상동성 말단 연결(NHEJ)을 증가시킴으로써 세포의 유전체에서 생성된 긴 결실의 수를 감소시킬 수 있다. (1) CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 폴리펩티드를 포함하는 융합 단백질; 및 (2) sgRNA를 포함하는 RNP를 사용하여 세포의 유전체를 편집하기 위한 예시적인 성분, 시스템, 제조 방법, 및 방법이 본원에 제공된다.

[0060] **포유동물 발현 플라스미드**

[0061] **[0081]** CRISPR 성분 코딩 서열, 즉, sgRNA 및 CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 융합 단백질을 본 발명의 개시의 렌티바이러스-유사 입자를 생성하는데 사용되는 포유동물 세포에 전달하는데 사용되는 포유동물 발현 플라스미드가 본원에 제공된다. 예를 들어, 비-바이러스 핵산 서열에 작동 가능하게 연결된 진핵 프로모터를 포함하는 포유동물 발현 플라스미드가 본원에 제공되며, 여기서 비-바이러스 핵산 서열은 (i) 융합 단백질을 인코딩하는 핵산 서열로서, 상기 융합 단백질이 (a) DNA 중합효소 도메인을 포함하는 폴리펩티드; 및 (b) CRISPR-관련 엔도뉴클레아제 코딩 서열을 포함하는, 핵산 서열; 및 (ii) 가이드 RNA(gRNA) 코딩 서열로서, 상기 gRNA 코딩 서열이 적어도 하나의 앱타머 코딩 서열을 포함하는, 가이드 RNA(gRNA) 코딩 서열을

포함한다.

[0062] **[0082]** 본원에 제공된 포유동물 발현 플라스미드에서, DNA 중합효소 도메인을 포함하는 폴리펩티드는 CRISPR-관련 엔도뉴클레아제에 융합되거나 연결될 수 있다. 선택적으로, CRISPR-관련 엔도뉴클레아제는 펩티드 링커를 통해 DNA 중합효소 도메인을 포함하는 폴리펩티드에 연결된다. 링커는 약 2 내지 약 25개의 아미노산 길이일 수 있다.

[0063] **[0083]** 일부 구현예에서, CRISPR-관련 엔도뉴클레아제 코딩 서열은 Cas9 단백질을 인코딩한다. 전장 Cas9는 DNA 서열에서 이중-가닥 파손을 생성하는 인식 도메인 및 2개의 뉴클레아제 도메인(각각 HNH 및 RuvC)을 포함하는 엔도뉴클레아제이다. Cas9은 Cas9에 의해 인식되는 NGG 모티프 바로 앞에 있는 20-뉴클레오티드 DNA 서열에 하이브리드화하는 가이드 RNA와 상호작용함으로써 세포의 유전체 부위를 표적화한다. 이는 세포의 유전체 DNA에서 이중-가닥 파손을 초래한다. 일부 예에서, 가이드 RNA에 의해 표적화된 영역의 바로 3'에 NGG 프로토스페이서 인접 모티프(PAM)를 필요로 하는 Cas9 뉴클레아제가 이용될 수 있다. 또 다른 예로서, 오소고날 PAM 모티프 요건을 갖는 Cas9 단백질은 인접한 NGG PAM 서열을 갖지 않는 서열을 표적화하는데 이용될 수 있다. 오소고날 PAM 서열 특이성을 갖는 예시적인 Cas9 단백질은 문헌[Esvelt *et al.*, Nature Methods 10: 1116-1121 (2013)]에 설명된 것들을 포함하지만 이에 제한되지 않는다.

[0064] **[0084]** 다양한 Cas9 뉴클레아제가 본원에 설명된 방법에서 이용될 수 있다. 예를 들어, SpCas9와 같은 가이드 RNA에 의해 표적화된 영역의 바로 3'에 NGG 프로토스페이서 인접 모티프(PAM)를 필요로 하는 Cas9 뉴클레아제가 이용될 수 있다. 이러한 Cas9 뉴클레아제는 NGG 서열을 함유하는 유전체의 임의의 영역으로 표적화될 수 있다. 또 다른 예에서, SaCas9와 같은 가이드 RNA에 의해 표적화된 영역의 바로 3'에 NNGRRT 또는 NNGRR(N) PAM을 필요로 하는 Cas9 뉴클레아제가 이용될 수 있다. 또 다른 예로서, 오소고날 PAM 모티프 요건을 갖는 Cas9 단백질은 인접한 NGG PAM 서열을 갖지 않는 서열을 표적화하는데 이용될 수 있다. 오소고날 PAM 서열 특이성을 갖는 예시적인 Cas9 단백질은 문헌[Nature Methods 10, 1116-1121 (2013)]에 설명된 것들, 및 문헌[Zetsche *et al.*, Cell, Volume 163, Issue 3, p759-771, 22 October 2015]에 설명된 것들을 포함하나 이에 제한되지 않는다. Cas9 단백질에 대한 예시적인 아미노산 서열은 본원에서 SEQ ID NO: 29로서 제시된다.

[0065] **[0085]** 일부 경우에, Cas9 단백질은 가이드 RNA와의 복합체의 일부로서 표적 핵산에 결합될 때, 단일 가닥 파손 또는 닉이 표적 핵산에 도입되도록 하는 닉카제이다. 각각 구조적으로 상이한 가이드 RNA에 결합된 한 쌍의 Cas9 닉카제는 표적 유전체 영역의 2개의 근위 부위로 표적화될 수 있고, 따라서 한 쌍의 근위 단일 가닥 파손을 표적 유전체 영역에 도입할 수 있다. 닉카제 쌍은 표적-외 효과가 단일 닉을 야기할 가능성이 있기 때문에 향상된 특이성을 제공할 수 있으며, 이는 일반적으로 염기-절제 복구 메커니즘에 의해 병변 없이 수복된다. 예시적인 Cas9 닉카제는 D10A 또는 H840A 돌연변이를 갖는 Cas9 뉴클레아제를 포함한다.

[0066] **[0086]** 일부 구현예에서, CRISPR-관련 엔도뉴클레아제는 Cpf1 폴리펩티드이다. Cpf1 단백질은 클래스 II, 유형 V CRISPR/Cas 시스템 단백질이다. Cpf1은 Cas9(예를 들어, spCas9)보다 작고 단순한 엔도뉴클레아제이다. Cpf1 단백질은 Cas9의 RuvC 도메인과 유사하지만 HNH 엔도뉴클레아제 도메인을 갖지 않는 RuvC-유사 엔도뉴클레아제 도메인을 갖는다. Cpf1의 N-말단 도메인은 또한 Cas9 단백질과 같은 알파-나선 인식 로브를 갖지 않는다. DNA를 절단할 때, Cpf1은 4 또는 5개의 뉴클레오티드 오버행을 갖는 점착성-말단-유사 DNA 이중-가닥 파손을 도입한다. Cpf1 단백질은 tracrRNA를 필요로 하지 않으며; 오히려, Cpf1 단백질은 crRNA만으로 기능한다. 본 발명의 개시의 맥락에서, CRISPR-관련 엔도뉴클레아제가 Cpf1 단백질인 경우, sgRNA는 tracr 서열을 포함하지 않는다. Cpf1 단백질과 함께 사용되는 sgRNA는 crRNA 서열(불변 영역)만을 포함할 수 있다. 일부 예에서, 가이드 RNA에 의해 표적화되는 영역의 바로 5'에 TTTN 또는 TTN PAM(종에 따라, "N"은 핵염기임)을 필요로 하는 Cpf1 단백질이 이용될 수 있다. 공지된 Cpf1 단백질 및 이의 유도체가 본 발명의 개시의 맥락에서 사용될 수 있다. 예를 들어, 일부 경우에, CRISPR-관련 엔도뉴클레아제는 FnCpf1p이고 PAM은 5' TTN이고, 여기서 N은 A/C/G 또는 T이다. 일부 예에서, CRISPR-관련 엔도뉴클레아제는 PaCpf1p이고 PAM은 5' TTTV이고, 여기서 V는 A/C 또는 G이다. 특정 예에서, CRISPR-관련 엔도뉴클레아제는 FnCpf1p이고 PAM은 5' TTN이고, 여기서 N은 A/C/G 또는 T이고, PAM은 프로토스페이서의 5' 말단의 상류에 위치한다. 특정 예에서, CRISPR-관련 엔도뉴클레아제는 FnCpf1p이고 PAM은 5' CTA이고 프로토스페이서 또는 표적 유전자좌의 5' 말단의 상류에 위치한다. 일 예에서, CRISPR-관련 엔도뉴클레아제는 AsCpf1p이고 PAM은 5' TTTN이다. Cpf1에 대한 예시적인 아미노산 서열은 본원에서 SEQ ID NO: 30으로 제시된다.

[0067] **[0087]** 전체적으로 사용되는 바와 같이, DNA 중합효소 도메인은 핵산 주형의 존재 하에 뉴클레오티드의 듀플렉스 DNA로의 5'-3' 중합을 촉매화하는 전장 DNA 중합효소의 단편이다. DNA 중합효소 도메인을 포함하는 임의의

DNA 중합효소 도메인 또는 폴리펩티드는 본원에 설명된 융합 단백질을 제조하는데 사용될 수 있다. DNA 중합효소 도메인을 CRISPR-관련 엔도뉴클레아제에 융합시킴으로써, DNA 중합효소 도메인은 세포의 유전체에서 CRISPR-관련 엔도뉴클레아제에 의해 만들어진 이중-가닥 파손 부위로 표적화된다. 일부 구현예에서, DNA 중합효소 도메인은 촉매적으로 활성이다. 일부 구현예에서, DNA 중합효소 도메인의 촉매 활성은 (예를 들어, 상응하는 야생형 DNA 중합효소 도메인에 비해), 예를 들어, 적어도 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 80%, 90% 또는 100% 감소된다. 일부 구현예에서, DNA 중합효소 도메인은 촉매 활성을 없애는 하나 이상의 돌연변이, 예를 들어, D705A 돌연변이를 포함한다.

[0068] **[0088]** 일부 경우에, DNA 중합효소 도메인을 포함하는 폴리펩티드는 중합효소 활성 및 3'-5' 엑소뉴클레아제 활성을 갖는다. 일부 경우에, DNA 중합효소 도메인을 포함하는 폴리펩티드는 3'-5' 엑소뉴클레아제 활성을 갖지만, 중합효소 활성이 (예를 들어, 상응하는 야생형 DNA 중합효소 도메인에 비해) 감소하거나 결여되어 있다.

[0069] **[0089]** DNA 중합효소 도메인을 포함하는 예시적인 폴리펩티드는 DNA 중합효소 I(DNA Pol I)이다. DNA Pol I은 5'-3' DNA 의존성 DNA 중합효소 활성, 3'-5' 엑소뉴클레아제 활성, 5'-3' 엑소뉴클레아제 활성 및 5'-3' RNA-의존성 DNA 중합효소 활성을 갖는 효소이다. DNA Pol I에 대한 예시적인 서열은 본원에서 SEQ ID NO: 27로서 제공된다. 일부 구현예에서, DNA 중합효소 도메인을 포함하는 DNA pol I 단편은 CRISPR-관련 엔도뉴클레아제에 융합된다. 일부 경우에, 단편은 5'-3'DNA 중합효소 활성을 갖는 아미노산 서열 및 3'-5' 엑소뉴클레아제 활성을 갖는 아미노산 서열(예를 들어, DNA Pol I의 클레노우 단편)을 포함한다. DNA Pol I의 클레노우 단편에 대한 예시적인 서열은 본원에서 SEQ ID NO: 28로서 제공된다. 다른 예시적인 DNA 중합효소는 써무스 아쿠아티쿠스(*Thermus Aquaticus*) DNA Pol I 및 바실러스 스테아로써모필루스(*Bacillus stearothermophilus*) DNA pol I을 포함하나 이에 제한되지 않는다.

[0070] **[0090]** 본원에서 제공되는 포유동물 발현 플라스미드는 CRISPR 성분 코딩 서열, 예를 들어, CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 폴리펩티드를 포함하는 융합 단백질에 대한 코딩 서열; 및 gRNA를 포함한다. 일부 예에서, gRNA 코딩 서열은 적어도 하나의 앱타머 코딩 서열을 포함한다. 일부 예에서, 적어도 하나의 앱타머 코딩 서열은 gRNA의 5' 말단 또는 3' 말단에 위치할 수 있다. 일부 예에서, 적어도 하나의 앱타머 코딩 서열은 gRNA 내의 내부 위치, 예를 들어, 폴딩된 gRNA에 형성된 루프 중 하나 이상에 삽입될 수 있다. 예를 들어, gRNA가 Cas9 단백질용인 경우, 적어도 하나의 앱타머 코딩 서열은 gRNA의 테트라 루프, 스템 루프 2(ST2), 또는 3' 말단에 위치할 수 있다. 일부 예에서, 1-30개 뉴클레오티드의 스페이서는 적어도 하나의 앱타머 코딩 서열의 gRNA 사이에, 또는 적어도 하나의 앱타머 코딩 서열의 측면에 위치할 수 있다.

[0071] **[0091]** 일부 예에서, 포유동물 발현 벡터는 앱타머-결합 단백질(ABP)에 의해 특이적으로 결합되는 앱타머 서열을 인코딩하는 적어도 하나의 앱타머 코딩 서열을 포함한다. 본 발명의 개시의 맥락에서, 앱타머 서열은 ABP에 의해 특이적으로 결합되는 3차 루프 구조를 형성하는 RNA 서열이다. ABP는 RNA-결합 단백질 또는 RNA-결합 단백질 도메인이다. 적합한 앱타머 코딩 서열은 공지된 박테리오파지 앱타머 서열을 인코딩하는 폴리뉴클레오티드 서열을 포함한다. 예시적인 앱타머 코딩 서열은 표 1에서 상기 제공된 앱타머 서열을 인코딩하는 것들을 포함한다. 일부 예에서, 앱타머는 ABP의 이량체에 의해 결합된다. 이러한 앱타머 서열은 박테리오파지 단백질에 의해 특이적으로 결합되는 것으로 공지된 RNA 서열이다. 일부 상황에서, 적어도 하나의 앱타머 코딩 서열은 MS2 코트 단백질, PP7 코트 단백질, 람다 N RNA-결합 도메인, 또는 Com 단백질로 구성된 군으로부터 선택된 ABP에 의해 특이적으로 결합된 앱타머 서열을 인코딩한다.

[0072] **표 1.** 앱타머-결합 단백질 및 상응하는 앱타머 서열

	앱타머-결합 단백질			
	MS2 코트 단백질	PP7 코트 단백질	람다 N 펩티드(아미노산 1-22)	Com 단백질
핵산 서열	SEQ ID NO: 1	SEQ ID NO :3	SEQ ID NO:5	SEQ ID NO:7
아미노산 서열	SEQ ID NO: 2	SEQ ID NO: 4	SEQ ID NO:6	SEQ ID NO:8
앱타머(RNA)	SEQ ID NO:9	SEQ ID NO:11	SEQ ID NO:13 (박스-B 앱타머)	SEQ ID NO:15
앱타머(DNA)	SEQ ID NO:10	SEQ ID NO:12	SEQ ID NO:14	SEQ ID NO:16

[0073]

[0074] **[0092]** 일부 예에서, 포유동물 발현 벡터는 이의 하류에 하나의 앱타머 코딩 서열을 포함하는 sgRNA를

포함한다. 다른 예에서, gRNA는 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 또는 12개의 앱타머 코딩 서열을 포함할 수 있다. 예를 들어, 일부 경우에, gRNA는 2개의 앱타머 코딩 서열을 나란히 포함할 수 있다.

[0075] **[0093]** 전체적으로 사용되는 바와 같이, sgRNA는 sgRNA 및 CRISPR-관련 엔도뉴클레아제가 세포의 유전체에서 표적 핵산에 공동-국소화되도록 하는, CRISPR-관련 엔도뉴클레아제(CRISPR 부위-특이적 뉴클레아제)와 상호작용하고 세포의 유전체 내의 표적 핵산(유전체 표적 서열)에 특이적으로 결합하거나 이에 하이브리드화하는 단일 가이드 RNA 서열이다. 각각의 sgRNA는 유전체에서 표적 DNA 서열에 특이적으로 결합하거나 하이브리드화되는 약 10 내지 50개 뉴클레오티드 길이의 DNA 표적화 서열 또는 프로토스페이서 서열을 포함한다. 예를 들어, DNA 표적화 서열은 약 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 또는 50개의 뉴클레오티드 길이일 수 있다. 예를 들어, DNA 표적화 서열은 약 15-30개의 뉴클레오티드, 약 15-25개의 뉴클레오티드, 약 10-25개의 뉴클레오티드, 또는 약 18-23개의 뉴클레오티드일 수 있다. 일 예에서, DNA 표적화 서열은 약 20개의 뉴클레오티드이다. 일부 구현예에서, sgRNA는 crRNA 서열 및 트랜스활성화 crRNA(tracrRNA) 서열을 포함한다. 일부 구현예에서, sgRNA는 tracrRNA 서열을 포함하지 않는다.

[0076] **[0094]** 일반적으로, DNA 표적화 서열은 표적 DNA 서열을 상보시키거나(예를 들어, 완전히 상보적) 또는 실질적으로 상보하도록(예를 들어, 1-4개의 미스매치를 가짐) 설계된다. 일부 경우에, DNA 표적화 서열은 다수의 유전적 요소에 결합하기 위해 워블(wobble) 또는 축퇴(degenerate) 염기를 혼입할 수 있다. 일부 경우에, 결합 영역의 3' 또는 5' 말단에 있는 19개의 뉴클레오티드는 표적 유전적 요소 또는 요소들에 완전히 상보적이다. 일부 경우에, 결합 영역은 안정성을 증가시키기 위해 변경될 수 있다. 예를 들어, 비-천연 뉴클레오티드는 분해에 대한 RNA 내성을 증가시키기 위해 혼입될 수 있다. 일부 경우에, 결합 영역은 결합 영역에서 이차 구조 형성을 피하거나 감소시키도록 변경되거나 설계될 수 있다. 일부 경우에, 결합 영역은 G-C 함량을 최적화하도록 설계될 수 있다. 일부 경우에, G-C 함량은 바람직하게는 약 40% 내지 약 60%(예를 들어, 40%, 45%, 50%, 55%, 60%)이다. 일부 경우에, 결합 영역은 sgRNA의 효율적인 전사를 촉진하는 서열로 시작하도록 선택될 수 있다. 예를 들어, 결합 영역은 G 뉴클레오티드로 5' 말단에서 시작할 수 있다. 일부 경우에, 결합 영역은 변형된 뉴클레오티드, 예를 들어, 비제한적으로 메틸화된 또는 인산화된 뉴클레오티드를 함유할 수 있다.

[0077] **[0095]** 본원에서 사용되는 용어 "상보적인" 또는 "상보성"은 뉴클레오티드 또는 핵산 사이의 염기쌍 형성, 예를 들어, 비제한적으로 sgRNA와 표적 서열 사이의 염기쌍 형성을 지칭한다. 상보적 뉴클레오티드는 일반적으로 A 및 T(또는 A 및 U), 및 G 및 C이다. 본원에 설명된 가이드 RNA는 서열, 예를 들어, 유전체 서열에 완전히 상보적이거나 실질적으로 상보적인(예를 들어, 1-4개의 미스매치를 갖는) DNA 표적화 서열을 포함할 수 있다.

[0078] **[0096]** sgRNA는 CRISPR-관련 엔도뉴클레아제와 상호작용하거나 이에 결합하는 sgRNA 불변 영역을 포함한다. 본원에 제공된 작제물에서, sgRNA의 불변 영역은 약 75 내지 250개의 뉴클레오티드 길이일 수 있다. 일부 예에서, 불변 영역은 스템, 스템 루프, 헤어핀, 헤어핀 사이의 영역, 및/또는 불변 영역의 넥서스 내의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10개 이상의 뉴클레오티드 치환을 포함하는 변형된 불변 영역이다. 일부 예에서, 변형된 불변 영역이 유래되는 천연 또는 야생형 sgRNA 불변 영역의 활성과 비교하여 적어도 80%, 85%, 90%, 또는 95% 활성을 갖는 변형된 불변 영역이 본원에 설명된 작제물에 사용될 수 있다. 특히, CRISPR-관련 엔도뉴클레아제와 직접 상호작용하는 뉴클레오티드 또는 불변 영역의 이차 구조에 중요한 뉴클레오티드에서 변형이 이루어지지 않아야 한다.

[0079] **[0097]** 포유동물 발현 플라스미드는 비-바이러스 핵산 서열에 작동 가능하게 연결된 진핵 프로모터를 포함한다. 일부 예에서, RNA 중합효소 II 프로모터는 CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 폴리펩티드를 포함하는 융합 단백질을 인코딩하는 핵산에 작동 가능하게 연결되고; RNA 중합효소 III 프로모터는 gRNA 코딩 서열에 작동 가능하게 연결된다.

[0080] **[0098]** RNA 중합효소 II 프로모터 서열은 포유동물 종으로부터 선택된다. RNA 중합효소 III 프로모터 서열은 포유동물 종으로부터 선택된다. 예를 들어, 이러한 프로모터 서열은 몇 가지 예를 들어, 인간, 소, 양, 버팔로, 돼지, 또는 마우스로부터 선택될 수 있다. 일부 예에서, RNA 중합효소 II 프로모터 서열은 CMV, FE1α, 또는 SV40 서열이다. 일부 예에서, RNA 중합효소 III 프로모터 서열은 U6 또는 H1 서열이다. 일부 예에서, RNA 중합효소 II 서열은 변형된 RNA 중합효소 II 서열이다. 예를 들어, 임의의 포유동물 종으로부터의 야생형 RNA 중합효소 II 프로모터 서열에 대해 적어도 80%, 85%, 90%, 95%, 또는 99% 동일성을 갖는 RNA 중합효소 II 서열이 본원에 제공된 작제물에 사용될 수 있다. 일부 예에서, RNA 중합효소 III 서열은 변형된 RNA 중합효소 III 서열이다. 예를 들어, 임의의 포유동물 종으로부터의 야생형 RNA 중합효소 III 프로모터 서열에 대해 적어도 80%,

85%, 90%, 95%, 또는 99% 동일성을 갖는 RNA 중합효소 III 서열이 본원에 제공된 작제물에 사용될 수 있다. 당업자는 2개의 폴리펩티드 또는 핵산의 동일성을 결정하는 방법을 용이하게 이해한다. 예를 들어, 동일성은 동일성이 가장 높은 수준이 되도록 2개의 서열을 정렬한 후에 계산될 수 있다. 동일성을 계산하는 또 다른 방법은 공개된 알고리즘에 의해 수행될 수 있다. 예를 들어, 비교를 위한 서열의 최적 정렬은 문헌[Needleman and Wunsch, J. Mol. Biol. 48(3): 443-453 (1970)]의 알고리즘을 사용하여 수행될 수 있다. 일부 예에서, 진핵 프로모터는 유도성 또는 조절 가능한 프로모터이다.

[0081] **[0099]** RNA pol II 프로모터로부터 전사된 코딩 서열은 코딩 서열의 하류의 폴리(A) 신호 및 전사 종료자 서열을 포함한다. 통상적으로 사용되는 포유동물 종료자(SV40, hGH, BGH, 및 rbGlob)는 폴리아데닐화 및 종료 둘 모두를 촉진하는 서열 모티프 AAUAAA(SEQ ID NO: 31)를 포함한다. RNA pol III 프로모터로부터 전사된 코딩 서열은 종료자 서열로서 코딩 서열의 하류에 T 잔기의 단순 실행을 포함한다. 서열-기반 요소인 종료자의 역할은 전사 단위(예를 들어, 유전자)의 말단을 규정하고 전사 기구로부터 새로 합성된 RNA를 방출하는 과정을 개시하는 것이다. 종료자는 전사될 유전자의 하류에서 발견되며, 전형적으로 폴리아데닐화 또는 폴리(A) 신호와 같은 임의의 3' 조절 요소 직후에 발생한다.

[0082] **[0100]** 일부 예에서, 포유동물 발현 플라스미드는 또한 CRISPR 성분 코딩 서열의 하류에 위치한 RNA-안정화 서열 또는 CRISPR 성분 코딩 서열의 하류에 위치하는 경우 앱타머 코딩 서열을 인코딩하는 적어도 하나의 폴리뉴클레오티드 서열을 포함할 수 있다. RNA-안정화 서열을 인코딩하는 폴리뉴클레오티드 서열은 CRISPR/Cas 시스템 성분 코딩 서열의 하류에서 전사되고, 전사된 RNA 서열의 수명을 안정화시킨다. 일 예에서, RNA-안정화 서열을 인코딩하는 폴리뉴클레오티드 서열은 촉매적으로 손상된 CRISPR-관련 엔도뉴클레아제 코딩 서열의 하류에 위치한다. 또 다른 예에서, RNA-안정화 서열을 인코딩하는 폴리뉴클레오티드 서열은 gRNA 코딩 서열의 하류에 위치한다. 예시적인 RNA-안정화 서열은 SEQ ID NO: 20(DNA) 및 SEQ ID NO: 21(RNA)에 제시된 인간 베타 글로빈 유전자의 3' UTR의 서열이다. RNA-안정화 서열의 또 다른 예는 SEQ ID NO: 20의 2개 복사체를 포함하는 SEQ ID NO: 22이다. 다른 RNA-안정화 서열은 문헌[Hayashi, T. et al., Developmental Dynamics 239(7):2034-2040 (2010) 및 Newbury, S. et al., Cell 48(2):297-310 (1987)]에 설명되어 있다. 일부 예에서, 1-30개의 뉴클레오티드의 스페이서는 CRISPR 성분 코딩 서열과 RNA-안정화 서열을 인코딩하는 적어도 하나의 폴리뉴클레오티드 서열 사이에 위치할 수 있다.

[0083] **[0101]** 일부 예에서, 포유동물 발현 플라스미드는 하나 이상의 발현 카세트를 포함할 수 있다. 일부 예에서, 포유동물 발현 플라스미드는 본원에 설명된 임의의 융합 단백질을 인코딩하는 제1 발현 카세트 및 적어도 하나의 앱타머를 포함하는 gRNA를 인코딩하는 제2 발현 카세트를 포함한다. 일부 예에서, 포유동물 발현 플라스미드는 또한 리포터 유전자를 포함할 수 있다.

[0084] **시스템**

[0085] **[0102]** 본 발명의 개시의 또 다른 양태는 렌티바이러스 패키징 시스템이다. 이러한 시스템은 본 발명의 개시에 설명된 포유동물 발현 플라스미드를 포함한다. 이러한 시스템은 포유동물 세포로의 도입을 위한 성분을 제공하여 본 발명의 개시에 설명된 렌티바이러스-유사 입자를 생성하는데 유용하다.

[0086] **[0103]** 일부 예에서, 시스템은 바이러스 서열, 예를 들어, Gag 뉴클레오티드 서열에 작동 가능하게 연결된 진핵 프로모터를 포함하는 렌티바이러스 패키징 플라스미드를 포함하며, 여기서 Gag 뉴클레오티드 서열은 뉴클레오캡시드(NC) 코딩 서열 및 매트릭스 단백질(MA) 코딩 서열을 포함하고, NC 코딩 서열 또는 MA 코딩 서열 중 하나 또는 둘 모두는 적어도 하나의 비-바이러스 앱타머-결합 단백질(ABP) 뉴클레오티드 서열을 포함하고, 패키징 플라스미드는 기능성 인테그라제 단백질을 인코딩하지 않는다.

[0087] **[0104]** 예를 들어, (a) Gag 뉴클레오티드 서열에 작동 가능하게 연결된 진핵 프로모터를 포함하는 패키징 플라스미드로서, 상기 Gag 뉴클레오티드 서열이 뉴클레오캡시드(NC) 코딩 서열 및 매트릭스 단백질(MA) 코딩 서열을 포함하고, 상기 NC 코딩 서열 또는 MA 코딩 서열 중 하나 또는 둘 모두가 적어도 하나의 비-바이러스 앱타머-결합 단백질(ABP) 뉴클레오티드 서열을 포함하고, 상기 패키징 플라스미드가 기능성 인테그라제 단백질을 인코딩하지 않는, 패키징 플라스미드; (b) (i) CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 융합 단백질을 인코딩하는 핵산 서열; 및 (ii) 본원에 설명된 gRNA를 포함하는 적어도 하나의 포유동물 발현 플라스미드; 및 (c) 외피 당단백질 코딩 서열을 포함하는 외피 플라스미드를 포함하는 렌트바이러스 패키징 시스템이 본원에 제공된다.

[0088] **[0105]** 시스템은 제2 세대 패키징 플라스미드 또는 제3 세대 패키징 플라스미드 또는 이들의 변형된 버전을 포

함할 수 있다. 일부 예에서, 패키징 플라스미드는 상기 설명된 바와 같은 Gag 뉴클레오티드 서열을 포함하고, Rev 뉴클레오티드 서열 및 Tat 뉴클레오티드 서열을 추가로 포함한다. 다른 예에서, 시스템은 상기 설명된 바와 같은 Gag 뉴클레오티드 서열을 포함하는 제1 패키징 플라스미드 및 Rev 뉴클레오티드 서열을 포함하는 제2 패키징 플라스미드를 포함한다. 각각의 패키징 플라스미드에서, 바이러스 단백질 코딩 서열은 진핵 프로모터, 예를 들어, 각각 개별적으로 또는 다중 단백질 코딩 서열에 대한 하나의 프로모터에 작동 가능하게 연결된다. 시스템은 제2 세대 패키징 플라스미드 또는 제3 세대 패키징 플라스미드 또는 이들의 변형된 버전을 포함할 수 있다.

[0089] **[0106]** 일부 예에서, ABP 코딩 서열은 바이러스 단백질 코딩 서열의 5' 말단 또는 3' 말단에, 즉, NC 또는 MA 코딩 서열의 5' 말단 또는 3' 말단에 있다. 일부 예에서, ABP 코딩 서열은 인코딩된 ABP가 바이러스 단백질에 융합되도록 바이러스 단백질 코딩 서열에 삽입될 수 있다. ABP 코딩 서열은 바이러스 단백질 코딩 서열 내의 내부 위치에서 프레임 내에 삽입될 수 있다. 바이러스 단백질 코딩 서열의 5' 또는 3' 말단 근처의 내부 위치에서 프레임 내에 위치할 때, ABP 코딩 서열은 전문이 본원에 참조로서 포함되는 문헌[Tritch, R.J. et al., *J. Virol.* 65(2):922-30 (1991) 및 Scarlata, S. and Carter, C., *Biochimica et Biophysica Acta* - Biomembranes 1614(1):62-72 (2003)]에 설명된 것들과 같은 프로세싱 서열을 방해하지 않도록 위치된다. 예를 들어, Gag 뉴클레오티드 서열은 특히 NC 코딩 서열 및 MA 코딩 서열을 인코딩하고, Gag 전구체 단백질은 단백질분해 절단에 의해 별도의 성숙한 바이러스 단백질로 프로세싱된다. ABP 코딩 서열의 프레임내 삽입은 단백질분해 절단을 위한 프로세싱 서열을 인코딩하는 뉴클레오티드를 방해하지 않을 것이다. 일부 예에서, 바이러스 단백질 코딩 서열의 뉴클레오티드는 ABP 단백질 코딩 서열로 대체될 수 있다. 일부 예에서, 3-6개의 아미노산을 인코딩하는 링커 서열은 발현시 단백질 도메인의 적절한 폴딩을 용이하게 하는 것을 돕기 위해 바이러스 단백질 코딩 서열과 ABP 코딩 서열 사이에 또는 ABP 코딩 서열의 측면에 위치할 수 있다.

[0090] **[0107]** 일 예에서, 변형된 바이러스 단백질은 NC이고 ABP 코딩 서열은 NC 코딩 서열의 5' 말단 또는 3' 말단에 삽입된다. 또 다른 예에서, 변형된 바이러스 단백질은 NC이고 ABP 코딩 서열은 징크 핑거(ZF) 도메인 중 하나의 앞 또는 뒤에 삽입된다. 예를 들어, ABP 코딩 서열은 제2 ZF(ZF2) 도메인의 마지막 코돈 뒤에 삽입될 수 있다. 또 다른 예에서, ABP 코딩 서열은 ZF2 도메인의 제1 코돈 앞에 삽입될 수 있다. 또 다른 예에서, ABP 코딩 서열은 제1 ZF(ZF1) 도메인의 제1 코돈 앞에 삽입될 수 있다. 또 다른 예에서, ABP 코딩 서열은 제1 ZF(ZF1) 도메인의 마지막 코돈 뒤에 삽입될 수 있다. 일부 예에서, ABP 코딩 서열은 NC 단백질에서 고도로 양성인 아미노산 스트레치를 방해하지 않는 방식으로 NC 코딩 서열에 삽입된다.

[0091] **[0108]** 또 다른 예에서, 변형된 바이러스 단백질은 MA이고 ABP 코딩 서열은 MA 코딩 서열의 5' 말단 또는 3' 말단에 삽입된다. 또 다른 예에서, ABP 코딩 서열은 MA 코딩 서열 내의 내부 위치에서 프레임 내에 삽입된다. 일부 예에서, MA 코딩 서열의 뉴클레오티드는 ABP 단백질 코딩 서열로 대체될 수 있다. 예를 들어, MA 단백질의 아미노산 44-132를 인코딩하는 뉴클레오티드는 ABP 코딩 서열로 대체될 수 있다. 또 다른 예에서, ABP 코딩 서열은 MA 단백질의 아미노산 44를 인코딩하는 코돈 앞에 삽입된다. 또 다른 예에서, ABP 코딩 서열은 MA 단백질의 아미노산 132를 인코딩하는 코돈 뒤에 삽입된다.

[0092] **[0109]** 일부 예에서, 시스템은 NEF 코딩 서열 또는 VPR 코딩 서열에 작동 가능하게 연결된 진핵 프로모터를 포함하는 패키징 플라스미드를 포함하고, 여기서 NEF 코딩 서열 또는 VPR 코딩 서열은 적어도 하나의 비-바이러스 ABP 뉴클레오티드 서열을 포함한다. 시스템은 제2 세대 패키징 플라스미드 또는 제3 세대 패키징 플라스미드 또는 이들의 변형된 버전을 포함할 수 있다. 일부 예에서, 패키징 플라스미드는 Gag 뉴클레오티드 서열, Rev 뉴클레오티드 서열, 및 Tat 뉴클레오티드 서열을 포함한다. 다른 예에서, 시스템은 Gag 뉴클레오티드 서열을 포함하는 제1 패키징 플라스미드 및 Rev 뉴클레오티드 서열을 포함하는 제2 패키징 플라스미드를 포함한다.

[0093] **[0110]** 일부 예에서, 변형된 바이러스 단백질은 VPR이고 ABP 코딩 서열은 VPR 코딩 서열의 5' 말단 또는 3' 말단에 삽입된다. 일 예에서, ABP 코딩 서열은 VPR 코딩 서열의 5' 말단에 삽입된다.

[0094] **[0111]** 다른 예에서, 변형된 바이러스 단백질은 NEF이고 ABP 코딩 서열은 NEF 코딩 서열의 5' 말단 또는 3' 말단에 삽입된다. 일 예에서, ABP 코딩 서열은 NEF 코딩 서열의 3' 말단에 삽입된다.

[0095] **[0112]** 일부 예에서, 바이러스 단백질의 코딩 서열은 SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, 또는 SEQ ID NO:25 중 하나일 수 있다. 일부 예에서, 바이러스 단백질의 아미노산 서열은 SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, 또는 SEQ ID NO:26 중 하나일 수 있다. 일부 예에서, 렌티바이러스 패키징 플라스미드는 진핵 프로모터에 작동 가능하게 연결된 SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, 또는 SEQ ID NO:26 중 적어도 하나를 인코딩하는 서열을 포함한다. 일부 예에서, 바이러스 단백질이 NEF인 경우, 폴리펩티드는 바이러스 캡시드에서 패키징을 향상시키는 3개의 돌연변이, 예를 들어, 다음 치환 돌연변이: G3C, V153L, 및 E177G를 포함할 수

있다.

[0096] **[0113]** 일부 예에서, 플라스미드는 이에 융합된 2개 이상의 앱타머-결합 단백질을 포함하는 하나 이상의 바이러스 단백질을 인코딩할 수 있다. 특정 예에서, 렌티바이러스 패키징 플라스미드의 Gag 뉴클레오티드 서열은 NC 코딩 서열 및 MA 코딩 서열을 포함할 수 있고, 여기서 NC 코딩 서열 또는 MA 코딩 서열 중 하나 또는 둘 모두는 제1 비-바이러스 ABP 뉴클레오티드 서열 및 제2 비-바이러스 ABP 뉴클레오티드 서열을 포함한다. 제1 비-바이러스 ABP 뉴클레오티드 서열 및 제2 비-바이러스 ABP 뉴클레오티드 서열은 둘 모두 동일한 ABP를 인코딩할 수 있다. 대안적으로, 제1 비-바이러스 ABP 뉴클레오티드 서열 및 제2 비-바이러스 ABP 뉴클레오티드 서열은 상이한 ABP를 인코딩한다. 일부 예에서, 렌티바이러스 패키징 플라스미드의 Gag 뉴클레오티드 서열은 적어도 하나의 제1 비-바이러스 ABP 뉴클레오티드 서열을 포함하는 NC 코딩 서열 및 적어도 하나의 제2 비-바이러스 ABP 뉴클레오티드 서열을 포함하는 MA 코딩 서열을 포함할 수 있다. 적어도 하나의 제1 비-바이러스 ABP 뉴클레오티드 서열 및 적어도 하나의 제2 비-바이러스 ABP 뉴클레오티드 서열은 둘 모두 동일한 ABP를 인코딩할 수 있다. 대안적으로, 적어도 하나의 제1 비-바이러스 ABP 뉴클레오티드 서열 및 적어도 하나의 제2 비-바이러스 ABP 뉴클레오티드 서열은 상이한 ABP를 인코딩한다.

[0097] **[0114]** 특정 예에서, 패키징 플라스미드는 VPR 코딩 서열 또는 NEF 코딩 서열을 인코딩할 수 있고, 여기서 VPR 코딩 서열 또는 NEF 코딩 서열은 제1 비-바이러스 ABP 뉴클레오티드 서열 및 제2 비-바이러스 ABP 뉴클레오티드 서열을 포함한다. 제1 비-바이러스 ABP 뉴클레오티드 서열 및 제2 비-바이러스 ABP 뉴클레오티드 서열은 둘 모두 동일한 ABP를 인코딩할 수 있다. 대안적으로, 제1 비-바이러스 ABP 뉴클레오티드 서열 및 제2 비-바이러스 ABP 뉴클레오티드 서열은 상이한 ABP를 인코딩한다.

[0098] **[0115]** 비-바이러스 앱타머-결합 단백질(ABP) 뉴클레오티드 서열은 RNA 앱타머 서열에 결합하는 폴리펩티드 서열을 인코딩한다. 여러 비-바이러스 ABP가 본 발명의 개시에 사용하기에 적합하다. 특히, 적합한 ABP는 RNA 앱타머 서열로 지칭되는 스템-루프 구조를 형성하는 RNA 서열에 특이적으로 결합하는 박테리오파지 RNA-결합 단백질을 포함한다. 예시적인 비-바이러스 앱타머 결합 단백질은 MS2 코트 단백질, PP7 코트 단백질, 람다 N 펩티드, 및 Com(mom의 대조군; Control of mom) 단백질을 포함한다. 람다 N 펩티드는 단백질의 RNA-결합 도메인인 람다 N 단백질의 아미노산 1-22일 수 있다. 일부 예에서, ABP는 이량체로서 이들의 앱타머에 결합한다. 이러한 ABP 및 이들이 결합하는 앱타머 서열에 대한 정보는 표 1에 제공된다. 일부 구현예에서, 적어도 하나의 비-바이러스 ABP 뉴클레오티드 서열은 SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:14, 또는 SEQ ID NO:16 중 어느 하나에 제시된 서열을 갖는 폴리펩티드를 인코딩한다. 일부 구현예에서, 적어도 하나의 비-바이러스 ABP 뉴클레오티드 서열은 SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:13, 또는 SEQ ID NO:15 중 어느 하나를 포함한다.

[0099] **[0116]** 본원에 제공된 렌티바이러스 패키징 플라스미드의 특징은 이들이 기능성 인테그라제 단백질을 인코딩하지 않을 수 있다는 것이다. 패키징 플라스미드가 기능성 인테그라제 단백질을 인코딩하지 않고 이들이 본원에 설명된 시스템 및 방법에서 사용될 때, 이러한 패키징 플라스미드를 사용하여 생산된 렌티바이러스-유사 입자에 의해 운반되는 핵산 분자가 형질도입된 진핵 세포의 유전체에 통합될 위험이 실질적으로 감소된다. 일부 예에서, 렌티바이러스 패키징 플라스미드는 내부에 인테그라제-불활성화 돌연변이를 갖는 인테그라제 코딩 서열을 포함한다. 예를 들어, 인테그라제-불활성화 돌연변이는 인테그라제 코딩 서열에 의해 인코딩된 인테그라제 단백질의 아미노산 위치 64에서의 아스파르트산에서 발린으로의 돌연변이(D64V)일 수 있다. 일부 예에서, 렌티바이러스 패키징 플라스미드는 인테그라제 코딩 서열의 전부 또는 일부의 결실을 포함한다.

[0100] **[0117]** 일부 구현예에서, 렌티바이러스 패키징 플라스미드는 Gag 뉴클레오티드 서열에 작동 가능하게 연결된 진핵 프로모터를 포함한다. 일부 구현예에서, 포유동물 발현 플라스미드는 VPR 코딩 서열 또는 NEF 코딩 서열에 작동 가능하게 연결된 진핵 프로모터를 포함한다. 일부 예에서, 진핵 프로모터는 RNA 중합효소 II 프로모터이다. RNA 중합효소 II 프로모터 서열은 포유동물 종으로부터 선택된다. 예를 들어, 프로모터 서열은 몇 가지 예를 들어, 인간, 소, 양, 버팔로, 돼지, 또는 마우스로부터 선택될 수 있다. 일부 예에서, RNA 중합효소 II 프로모터 서열은 CMV, FE1α, 또는 SV40 서열이다. 일부 예에서, RNA 중합효소 II 서열은 변형된 RNA 중합효소 II 서열이다. 예를 들어, 임의의 포유동물 종으로부터의 야생형 RNA 중합효소 II 프로모터 서열에 대해 적어도 80%, 85%, 90%, 95%, 또는 99% 동일성을 갖는 RNA 중합효소 II 서열은 본원에 제공된 작제물에 사용될 수 있다. 당업자는 상기 설명된 바와 같이 2개의 폴리펩티드 또는 핵산의 동일성을 결정하는 방법을 용이하게 이해한다.

[0101] **[0118]** RNA pol II 프로모터로부터 전사된 코딩 서열은 코딩 서열의 하류에 폴리(A) 신호 및 전사 종료자 서열을 포함한다. 통상적으로 사용되는 포유동물 종료자(예를 들어, SV40, hGH, BGH, 및 rbGlob)는 폴리아데닐화 및

종료 둘 모두를 촉진하는 서열 모티프 AAUAAA를 포함한다. 서열-기반 요소인 종료자의 역할은 전사 단위(예를 들어, 유전자)의 말단을 규정하고 전사 기구로부터 새로 합성된 RNA를 방출하는 과정을 개시하는 것이다. 종료자는 전사될 유전자의 하류에서 발견되며, 전형적으로 폴리아데닐화 또는 폴리(A) 신호와 같은 임의의 3' 조절 요소 직후에 발생한다.

[0102] **[0119]** 일부 예에서, 렌티바이러스 패키징 플라스미드는 하나 이상의 발현 카세트를 포함할 수 있다.

[0103] **[0120]** 시스템은 또한 바이러스 외피 당단백질을 인코딩하는 외피 코딩 서열을 갖는 외피 플라스미드를 포함할 수 있다. 예를 들어, Env 뉴클레오티드 서열은 VSV-G를 인코딩할 수 있다. 외피 코딩 서열은 진핵 프로모터에 작동 가능하게 연결된다. 적절한 진핵 프로모터는 상기 설명되어 있다. 일부 예에서, 진핵 프로모터는 RNA pol II 프로모터이다.

[0104] **[0121]** 시스템은 패키징 플라스미드, 외피 플라스미드 및 포유동물 발현 플라스미드, 즉, (i) CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 폴리펩티드를 포함하는 융합 단백질을 인코딩하는 핵산 서열; 및 (ii) 본원에 설명된 적어도 하나의 앱타머를 포함하는 gRNA를 포함하는 포유동물 발현 플라스미드 중 임의의 것을 포함할 수 있다. 본원에 설명된 패키징 플라스미드, 포유동물 발현 플라스미드 및 외피 플라스미드 중 임의의 것이 시스템으로서 진핵 세포에 전달되는 경우, 포유동물 발현 플라스미드에 의해 발현된 gRNA는 포유동물 발현 플라스미드에 의해 발현된 CRISPR-관련 엔도뉴클레아제와 복합체를 형성하여 RNP를 형성하며, 이는 gRNA에 융합되거나 연결된 앱타머와 패키징 플라스미드에 의해 발현된 바이러스 단백질에 연결된 ABP 사이의 상호작용을 통해 진핵 세포에 의해 생산된 바이러스 입자에 의해 패키징된다.

[0105] **[0122]** 또한, 본 발명의 개시에 설명된 시스템의 성분을 포함하는 키트가 본원에 제공된다. 일부 구현예에서, 키트는 본원에 설명된 플라스미드 중 하나 이상을 포함한다.

[0106] **렌티바이러스-유사 입자**

[0107] **[0123]** 또 다른 양태에서, 렌티바이러스-유사 입자, 예를 들어, 본원에 설명된 임의의 방법에 의해 제조된 렌티바이러스-유사 입자가 제공된다. 본원에서 사용되는 렌티바이러스-유사 입자는 진정한 고유 바이러스의 조직 및 입체형태를 모방하지만 바이러스 유전체가 없는 다중단백질 구조이다. 복수의 렌티바이러스-유사 입자가 또한 제공된다. 렌티바이러스-유사 입자는 적어도 하나의 앱타머-결합 단백질이 하나 이상의 바이러스 단백질에 융합된 융합 단백질인 변형된 렌티바이러스 단백질을 함유한다. 본 발명의 개시의 맥락에서, 변형된 바이러스 단백질은 구조적 또는 비구조적일 수 있다. 예시적인 구조 단백질은 렌티바이러스 뉴클레오캡시드(NC) 단백질 및 매트릭스(MA) 단백질이다. 예시적인 비-구조 단백질은 바이러스 단백질 R(VPR) 및 음성 조절 인자(NEF)이다. 일부 예에서, 입자는 NC 단백질 및 MA 단백질을 포함하는 융합 단백질을 함유하고, 여기서 이들 중 하나 또는 둘 모두는 적어도 하나의 비-바이러스 앱타머 결합 단백질(ABP)과 융합된다. 입자의 NC 단백질은 2개의 기능성 징크 핑거 단백질 도메인을 가질 수 있다. 특히, 제2 NC 징크 핑거 도메인의 보유는 바이러스 조립 및 버딩(budding)의 효율을 보존할 수 있다. 일부 예에서, 입자는 VPR 단백질 또는 NEF 단백질을 포함하는 융합 단백질을 함유하며, 여기서 VPR 단백질 또는 NEF 단백질은 적어도 하나의 비-바이러스 ABP와 융합된다. 입자는 또한 (i) CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 폴리펩티드를 포함하는 융합 단백질; 및 (ii) gRNA를 포함하는 RNP를 함유한다. 적어도 하나의 앱타머가 gRNA 서열에 부착되거나 삽입된 비-바이러스 핵산 서열을 포함하는 본원에 설명된 임의의 포유동물 발현 플라스미드는 RNP를 함유하는 렌티바이러스-유사 입자를 생성하는데 사용될 수 있다. 일부 예에서, 렌티바이러스-유사 입자는 기능성 인테그라제 단백질을 함유하지 않는다. 이러한 바이러스-유사 입자는 관심 진핵 세포를 형질도입하는데 유용하다.

[0108] **[0124]** 입자는 하나 이상의 ABP를 포함하는 바이러스 융합 단백질을 포함할 수 있다. 일부 예에서, 입자는 NC 단백질, MA 단백질, 또는 둘 모두를 함유하며, 여기서 NC 단백질 또는 MA 단백질 중 하나 또는 둘 모두는 하나 이상의 비-바이러스 ABP와 융합된다. 일부 예에서, 렌티바이러스-유사 입자는 적어도 하나의 비-바이러스 ABP와 융합된 NC 단백질을 포함한다. 일부 예에서, 렌티바이러스-유사 입자는 적어도 하나의 비-바이러스 ABP와 융합된 MA 단백질을 포함한다. 일부 예에서, 렌티바이러스-유사 입자는 NC 단백질 및 MA 단백질을 포함할 수 있고, 여기서 NC 단백질 또는 MA 단백질 중 하나 또는 둘 모두는 2개의 비-바이러스 ABP 단백질, 제1 비-바이러스 ABP 및 제1 비-바이러스 ABP의 C' 말단에 융합된 제2 비-바이러스 ABP와 융합(즉, 일렬로)될 수 있다. 특정 예에서, 입자는 제1 비-바이러스 ABP 및 제2 비-바이러스 ABP와 융합된 NC 단백질 또는 MA 단백질 중 하나 또는 둘 모두를 함유할 수 있다.

[0109] **[0125]** 일부 예에서, 렌티바이러스-유사 입자는 VPR 단백질 또는 NEF 단백질을 함유하고, 여기서 VPR 단백질 또

는 NEF 단백질은 하나 이상의 비-바이러스 ABP에 융합된다. 일부 예에서, 렌티바이러스-유사 입자는 2개의 비-바이러스 ABP, 제1 비-바이러스 ABP 및 제1 비-바이러스 ABP의 C' 말단에 융합된 제2 비-바이러스 ABP에 융합(즉, 일렬로)된 VPR 단백질 또는 NEF 단백질을 함유한다. 일부 예에서, 렌티바이러스-유사 입자는 제1 비-바이러스 ABP 및 제2 비-바이러스 ABP에 융합된 VPR 단백질 또는 NEF 단백질을 함유한다. 제1 비-바이러스 ABP 및 제2 비-바이러스 ABP는 둘 모두 동일한 ABP일 수 있다. 대안적으로, 제1 비-바이러스 ABP 및 제2 비-바이러스 ABP는 상이한 ABP일 수 있다. 일부 예에서, 렌티바이러스-유사 입자는 MA 단백질에 융합된 적어도 하나의 제1 비-바이러스 ABP와 이의 C' 말단에 융합된 적어도 하나의 제2 비-바이러스 ABP를 갖는 NC 단백질을 포함할 수 있다. 적어도 하나의 제1 비-바이러스 ABP 및 적어도 하나의 제2 비-바이러스 ABP 둘 모두는 동일한 ABP이다. 대안적으로, 적어도 하나의 제1 비-바이러스 ABP 단백질 및 적어도 하나의 제2 비-바이러스 ABP는 상이한 ABP일 수 있다. 제1 비-바이러스 ABP 및 제2 비-바이러스 ABP는 둘 모두 동일한 ABP일 수 있다. 대안적으로, 제1 비-바이러스 ABP 및 제2 비-바이러스 ABP는 상이한 ABP일 수 있다.

[0110] **[0126]** 비-바이러스 ABP는 RNA 앱타머 서열에 결합하는 폴리펩티드 서열이다. 여러 비-바이러스 ABP가 본 발명의 개시에 사용하기에 적합하다. 특히, 적합한 ABP는 스템-루프 구조를 형성하는 RNA 서열인 공지된 RNA 앱타머 서열에 특이적으로 결합하는 박테리오파지 RNA-결합 단백질을 포함한다. 예시적인 비-바이러스 앱타머 결합 단백질은 MS2 코트 단백질, PP7 코트 단백질, 람다 N 펩티드, 및 Com(mom의 대조군) 단백질을 포함한다. 람다 N 펩티드는 단백질의 RNA-결합 도메인인 람다 N 단백질의 아미노산 1-22일 수 있다. 이러한 ABP 및 이들이 결합하는 앱타머 서열에 대한 정보는 상기 표 1에 제공된다.

[0111] **[0127]** 렌티바이러스-유사 입자는 다양한 렌티바이러스 단백질을 포함할 수 있다. 그러나, 일부 경우에, 렌티바이러스-유사 입자는 고유 렌티바이러스에서 발견되는 모든 유형의 단백질 또는 핵산을 포함하지 않는다. 일부 예에서, 입자는 NC, MA, CA, SP1, SP2, P6, POL, ENV, TAT, REV, VIF, VPU, VPR, 및/또는 NEF 단백질, 또는 이들 중 임의의 유도체, 조합, 또는 일부를 함유할 수 있다. 일부 예에서, 입자는 NC, MA, CA, SP1, SP2, P6, 및 POL을 함유할 수 있다. 일부 예에서, 렌티바이러스-유사 입자는 바이러스 쉘(캡시드)을 형성하는 단백질만을 포함할 수 있다. 일부 예에서, 하나 이상의 렌티바이러스 단백질은 렌티바이러스-유사 입자로부터 전체적으로 또는 부분적으로 배제될 수 있다. 예를 들어, 일부 경우에, 렌티바이러스-유사 입자는 POL 단백질을 함유하지 않을 수 있거나, 예를 들어, 불활성화 점 돌연변이 또는 불활성화 트렁케이션을 갖는 POL 단백질과 같은 POL 단백질의 비-기능성 버전을 포함할 수 있다. 또 다른 예에서, 렌티바이러스-유사 입자는 인테그라제 단백질을 함유하지 않을 수 있거나, 예를 들어, 불활성화 점 돌연변이 또는 불활성화 트렁케이션을 갖는 인테그라제 단백질과 같은 인테그라제 단백질의 비-기능성 버전을 포함할 수 있다. 예를 들어, 렌티바이러스-유사 입자는 아미노산 위치 64에서 아스파르트산에서 발린으로의 돌연변이(D64V)를 포함하는 비-기능성 인테그라제 단백질을 함유할 수 있다. 또 다른 예에서, 렌티바이러스-유사 입자는 역전사효소 단백질을 함유하지 않을 수 있거나, 예를 들어, 불활성화 점 돌연변이 또는 불활성화 트렁케이션을 갖는 역전사효소 단백질과 같은 역전사효소 단백질의 비-기능성 버전을 포함할 수 있다.

[0112] **[0128]** 상기 기재된 바와 같이, gRNA는 일반적으로 DNA 표적화 서열 및 CRISPR-관련 엔도뉴클레아제와 상호작용하는 불변 영역을 포함한다. 일부 예에서, gRNA는 트랜스활성화 crRNA(tracrRNA) 서열을 포함할 수 있다. 예를 들어, gRNA는 Cas9 단백질 또는 유도체와 함께 사용되는 tracrRNA를 포함할 수 있다. 다른 예에서, gRNA는 tracrRNA 서열을 포함하지 않는다. 예를 들어, gRNA는 Cpf1 단백질 또는 유도체와 함께 사용되는 tracrRNA 서열을 포함하지 않을 수 있다.

[0113] **[0129]** 일부 예에서, gRNA는 적어도 하나의 앱타머 서열을 포함한다. 일부 예에서, 적어도 하나의 앱타머 서열은 gRNA의 5' 말단 또는 3' 말단에 위치할 수 있다. 일부 예에서, 적어도 하나의 앱타머 서열은 gRNA 내의 내부 위치, 예를 들어, 폴딩된 gRNA에 형성된 루프 중 하나 이상에 삽입될 수 있다. 예를 들어, gRNA가 Cas9 단백질용인 경우, 적어도 하나의 앱타머 서열은 gRNA의 테트라 루프, 스템 루프 2(ST2), 또는 3' 말단에 위치할 수 있다. 일부 예에서, 1-30개의 리보뉴클레오티드의 스페이서는 gRNA와 적어도 하나의 앱타머 서열 사이에, 또는 적어도 하나의 앱타머 서열의 측면에 위치할 수 있다. 특정 예에서, 적어도 하나의 앱타머 서열은 진핵 세포의 렌티바이러스-유사 입자 형질도입을 방해하지 않는다. 예를 들어, NC 단백질, MA 단백질, VPR 단백질, 또는 NEF 단백질 중 하나 이상에 융합된 적어도 하나의 비-바이러스 ABP는 진핵 세포의 렌티바이러스-유사 입자 형질도입을 방해하지 않을 수 있다.

[0114] **유전자 편집 방법**

[0115] **[0130]** 진핵 세포의 유전체에서 DNA 표적, 예를 들어, 유전자를 편집하기 위한 CRISPR/Cas 시스템에서 본 발명

의 개시에 제공된 플라스미드 및 시스템을 사용하는 방법이 본원에 설명된다. 본원에 제공된 방법에서, 변형될 관심 표적 유전체 서열을 포함하는 진핵 세포는 적어도 하나의 앱타머-결합 단백질(ABP)에 융합된 바이러스 단백질 및 (1) gRNA 및 (2) CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 폴리펩티드를 포함하는 융합 단백질을 포함하는 RNP를 포함하는 바이러스 융합 단백질을 함유하는 렌티바이러스-유사 입자로 형질도입된다.

[0116] **[0131]** 본원에 설명된 방법은 임의의 크기의 표적-내 결실을 감소시키면서 세포의 유전체를 편집하는데 사용될 수 있다. 예를 들어, 100 bp, 90bp, 80 bp, 70 bp, 60 bp, 50 bp, 40 bp 30 bp, 20 bp, 10 bp, 9 bp, 8 bp, 7 bp, 6 bp, 5 bp, 4 bp, 3 bp, 2 bp 또는 1 bp 미만의 결실이 감소될 수 있다. 일부 구현예에서, 크기가 100 bp, 200 bp, 300 bp, 400 bp, 또는 500 bp 초과의 큰 유전체 결실이 감소될 수 있다. 본원에서 사용되는 용어 "표적-내 결실"은 세포의 유전체에서 sgRNA 표적 부위 또는 그 근처에서, 즉, 변형될 관심 표적 유전체 서열의 부근에서 발생하는 결실을 지칭한다. 본원에 제공된 방법에서, 큰 유전체 결실은 세포가 DNA 중합효소 도메인을 포함하는 폴리펩티드에 융합되거나 연결되지 않은 CRISPR-관련 엔도뉴클레아제로 편집되는 경우의 큰 유전체 결실의 수와 비교하여 적어도 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 90%, 95%, 99% 이상만큼 감소될 수 있다.

[0117] **[0132]** 일부 예에서, 제공된 방법은 DNA 중합효소 도메인을 포함하는 폴리펩티드와 융합되거나 회합되지 않은 CRISPR-관련 엔도뉴클레아제로 편집된 세포에서 비-NHEJ와 비교하여 비-상동성 말단 연결(NHEJ)을 증가시킨다. 일부 구현예에서, NJEH는 세포에서 비-NHEJ 말단 연결과 비교하여 증가된다. 일부 구현예에서, 비-NJEH는 미세상동성-매개 말단 연결(MMEJ) 및/또는 단일-가닥 어닐링(SSA)이다. 일부 구현예에서, 증가는 적어도 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, 또는 400% 증가이다. 일부 구현예에서, 증가는 2배, 3배, 4배, 5배, 6배, 7배, 8배, 9배, 10배, 20배, 30배, 40배, 50배 증가 또는 그 초과이다.

[0118] **[0133]** 일부 구현예에서, NHEJ 대 비-NHEJ의 비는 세포에서 증가한다. 일부 방법에서, 표적-내 1 염기쌍(1-bp) 결실 대 1-bp 초과의 표적-내 결실의 비는 세포에서 증가한다. 예를 들어, 일부 구현예에서, 표적-내 1 염기쌍(1-bp) 결실 대 500개 염기쌍 초과의 결실의 비는 세포에서 증가한다.

[0119] **[0134]** 일부 구현예에서, 비-주형 삽입(비-TIS)의 수는 세포에서 감소한다. 일부 구현예에서, 감소는 적어도 10%, 20%, 30%, 40%, 50%, 60%, 80%, 90%, 95%, 99%, 또는 100%이다.

[0120] **[0135]** 일부 구현예에서, 1-bp 주형 삽입(TIS)의 수는 세포에서 증가한다. 일부 구현예에서, 증가는 적어도 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 100%, 200%, 300%, 또는 400% 증가이다. 일부 구현예에서, 증가는 2배, 3배, 4배, 5배, 6배, 7배, 8배, 9배, 10배, 20배, 30배, 40배, 50배 증가 또는 그 초과이다.

[0121] **[0136]** 일부 구현예에서, 1-bp TIS 대 1-bp 비-TIS의 비는 세포에서 증가한다. 본원에서 사용되는 "주형 삽입"은 Cas9-유도된 이중 가닥 파손이 후속하여 DNA 중합효소에 의해 채워지고 라이게이션되는 1-nt 5' 오버행 염기를 남긴 후에 발생하는 비무작위 삽입이다.

[0122] **[0137]** 일부 구현예에서, MRE11 이중 가닥 파손 복구 뉴클레아제(MRE11)의 엑소뉴클레아제 활성은 (예를 들어, 야생형 MRE11에 비해) 감소된다. 예를 들어, 본원에 제공된 방법에서, MRE11의 엑소뉴클레아제 활성은 적어도 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 90%, 95%, 99% 또는 그 초과만큼 감소될 수 있다.

[0123] **[0138]** 인테그라제 활성이 없는 렌티바이러스-유사 입자가 방법에 사용되는 경우, 입자에 의해 운반되는 임의의 핵산의 세포 유전체로의 통합 위험이 감소된다. 일부 예에서, 사용되는 렌티바이러스-입자는 바이러스 복제에 필수적인 렌티바이러스 유전체 서열의 일부가 결여되어 있고, 따라서, 계속되는 입자 생산의 위험을 감소시킨다. 제공된 성분의 또 다른 이점은 바이러스 융합 단백질이 렌티바이러스-유사 입자로의 RNP의 패키징을 증가시킬 수 있고, 이는 차례로 유전체 편집 효율을 증가시킬 수 있다는 것이다.

[0124] **[0139]** 일부 예에서, 형질도입된 진핵 세포는 포유동물 세포이다. 일부 예에서, 진핵 세포는 시험관내 배양된 세포일 수 있다. 일부 예에서, 진핵 세포는 대상체로부터 획득된 생체외 세포일 수 있다. 다른 예에서, 진핵 세포는 대상체에 존재한다. 전체에 걸쳐 사용되는 바와 같이, 대상체는 개인을 의미한다. 예를 들어, 대상체는 영장류와 같은 포유동물, 및 보다 구체적으로 인간이다. 비-인간 영장류도 대상체이다. 용어 대상체는 길들여진 동물, 예를 들어, 고양이, 개 등, 가축(예를 들어, 소, 말, 돼지, 양, 염소 등) 및 실험 동물(예를 들어, 흰 족제비, 친칠라, 마우스, 토끼, 래트, 저빌, 기니피그 등)을 포함한다. 따라서, 수의학적 용도 및 의학적 용도 및 제형이 본원에서 고려된다. 이 용어는 특정 연령 또는 성별을 나타내지 않는다. 따라서, 남성이든 여성이든 성인 및 신생아 대상체가 포함되도록 의도된다. 본원에서 사용되는 바와 같이, 환자 또는 대상체는 상호교환적으로 사용될 수 있고, 질병 또는 장애를 앓고 있는 대상체를 지칭할 수 있다. 본원에 제공된 렌티바이러스-유사

입자는 대상체에게 투여될 수 있고, 예를 들어, 공지된 일상적인 방법에 따라 대상체에 주사될 수 있다. 예시적인 투여 방식은 경구, 직장, 경점막, 국소, 비강내, 흡입(예를 들어, 에어로졸을 통함), 협측(예를 들어, 설하), 질, 척수강내, 안내, 경피, 피내, 흉막내, 뇌내, 및 관절내), 국소 등, 뿐만 아니라 직접 조직 또는 기관 주사를 포함한다. 투여는 또한 종양일 수 있다. 임의의 주어진 경우에 가장 적합한 경로는 치료되는 질환의 성질 및 중증도 및 사용되는 특정 렌티바이러스-유사 입자의 성질에 의존할 것이다. 일부 예에서, 렌티바이러스-유사 입자는 정맥내(IV), 복강내(IP), 근육내, 또는 특정 기관 또는 조직으로 주사된다. 일부 구현예에서, 다양한 간격, 예를 들어, 매일, 매주, 매월, 매년 등의 기간에 걸쳐 요망되는 수준의 유전자 편집을 달성하기 위해 1회 초과의 투여(예를 들어, 2회, 3회, 4회 또는 그 초과의 투여)가 이용될 수 있다.

[0125] **[0140]** 본원에 설명된 임의의 재조합 렌티바이러스-유사 입자의 유효량은 다양할 것이며, 실험 및/또는 임상 시험을 통해 당업자에 의해 결정될 수 있다. 예를 들어, 유효 용량은 약 10^6 내지 약 10^{15}개의 렌티바이러스-유사 입자, 예를 들어, 약 10^6 내지 약 10^{14}, 약 10^6 내지 약 10^{13}, 약 10^6 내지 약 10^{12}개의 렌티바이러스-유사 입자, 약 10^6 내지 약 10^{12}, 약 10^6 내지 약 10^{11}, 또는 약 10^6 내지 약 10^{11}개의 렌티바이러스-유사 입자일 수 있다. 다른 효과적인 투여량은 용량 반응 곡선을 확립하는 일상적인 시험을 통해 당업자에 의해 용이하게 확립될 수 있다. 예를 들어, 문헌[Mangeot et al. "Genome editing in primary cells and in vivo using viral-derived Nanoblades loaded with Cas9-sgRNA ribonucleoproteins," *Nat Commun* 10, 45 (2019). doi.org/10.1038/s41467-018-07845-z]을 참조한다.

[0126] **[0141]** 일부 예에서, 제공된 방법은 관심 표적 유전자좌를 변형시키기 위한 것이고, 상기 방법은 복수의 바이러스 입자로 복수의 진핵 세포를 형질도입시키는 것을 포함하고, 상기 복수의 바이러스 입자는 (i) 바이러스 단백질, 예를 들어, NC, MA, VRP, 또는 NEF를 포함하는 융합 단백질로서, 상기 바이러스 단백질이 적어도 하나의 비-바이러스 앱타머 결합 단백질(ABP)를 포함하는, 융합 단백질; 및 (ii) (1) gRNA 및 (2) CRISPR-관련 엔도뉴클레아제 및 DNA 중합효소 도메인을 포함하는 폴리펩티드를 포함하는 융합 단백질을 포함하는 리보뉴클레오티드 단백질(RNP) 복합체로서, 상기 RNP 복합체가 gRNA를 통해 세포의 유전체 DNA 내의 유전체 표적 서열에 결합(예를 들어, 우선적으로 결합)할 수 있고 CRISPR-관련 엔도뉴클레아제가 세포의 유전체 DNA를 변경시키는, 리보뉴클레오티드 단백질(RNP) 복합체를 포함한다. 상기 설명된 바와 같이, RNP 복합체는 CRISPR-관련 엔도뉴클레아제와 복합체를 형성하는 gRNA 서열에 부착되거나 삽입된 앱타머 서열의 상호작용을 통해 바이러스 입자에 패키징된다.

[0127] **[0142]** 일반적으로, sgRNA는 유전자 또는 그 근처의 특정 영역을 표적으로 한다. 일부 예에서, sgRNA는, 예를 들어, 겸상적혈구 빈혈을 유발하는 인간 베타-글로빈 유전자에서 단일 염기 돌연변이를 교정하기 위해 단일 염기 변화가 필요한 영역으로 표적화될 수 있다. sgRNA는 본원에 설명된 RNP 복합체가 세포의 유전체 서열의 특정 부위에 있도록 한다.

[0128] **[0143]** 일부 예에서, 본 발명의 개시에 설명된 바와 같은 시스템 성분에 대한 변형은 진핵 세포로의 형질도입 후 시스템 성분이 기능하는 방법을 손상시키지 않는다. 오히려, 성분은 진핵 세포로의 형질도입시 비변형된 성분과 유사하거나 더 잘 기능할 수 있다. 예를 들어, 렌티바이러스-유사 입자의 바이러스 융합 단백질은 진핵 세포의 렌티바이러스-유사 입자 형질도입을 방해하지 않을 수 있다. 유사하게, 렌티바이러스-유사 입자에 패키징된 RNP 복합체가 적어도 하나의 앱타머 서열을 포함하는 경우, 적어도 하나의 앱타머 서열은 진핵 세포의 렌티바이러스-유사 입자 형질도입을 방해하지 않을 수 있다. 일부 예에서, 바이러스 융합 단백질을 함유하는 렌티바이러스-유사 단백질은 바이러스 융합 단백질을 포함하지 않는 렌티바이러스-유사 입자에 비해 진핵 세포로의 형질도입시 더 큰 유전자 편집을 야기할 수 있다. 일 예에서, 바이러스 융합 단백질은 NC-ABP 융합 단백질, 예를 들어, NC-MS2 융합 단백질 또는 NC-PP7 융합 단백질일 수 있다. 일 예에서, NC 융합 단백질은 1개 또는 2개의 ABP, 예를 들어, 1개 또는 2개의 MS2 단백질, 1개 또는 2개의 PP7 단백질, 또는 1개의 MS2 단백질 및 1개의 PP7 단백질에 융합된다.

[0129] **[0144]** 진핵 세포는 시험관내, 생체외 또는 생체내 진핵 세포일 수 있다. 일부 구현예에서, 세포는 일차 세포(대상체로부터 분리됨)이다. 본원에서 사용되는 바와 같이, 일차 세포는 형질전환되거나 무한증식화되지 않은 세포이다. 이러한 일차 세포는 제한된 횟수로 배양, 계대배양, 또는 계대될 수 있다(예를 들어, 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20회 배양됨). 일부 경우에, 일차 세포는 시험관내 배양 조건에 적응된다. 일부 경우에, 일차 세포는 유기체, 시스템, 기관 또는 조직으로부터 분리되고, 선택적으로 분류되고, 배양 또는 계대배양 없이 직접 이용된다. 일부 경우에, 일차 세포는 자극되거나, 활성화되거나, 분화된다. 일부 구현예에서, 세포는 변형된 세포의 집단을 확장시키는데 효과적인 조건 하에 배양된다.

일부 구현예에서, 본원에 제공된 임의의 방법에 의해 변형된 세포가 정제된다. 일부 경우에, 세포는 대상체로부터 제거되고, 본원에 설명된 임의의 방법을 사용하여 변형되고, 환자에게 재투여된다.

[0130] **[0145]** 일부 예에서, 세포가 상기 설명된 바이러스 입자로 형질도입되면, 세포는 유전자 편집이 일어나도록 하기에 충분한 시간 동안 배양되어, 검출 가능한 표현형을 발현하는 세포의 푸울이 복수의 형질도입된 세포로부터 선택될 수 있다. 표현형은, 예를 들어, 세포 성장, 생존, 또는 증식일 수 있다. 일부 예에서, 표현형은 제제, 예를 들어, 세포독성제, 종양발생 유전자, 종양 억제인자, 전사 인자, 키나제(예를 들어, 수용체 티로신 키나제), 프로모터(예를 들어, 이종성 프로모터)의 제어 하에 있는 유전자(예를 들어, 외인성 유전자), 체크포인트 유전자 또는 세포 주기 조절인자, 성장 인자, 호르몬, DNA 손상제, 약물, 또는 화학요법제의 존재 하에서의 세포 성장, 생존, 또는 증식이다. 표현형은 또한 단백질 발현, RNA 발현, 단백질 활성, 또는 세포 운동성, 이동성, 또는 침습성일 수 있다. 일부 예에서, 표현형에 기초하여 세포를 선택하는 것은 형광 활성화 세포 분류, 세포의 친화성 정제, 또는 세포 운동성에 기초한 선택을 포함한다.

[0131] **[0146]** 일부 예에서, 세포를 선택하는 것은 증폭, 시퀀싱, SNP 분석 등에 의한 세포의 유전체 DNA의 분석을 포함한다. 시퀀싱 방법은 샷건 시퀀싱, 브리지 PCR, 생거(Sanger) 시퀀싱(미세유체 생거 시퀀싱 포함), 파이로시퀀싱, 대규모 병렬 시그니처 시퀀싱, 나노포어 DNA 시퀀싱, 단일 분자 실시간 시퀀싱(SMRT)(Pacific Biosciences, Menlo Park, CA), 이온 반도체 시퀀싱, 라이게이션 시퀀싱, 합성에 의한 시퀀싱(Illumina, San Diego, Ca), 폴로니(Polony) 시퀀싱, 454 시퀀싱, 고체 상 시퀀싱, DNA 나노볼 시퀀싱, 헬리스코프(heliscope) 단일 분자 시퀀싱, 질량 분광학 시퀀싱, 파이로시퀀싱, 지지된 올리고 라이게이션 검출(SOLiD) 시퀀싱, DNA 마이크로어레이 시퀀싱, RNAP 시퀀싱, 터널링 전류 DNA 시퀀싱, 및 미래에 확인되는 임의의 다른 DNA 시퀀싱 방법을 포함하나 이에 제한되지 않는다. 본원에 설명된 하나 이상의 시퀀싱 방법은 고처리량 시퀀싱 방법에 사용될 수 있다. 본원에서 사용되는 용어 "고처리량 시퀀싱"은 하나 초과의 핵산 서열이 주어진 시간에 시퀀싱되는 핵산 시퀀싱과 관련된 모든 방법을 지칭한다.

[0132] **치료 방법**

[0133] **[0147]** 본원에 설명된 임의의 방법 및 조성물은 대상체에서 질환(예를 들어, 암, 혈액 장애(예를 들어, 겸상적혈구 빈혈 또는 베타 지중해빈혈), 감염성 질환, 자가면역 질환, 이식 거부, 이식편 대 숙주병 또는 다른 염증성 장애)을 치료하는데 사용될 수 있다. 일부 구현예에서, 감염성 질환은 바이러스 감염, 예를 들어, 비제한적인 예로, 인간 면역결핍 바이러스(HIV), 단순 헤르페스 바이러스 타입 1(헤르페스 기질 각막염), 및 인간 유두종 바이러스(HPV)이다. 일부 구현예에서, 질환은 신경퇴행성 또는 근위축성 질환, 예를 들어, 뒤쉬엔느 근이영양증이다.

[0134] **[0148]** 일부 방법에서, 치료될 암은 B-세포 기원의 암, 유방암, 위암, 신경모세포종, 골육종, 폐암, 결장암, 만성 골수성 암, 백혈병(예를 들어, 급성 골수성 백혈병, 만성 림프구성 백혈병(CLL) 또는 급성 림프구성 백혈병(ALL)), 전립선암, 결장암, 신장 세포 암종, 간암, 신장암, 난소암, 위암, 고환암, 횡문근육종, 및 호지킨 림프종으로부터 선택된다. 일부 구현예에서, B-세포 기원의 암은 B-계통 급성 림프모구 백혈병, B-세포 만성 림프구성 백혈병, 및 B-세포 비-호지킨 림프종으로 구성된 군으로부터 선택된다.

[0135] **[0149]** 일부 방법에서, 대상체의 세포는 생체내에서 변형된다. 일부 방법에서, 대상체에서 질환을 치료하는 방법은 a) 대상체로부터 세포를 획득하고; b) 본원에 제공된 임의의 방법을 사용하여 세포를 변형시키고; c) 변형된 세포를 대상체에게 투여하는 것을 포함한다. 예를 들어, 문헌[Milone and O'Doherty "Clinical sue of lentiviral vectors," *Leukemia* 32, 1529-1541 (2018)]을 참조한다. 선택적으로, 질환은 대상체에서의 암, 혈액 장애(예를 들어, 겸상적혈구 빈혈 또는 베타 지중해빈혈), 감염성 질환, 자가면역 질환, 이식 거부, 이식편 대 숙주병 또는 다른 염증성 장애로 구성된 군으로부터 선택된다. 암을 치료하기 위한 일부 방법에서, 대상체로부터 획득된 세포는 종양 특이적 항원을 발현하도록 변형된다. 전체에 걸쳐 사용되는 바와 같이, 어구 "종양-특이적 항원"은 암 세포에 고유하거나 비-암성 세포에서보다 암 세포에서 더 풍부하게 발현되는 항원을 의미한다. 선택적으로, 대상체로부터 획득된 세포는 T 세포이다. 선택적으로, 변형된 세포는 대상체에 투여되기 전에 확장된다.

[0136] **[0150]** 본원에 설명된 렌티바이러스-유사 입자 또는 세포는 약학적 조성물로서 제형화될 수 있다. 따라서, 본원에 설명된 임의의 렌티바이러스-유사 입자를 포함하는 약학적 조성물이 본원에 제공된다. 본원에 설명된 임의의 변형된 세포를 포함하는 약학적 조성물이 또한 제공된다. 선택적으로, 약학적 조성물은 담체를 추가로 포함할 수 있다. 용어 담체는 렌티바이러스-유사 입자 또는 세포와 조합되는 경우, 의도된 용도 또는 목적을 위한 렌티바이러스-유사 입자 또는 세포의 제조, 저장, 투여, 전달, 유효성, 선택성, 또는 임의의 다른 특징을 돕거나 용

이하게 하는 화합물, 조성물, 물질 또는 구조를 의미한다. 예를 들어, 담체는 활성 성분의 임의의 분해를 최소화하고 대상체에서 임의의 부작용을 최소화하도록 선택될 수 있다. 이러한 약학적으로 허용되는 담체는 염수, 완충된 염수, 인공 뇌척수액, 덱스트로스, 및 물을 포함하나 이에 제한되지 않는 멸균 생체적합성 약학적 담체를 포함한다. 약학적으로 허용되는 것은 생물학적으로 또는 달리 바람직하지 않은 물질을 의미하며, 이는 허용되지 않는 생물학적 효과를 야기하거나 함유된 약학적 조성물의 다른 성분과 유해한 방식으로 상호작용하지 않으면서 선택된 제제와 함께 개인에게 투여될 수 있다.

[0137] **구현예**

[0138] **[0151]** 1. 비-바이러스 핵산 서열에 작동 가능하게 연결된 진핵 프로모터를 포함하는 포유동물 발현 플라스미드로서, 상기 비-바이러스 핵산 서열이 (i) 융합 단백질을 인코딩하는 핵산 서열로서, 상기 융합 단백질이 (a) DNA 중합효소 도메인을 포함하는 폴리펩티드; 및 (b) CRISPR-관련 엔도뉴클레아제 코딩 서열을 포함하는, 핵산 서열; 및 (ii) 가이드 RNA(gRNA) 코딩 서열로서, 상기 gRNA 코딩 서열이 적어도 하나의 앱타머 코딩 서열을 포함하는, 가이드 RNA(gRNA) 코딩 서열을 포함하는, 포유동물 발현 플라스미드.

[0139] **[0152]** 2. 구현예 1에 있어서, CRISPR-관련 엔도뉴클레아제 코딩 서열이 Cas9 단백질을 인코딩하는 포유동물 발현 플라스미드.

[0140] **[0153]** 3. 구현예 1 또는 구현예 2에 있어서, DNA 중합효소 도메인을 포함하는 폴리펩티드가 E. 콜리 DNA 중합효소 I(DNA PolI)인 포유동물 발현 플라스미드.

[0141] **[0154]** 4. 구현예 1 또는 구현예 2에 있어서, DNA 중합효소 도메인을 포함하는 폴리펩티드가 DNA PolI의 클레노우 단편인 포유동물 발현 플라스미드.

[0142] **[0155]** 5. 구현예 1 내지 구현예 4 중 어느 한 구현예에 있어서, DNA 중합효소 도메인을 포함하는 폴리펩티드가 감소된 중합효소 활성을 갖는 포유동물 발현 플라스미드.

[0143] **[0156]** 6. 구현예 1 내지 구현예 5 중 어느 한 구현예에 있어서, 적어도 하나의 앱타머 코딩 서열이 MS2 코트 단백질, PP7 코트 단백질, 람다 N RNA-결합 도메인, 또는 Com 단백질로 구성된 군으로부터 선택된 ABP에 의해 특이적으로 결합된 앱타머 서열을 인코딩하는 포유동물 발현 플라스미드.

[0144] **[0157]** 7. 구현예 1 내지 구현예 6 중 어느 한 구현예에 있어서, 앱타머가 MS2 앱타머 서열 또는 com 앱타머 서열인 포유동물 발현 플라스미드.

[0145] **[0158]** 8. 구현예 1 내지 구현예 7 중 어느 한 구현예에 있어서, sgRNA 코딩 서열이 sgRNA 코딩 서열의 테트라루프 또는 ST2 루프에 삽입된 적어도 하나의 앱타머를 포함하는 포유동물 발현 플라스미드.

[0146] **[0159]** 9. 구현예 8에 있어서, sgRNA 코딩 서열이 gRNA 코딩 서열의 ST2 루프에 삽입된 적어도 하나의 com 앱타머를 포함하는 포유동물 발현 플라스미드.

[0147] **[0160]** 10. a) Gag 뉴클레오티드 서열에 작동 가능하게 연결된 진핵 프로모터를 포함하는 패키징 플라스미드로서, 상기 Gag 뉴클레오티드 서열이 뉴클레오캡시드(NC) 코딩 서열 및 매트릭스 단백질(MA) 코딩 서열을 포함하고, 상기 NC 코딩 서열 또는 MA 코딩 서열 중 하나 또는 둘 모두가 적어도 하나의 비-바이러스 앱타머-결합 단백질(ABP) 뉴클레오티드 서열을 포함하고, 상기 패키징 플라스미드가 기능성 인테그라제 단백질을 인코딩하지 않는, 패키징 플라스미드; b) 구현예 1 내지 구현예 9 중 어느 한 구현예의 적어도 하나의 포유동물 발현 플라스미드; 및 c) 외피 당단백질 코딩 서열을 포함하는 외피 플라스미드를 포함하는 렌티바이러스 패키징 시스템.

[0148] **[0161]** 11. 구현예 10에 있어서, 패키징 플라스미드가 Rev 뉴클레오티드 서열 및 Tat 뉴클레오티드 서열을 추가로 포함하는 렌티바이러스 패키징 시스템.

[0149] **[0162]** 12. 구현예 10 또는 구현예 0에 있어서, Rev 뉴클레오티드 서열을 포함하는 제2 패키징 플라스미드를 추가로 포함하는 렌티바이러스 패키징 시스템.

[0150] **[0163]** 13. 구현예 10 내지 구현예 12 중 어느 한 구현예에 있어서, 적어도 하나의 비-바이러스 ABP 뉴클레오티드 서열이 MS2 코트 단백질, PP7 코트 단백질, 람다 N 펩티드, 또는 Com 단백질을 인코딩하는 렌티바이러스 패키징 시스템.

[0151] **[0164]** 14. A) 뉴클레오캡시드(NC) 단백질 또는 매트릭스(MA) 단백질을 포함하는 융합 단백질로서, 상기 NC 단백질 또는 MA 단백질이 적어도 하나의 비-바이러스 앱타머 결합 단백질(ABP)을 포함하는, 융합 단백질; 및 B)

(i) 융합 단백질을 인코딩하는 핵산 서열로서, 상기 융합 단백질이 (a) DNA 중합효소 도메인을 포함하는 폴리펩티드, 및 (b) CRISPR-관련 엔도뉴클레아제 코딩 서열을 포함하는, 핵산 서열; 및 (ii) 가이드 RNA(gRNA) 코딩 서열로서, 상기 gRNA 코딩 서열이 적어도 하나의 앱타머 코딩 서열을 포함하고, 여기서 렌티바이러스-유사 입자가 기능성 인테그라제 단백질을 포함하지 않는, 가이드 RNA(gRNA) 코딩 서열을 포함하는, 리보뉴클레오티드 단백질(RNP) 복합체를 포함하는 렌티바이러스 입자.

[0152] **[0165]** 15. 구현예 14에 있어서, CRISPR-관련 엔도뉴클레아제 코딩 서열이 Cas9 단백질을 인코딩하는 렌티바이러스 입자.

[0153] **[0166]** 16. 구현예 14 또는 구현예 15에 있어서, DNA 중합효소 도메인을 포함하는 폴리펩티드가 E. 콜리 DNA 중합효소 I(DNA PolI)인 렌티바이러스 입자.

[0154] **[0167]** 17. 구현예 14 또는 구현예 15에 있어서, DNA 중합효소 도메인을 포함하는 폴리펩티드가 DNA PolI의 클레노우(Klenow) 단편인 렌티바이러스 입자.

[0155] **[0168]** 18. 구현예 14 내지 구현예 17 중 어느 한 구현예에 있어서, DNA 중합효소 도메인을 포함하는 폴리펩티드가 감소된 중합효소 활성을 갖는 렌티바이러스 입자.

[0156] **[0169]** 19. 렌티바이러스 입자를 생산하는 방법으로서, 상기 방법이 a) 복수의 진핵 세포를 구현예 10 내지 구현예 13 중 어느 한 구현예의 시스템의 패키징 플라스미드, 적어도 하나의 포유동물 발현 플라스미드, 및 외피 플라스미드로 형질감염시키고; b) 렌티바이러스 입자가 생산되기에 충분한 시간 동안 형질감염된 진핵 세포를 배양하는 것을 포함하는, 방법.

[0157] **[0170]** 20. 구현예 19에 있어서, 렌티바이러스 입자가 (i) 융합 단백질을 인코딩하는 핵산 서열로서, 상기 융합 단백질이 (a) DNA 중합효소 도메인을 포함하는 폴리펩티드 및 (b) CRISPR-관련 엔도뉴클레아제 코딩 서열을 포함하는, 핵산 서열; 및 (ii) 가이드 RNA를 포함하는 리보뉴클레오티드 단백질(RNP) 복합체를 포함하는, 방법.

[0158] **[0171]** 21. 구현예 20에 있어서, 복수의 진핵 세포가 포유동물 세포인 방법.

[0159] **[0172]** 22. 구현예 19 내지 구현예 21 중 어느 한 구현예의 방법에 의해 제조된 렌티바이러스 입자.

[0160] **[0173]** 23. 세포에서 유전체 표적 서열을 변형시키는 방법으로서, 상기 방법이 복수의 바이러스 입자로 복수의 진핵 세포를 형질도입시키는 것을 포함하고, 상기 복수의 바이러스 입자가 구현예 14 내지 구현예 18 중 어느 한 구현예에 따른 렌티바이러스-유사 입자를 포함하고, 여기서 RNP 복합체가 세포의 유전체 DNA 내의 유전체 표적 서열에 결합하고, 여기서 CRISPR-관련 엔도뉴클레아제가 유전체 표적 서열을 절단하여 이중-가닥 파손을 생성함으로써 유전체 표적 서열을 변형시키는, 방법.

[0161] **[0174]** 24. 구현예 23에 있어서, 비-상동성 말단 연결(NHEJ)이 DNA 중합효소 도메인에 융합되지 않은 CRISPR-관련 엔도뉴클레아제에 의해 변형된 세포와 비교하여 세포에서 증가하는 방법.

[0162] **[0175]** 25. 구현예 23에 있어서, NHEJ 대 비-NHEJ의 비가 세포에서 증가하는 방법.

[0163] **[0176]** 26. 구현예 25에 있어서, 비-NHEJ 말단 연결이 미세상동성-매개 말단 연결(MMEJ) 및/또는 단일-가닥 어닐링(SSA)인 방법.

[0164] **[0177]** 27. 구현예 23 내지 구현예 26 중 어느 한 구현예에 있어서, 표적-내 결실의 수가 세포에서 감소되는 방법.

[0165] **[0178]** 28. 구현예 27에 있어서, 500개 염기쌍 초과의 크기의 표적-내 결실의 수가 세포에서 감소하는 방법.

[0166] **[0179]** 29. 구현예 23 내지 구현예 27 중 어느 한 구현예에 있어서, 표적-내 1 염기쌍(1-bp) 결실 대 1-bp 초과의 표적-내 결실의 비가 세포에서 증가하는 방법.

[0167] **[0180]** 30. 구현예 29에 있어서, 표적-내 1 염기쌍(1-bp) 결실 대 500개 염기쌍 초과의 결실의 비가 세포에서 증가하는 방법.

[0168] **[0181]** 31. 구현예 23 내지 구현예 30 중 어느 한 구현예에 있어서, 주형 삽입(TIS)의 수가 세포에서 증가하는 방법.

[0169] **[0182]** 32. 구현예 31에 있어서, TIS 대 비-TIS의 비가 세포에서 증가하는 방법.

[0170] **[0183]** 34. 구현예 23 내지 구현예 32 중 어느 한 구현예에 있어서, 복수의 진핵 세포가 포유동물 세포인 방법.

[0171] **[0184]** 35. 구현예 23 내지 구현예 33 중 어느 한 구현예에 있어서, 복수의 진핵 세포가 대상체에 존재하는 세포인 방법.

[0172] **[0185]** 36. 구현예 34에 있어서, 대상체가 인간 대상체인 방법.

[0173] **[0186]** 37. 구현예 35에 있어서, 대상체가 복수의 바이러스 입자로 주사되는 방법.

[0174] **[0187]** 38. 구현예 1 내지 구현예 9 중 어느 한 구현예의 플라스미드를 함유하는 세포.

[0175] **[0188]** 39. 구현예 10 내지 구현예 13 중 어느 한 구현예의 렌티바이러스 패키징 시스템을 함유하는 세포.

[0176] **[0189]** 40. 구현예 14 내지 구현예 16 중 어느 한 구현예의 렌티바이러스 입자를 함유하는 세포.

[0177] **[0190]** 41. 구현예 23 내지 구현예 36 중 어느 한 구현예의 방법을 사용하여 변형된 세포.

[0178] **[0191]** 42. a) 대상체로부터 세포를 획득하고; b) 구현예 23 내지 구현예 36 중 어느 한 구현예의 방법을 사용하여 대상체의 세포를 변형시키고; c) 변형된 세포를 대상체에게 투여하는 것을 포함하는, 대상체에서 질환을 치료하기 위한 방법.

[0179] **[0192]** 43. 구현예 41에 있어서, 질환이 암인 방법.

[0180] **[0193]** 44. 구현예 42에 있어서, 질환이 뒤시엔느 근이영양증인 방법.

[0181] **[0194]** 45. 구현예 41 내지 구현예 43 중 어느 한 구현예에 있어서, 세포가 T 세포인 방법.

[0182] **[0195]** 본 발명의 개시에서 논의된 모든 특허, 특허 간행물, 특허 출원, 저널 기사, 서적, 기술 참고문헌 등은 모든 목적을 위해 그 전체가 참조로서 본원에 포함된다.

[0183] **[0196]** 본 발명의 개시의 도면 및 설명은 본 발명의 개시의 명확한 이해와 관련된 요소를 예시하기 위해 단순화된 것으로 이해되어야 한다. 도면은 구성 도면이 아니라 예시의 목적으로 제공되는 것으로 이해되어야 한다. 생략된 세부사항 및 수정 또는 대안적인 구현예는 당업자의 범위 내에 있다.

[0184] **[0197]** 본 발명의 개시의 특정 양태에서, 단일 성분은 다수의 성분으로 대체되고, 다수의 성분은 단일 성분으로 대체되어, 요소 또는 구조를 제공하거나 주어진 기능 또는 기능들을 수행할 수 있다는 것이 이해될 수 있다. 이러한 치환이 본 발명의 개시의 특정 구현예를 실시하는데 작동하지 않는 경우를 제외하고, 이러한 치환은 본 발명의 개시의 범위 내에 있는 것으로 간주된다.

[0185] **[0198]** 본원에 제시된 예는 본 발명의 개시의 잠재적이고 특정한 구현을 예시하기 위한 것이다. 실시예는 주로 당업자를 위한 본 발명의 개시의 예시의 목적으로 의도된 것으로 이해될 수 있다. 본 발명의 개시의 사상을 벗어나지 않으면서 본원에 설명된 이러한 도면 또는 작업에 대한 변형이 있을 수 있다. 예를 들어, 특정 경우에, 방법 단계 또는 작업은 상이한 순서로 수행되거나 실행될 수 있거나, 작업이 추가, 삭제 또는 수정될 수 있다.

[0186] **[0199]** 값의 범위가 제공되는 경우, 문맥에서 명백히 달리 지시하지 않는 한, 상기 범위의 상한과 하한 사이의 하한 단위의 가장 작은 분율까지의 각각의 개재 값이 또한 구체적으로 개시되는 것으로 이해된다. 언급된 범위의 임의의 언급된 값 또는 언급되지 않은 개재 값과 해당 언급된 범위의 임의의 다른 언급된 또는 개재 값 사이의 임의의 더 작은 범위가 포함된다. 이러한 더 작은 범위의 상한 및 하한은 독립적으로 범위에 포함되거나 배제될 수 있고, 어느 하나, 둘 모두가 아닌 또는 둘 모두의 한계가 더 작은 범위에 포함되는 각각의 범위는 또한 언급된 범위 내의 임의의 구체적으로 배제된 제한을 조건으로 하여 기술 내에 포함된다. 언급된 범위가 한계 중 하나 또는 둘 모두를 포함하는 경우, 포함된 한계 중 어느 하나 또는 둘 모두를 제외한 범위가 또한 포함된다.

[0187] **[0200]** 도면에 도시되거나 상기 설명된 성분의 상이한 배열 뿐만 아니라 제시되지 않거나 설명되지 않은 성분 및 단계가 가능하다. 유사하게, 일부 특징 및 하위조합이 유용하고, 다른 특징 및 하위조합을 참조하지 않고 사용될 수 있다. 본 발명의 개시의 구현예는 제한적인 것이 아니라 예시 목적을 위해 설명되었으며, 대안적인 구현예는 본 특허의 독자에게 명백해질 것이다. 따라서, 본 발명의 개시는 상기 설명되거나 도면에 도시된 구현예로 제한되지 않으며, 하기 청구범위의 범위를 벗어나지 않으면서 다양한 구현예 및 수정이 이루어질 수 있다.

[0188] **실시예**

[0189] **[0201]** 에스케리키아 콜리(*Escherichia. coli*) DNA 중합효소 I(pol I)은 DNA 복구 및 래깅(lagging)-가닥 염색체 DNA의 복제에서 기능한다. 이의 작은 N-말단 도메인은 5'-3' 엑소뉴클레아제 활성을 함유하고, 전장 중합효소의 부분 분해에 의해 생성될 수 있는 큰 C-말단 클레노우 단편은 중합효소 및 3'-5' 엑소뉴클레아제 활성을

보유한다. MMEJ 및 SSA는 큰 DNA 결실을 생성할 수 있고 둘 모두 DNA 절제에 의존하기 때문에, DNA 중합효소 I로 DNA 절제에 대응하는 것이 MMEJ 및 SSA보다 표준 NHEJ를 선호할 수 있고, 큰 결실의 생성을 감소시킬 수 있는 것으로 가정되었다. 따라서, E. 콜리 pol I은 DNA 절제에 대응하고 큰 결실을 생성할 기회를 감소시키기 위해 DSM을 표적으로 하였다(도 1a). CRISPR/Cas9에 의해 생성된 5' 오버행이 채워지고 TIS의 비율(도 1b)이 증가하여 INDEL을 추가로 정제하는지를 결정하기 위한 연구가 또한 수행되었다.

[0190] **A. 재료 및 방법**

[0191] **[0202] 작제물.** 이 연구에 사용된 작제물은 표 2에 설명되어 있다. 렌티바이러스 슈도타이핑을 위한 외피 플라스미드 pMD2.G는 Addgene(Addgene 12259, Watertown, MA)에서 구입하였다. 이 연구에 사용된 플라스미드 중 일부는 Addgene으로부터 입수 가능하다(플라스미드 ID: 176234, 76235, 176236, 176237, 176238). 다른 것들은 요청 시 저자로부터 입수 가능하다. 이 연구에 사용된 프라이머에 대한 서열은 표 3에 나열되어 있다. 단일 가이드 RNA(sgRNA) 서열은 표 4에 열거되어 있다.

[0192] **표 2. 작제물**

[0193]

번호	명칭	목적	생성 전략
1	pspCas9-POLI-테트라-com 벡터	spCas9-POLI 융합 단백질 및 sgRNA 를 발현시키기 위한 벡터 플라스미드. 발현을 향상시키기 위해 spCas9-POLI 의 3'UTR 에 HBB 3'UTR 의 2 개 복사체가 존재한다. sgRNA 스캐폴드의 테트라루프를 com-앱타머로 대체하여 com-COM 상호작용을 통해 바이러스 캡시드로의 Cas9 RNP 캡슐화를 가능하게 하였다.	E. 콜리 DNA POLI(POLA 에 의해 인코딩됨)의 DNA 서열을 "GGATCCTCTGAGACACCAGGCACCTCCGAGTCTGCCA CCCCTGAGGGAGGCAGCGGAGGCAGCGGCTCCGTGCA GATCCCACAGAACCCCCTGATCCTGGTGGACGGCAGCT CCTACCTGTATCGGGCCTACCACGCCTTCCCACCTCTG ACAAACTCCGCCGGAGAGCCAACCGGAGCCATGTATG GCGTGCTGAATATGCTGAGGAGCCTGATCATGCAGTAC AAGCCTACACACGCCGCCGTGGTGTTTGATGCCAAGGG CAAGACCTTCCGCGACGAGCTGTTTGAGCACTACAAGA GCCACAGGCCACCAATGCCTGACGATCTGAGGGCACA GATCGAGCCACTGCACGCAATGGTGAAGGCCATGGGC CTGCCTCTGCTGGCCGTGAGCGGAGTGGAGGCCGACG ATGTGATCGGCACACTGGCAAGGGAGGCAGAGAAGGC AGGCCGCCCAGTGCTGATCTCCACCGGCGACAAGGAT ATGGCCCAGCTGGTGACACCAAACATCACCCTGATCAA CACCATGACAAATACCATCCTGGGCCCCGAGGAGGTG GTGAATAAGTATGGCGTGCCTCCAGAGCTGATCATCGA TTTCCTGGCCCTGATGGGCGACTCTAGCGATAACATCC CTGGAGTGCCAGGAGTGGGAGAAAAGACCGCACAGGC CCTGCTGCAGGGCCTGGGAGGCCTGGACACCCTGTACG CCGAGCCAGAGAAGATCGCCGGCCTGTCCTTTAGGGGC GCCAAGACAATGGCCGCCAAGCTGGAGCAGAATAAGG AGGTGGCCTACCTGTCTTATCAGCTGGCCACAATCAAG ACCGACGTGGAGCTGGAGCTGACCTGCGAGCAGCTGG AGGTGCAGCAGCCTGCAGCAGAGGAGCTGCTGGGCCT GTTCAAGAAGTACGAGTTTAAGAGATGGACAGCCGAT GTGGAGGCCGGCAAGTGGCTGCAGGCAAAGGGAGCAA AGCCAGCAGCAAAGCCACAGGAGACAAGCGTGGCAGA CGAGGCACCAGAGGTGACAGCCACCGTGATCTCCTAC GATAACTATGTGACAATCCTGGACGAGGAGACACTGA AGGCCTGGATCGCCAAGCTGGAGAAGGCCCCCGTGTTC GCCTTTGATACAGAGACAGACAGCCTGGATAACATCTC CGCCAATCTGGTGGGCCTGTCTTTCGCCATCGAGCCTG GCGTGGCCGCCTATATCCCAGTGGCCCACGACTACCTG GATGCCCCCGACCAGATCAGCAGGGAGAGAGCCCTGG AGCTGCTGAAGCCTCTGCTGGAGGATGAGAAGGCCCT GAAGGTCGGCCAGAACCTGAAGTATGACAGGGGCATC CTGGCCAATTACGGCATCGAGCTGAGAGGCATCGCCTT TGACACCATGCTGGAGTCTTATATCCTGAATAGCGTGG CAGGCCGGCACGACATGGATTCCCTGGCCGAGAGGTG GCTGAAGCACAAGACAATCACCTTCGAGGAGATCGCC GGCAAGGGCAAGAACCAGCTGACCTTCAACCAGATCG CCCTGGAGGAGGCAGGCAGGTACGCAGCAGAGGACGC AGATGTGACCCTGCAGCTGCACCTGAAGATGTGGCCAG ATCTGCAGAAGCACAAGGGCCCCCTGAACGTGTTCGA GAATATCGAGATGCCCCTGGTGCCTGTGCTGAGCCGGA TCGAGCGCAACGGCGTGAAGATCGACCCTAAGGTGCT GCACAATCACTCCGAGGAGCTGACCCTGAGACTGGCC GAGCTGGAGAAGAAGGCCCACGAGATCGCCGGCGAGG AGTTCAACCTGTCCTCTACAAAGCAGCTGCAGACCATC CTGTTTGAGAAGCAGGGCATCAAGCCCCTGAAGAAAA CCCCTGGAGGAGCACCATCTACCAGCGAGGAGGTGCT GGAGGAGCTGGCCCTGGATTATCCCCTGCCTAAAGTGA

번호	명칭	목적	생성 전략
			TCCTGGAGTACCGGGGCCTGGCCAAGCTGAAGTCTACA TATACCGACAAGCTGCCCCTGATGATCAACCCTAAGAC AGGAAGGGTGCACACCAGCTACCACCAGGCAGTGACA GCAACCGGCCGCCTGAGCTCCACCGATCCAAACCTGCA GAATATCCCCGTGAGGAATGAGGAGGGCAGGAGAATC AGACAGGCCTTCATCGCCCCCGAGGATTATGTGATCGT GTCCGCCGACTACTCTCAGATCGAGCTGAGGATCATGG CCCACCTGTCCAGAGATAAGGGCCTGCTGACAGCCTTC GCCGAGGGCAAGGACATCCACAGGGCAACCGCAGCAG AGGTGTTTGGCCTGCCTCTGGAGACAGTGACCTCCGAG CAGCGGCGCTCTGCCAAGGCCATCAACTTCGGCCTGAT CTATGGCATGTCTGCCTTTGGCCTGGCCAGGCAGCTGA ATATCCCTAGAAAGGAGGCCCAGAAGTACATGGACCT GTATTTCGAGAGGTACCCAGGCGTGCTGGAGTACATGG AGAGGACAAGGGCACAGGCAAAGGAGCAGGGCTATGT GGAGACACTGGATGGCAGGAGACTGTACCTGCCAGAC ATCAAGTCTAGCAACGGAGCAAGGAGGGCAGCAGCAG AGAGGGCCGCCATCAATGCCCCCATGCAGGGCACAGC CGCCGATATCATCAAGAGAGCCATGATCGCAGTGGAC GCATGGCTGCAGGCAGAGCAGCCAAGGGTGAGAATGA TCATGCAGGTGCACGATGAGCTGGTGTTTGAGGTGCAC AAGGACGATGTGGACGCCGTGGCCAAGCAGATCCACC AGCTGATGGAGAACTGTACCAGGCTGGATGTGCCACTG CTGGTGGAAGTGGGCAGCGGAGAGAATTGGGACCAGG CCCACTGAGGATCC"(SEQ ID NO: 36)로서 인간 및 마우스 발현에 대해 최적화되었으며, pspCas9-테트라-com 벡터의 BamH1 부위에 삽입되었다. 폴리펩티드 서열은 (SEQ ID NO: 37)이다: *MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPA* *ADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI* *KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS* *NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY* *HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE* *GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS* *ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF* *DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS* *DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ* *QLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM* *DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR* *QEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR* *KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP* *KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDL* *LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY* *HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLK* *TYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI* *LDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL* *HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIE* *MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT* *QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQS* *FLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL* *LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITK* *HVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ* *FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY* *KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE*

[0194]

번호	명칭	목적	생성 전략
			IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKT EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE LALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLY ETRIDLSQLGGDKRPAATKKAGQAKKKK<u>GSSETPGTSESAT PEGGSGGSGS</u>**VQIPQNPLILVDGSSYLYRAYHAFPPLTN SAGEPTGAMYGVLNMLRSLIMQYKPTHAAVVFDAKG KTFRDELFEHYKSHRPPMPDDLRAQIEPLHAMVKAM GLPLLAVSGVEADDVIGTLAREAEKAGRPVLISTGDKD MAQLVTPNITLINTMTNTILGPEEVVNKYGVPPELIIDF LALMGDSSDNIPGVPGVGEKTAQALLQGLGGLDTLYA EPEKIAGLSFRGAKTMAAKLEQNKEVAYLSYQLATIK TDVELELTCEQLEVQQPAAEELLGLFKKYEFKRWTA DVEAGKWLQAKGAKPAAKPQETSVADEAPEVTATVIS YDNYVTILDEETLKAWIAKLEKAPVFAFDTETDSLDNI SANLVGLSFAIEPGVAAYIPVAHDYLDAPDQISRERALE LLKPLLEDEKALKVGQNLKYDRGILANYGIELRGIAFD TMLESYILNSVAGRHDMDSLAERWLKHKTITFEEIAG KGKNQLTFNQIALEEAGRYAAEDADVTLQLHLKMWP DLQKHKGPLNVFENIEMPLVPVLSRIERNGVKIDPKVL HNHSEELTLRLAELEKKAHEIAGEEFNLSSTKQLQTIL FEKQGIKPLKKTPGGAPSTSEEVLEELALDYPLPKVIL EYRGLAKLKSTYTDKLPLMINPKTGRVHTSYHQAVTA TGRLSSTDPNLQNIPVRNEEGRRIRQAFIAPEDYVIVSA DYSQIELRIMAHLSRDKGLLTAFAEGKDIHRATAAEVF GLPLETVTSEQRRSAKAINFGLIYGMSAFGLARQLNIP RKEAQKYMDLYFERYPGVLEYMERTRAQAKEQGYV ETLDGRRLYLPDIKSSNGARRAAAERAAINAPMQGTA ADIIKRAMIAVDAWLQAEQPRVRMIMQVHDELVFEV HKDDVDAVAKQIHQLMENCTRLDVPLLVEVGSGENW DQAH* (키 : *Cas9* (SEQ ID NO: 29)-링커 (SEQ ID NO: 32)-**중합효소 (SEQ ID NO: 27)**
2	pspCas9-poli-테트라-com-clcn5-sp-g1	인간 CLCN5 유전자를 표적화하는 spCas9-polI 융합 단백질 및 sgRNA 를 발현시키기 위한 벡터 플라스미드.	hCLCN5-sp-g1-F (ACCGGAGGACAAGTCGTACAATGG) (SEQ ID NO: 38) 및 hCLCN5-sp-g1-R (AAACCCATTGTACGACTTGTCCTC) (SEQ ID NO: 39)의 어닐링된 생성물은 T4 DNA 리가제에 의해 pspCas9-POLI-테트라-com 벡터의 BbsI 부위 사이에 삽입되었다.
3	pspCas9-클레노우-테트라-com-CLCN5-sp-g1	인간 CLCN5 유전자를 표적화하는 spCas9-클레노우 융합 단백질 및 sgRNA 를	pspCas9-POLI-테트라-com-hCLCN5-sp-g1 을 주형으로 사용하고, 클레노우-F(GTCTGCCACCCCTGAGGGAGGAGCACCAGAGGTGAC AGCCAC)(SEQ ID NO: 40) 및 클레노우-R(TTGTGCAGCACCTTAGGGTCG)(SEQ ID NO: 41)을 735 bp 단편을 증폭시키기 위한 프라이머로 사용한다. 이

[0195]

번호	명칭	목적	생성 전략
		발현시키기 위한 플라스미드.	DNA 를 주입 반응에 의해 pspCas9-POLI-테트라-com-hCLCN5-sp-g1 의 2 개의 Bsu36I 부위 사이에 삽입하였다. 최종 결과는 E. 콜리 POLI 의 5' 엑소뉴클레아제 활성을 코딩하는 DNA 의 제거이다. 핵산 서열은 (SEQ ID NO: 42)이다: ATGGACTATAAGGACCACGACGGAGACTACAAGGATC ATGATATTGATTACAAAGACGATGACGATAAGATGGC CCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTC CCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACA TCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGAC GAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGG GCAACACCGACCGGCACAGCATCAAGAAGAACCTGAT CGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAG GCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACA CCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGAT CTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCT TCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGAT AAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCG TGGACGAGGTGGCCTACCACGAGAAGTACCCCACCAT CTACCACCTGAGAAAGAAACTGGTGGACAGCACCGAC AAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCA CATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCG ACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTT CATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGG AAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGC CATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTG GAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGA ATGGCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGC CTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGA GGATGCCAAACTGCAGCTGAGCAAGGACACCTACGAC GACGACCTGGACAACCTGCTGGCCCAGATCGGCGACC AGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCC GACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACAC CGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCA AGAGATACGACGAGCACCACCAGGACCTGACCCTGCT GAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTAC AAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACG CCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTT CTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACG GCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGA CCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGC ATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCAT TCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGG ACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCG CATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACA GCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAAC CATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAG GGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAA CTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCA AGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAAC GAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGA GAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGC CATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGA CCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAAT CGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAG ATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTG

[0196]

[0197]

번호	명칭	목적	생성 전략
			CTGAAAATTATCAAGGACAAGGACTTCCTGGACAATG AGGAAAACGAGGACATTCTGGAAGATATCGTGCTGAC CCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAA CGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGT GATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGG GGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGG ACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAG TCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGAT CCACGACGACAGCCTGACCTTTAAAGAGGACATCCAG AAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACG AGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAG AAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGC TCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACAT CGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAG AAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGG ATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCC TGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAA CGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGG GATATGTACGTGGACCAGGAACTGGACATCAACCGGC TGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGC TTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGAC CAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTG CCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACT GGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAG AAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGC CTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGAC AGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGC ACAGATCCTGGACTCCCGGATGAACACTAAGTACGAC GAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCA CCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGAT TTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCA CCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAA CCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGA GTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGA AGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGC TACCGCCAAGTACTTCTTCTACAGCAACATCATGAACT TTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATC CGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCG GGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCAC CGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATC GTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCA AAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCT GATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTAC GGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCT GGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAA CTGAAGAGTGTGAAAAGAGCTGCTGGGGATCACCATCA TGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTT CTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACC TGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTG GAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCG AACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAA ATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGA AGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACA GCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAG ATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGAT CCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCT ACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGC

[0198]

번호	명칭	목적	생성 전략
			CGAGAATATCATCCACCTGTTTACCCTGACCAATCTGG GAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATC GACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGG ACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTAC GAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACA AAAGGCCGGCGGCCACGAAAAAGGCCGGccaggcaaaaaag aaaaaggGATCCTCTGAGACACCAGGCACCTCCGAGTCTG CCACCCCTGAGGGAGGAGCACCAGAGGTGACAGCCAC CGTGATCTCCTACGATAACTATGTGACAATCCTGGACG AGGAGACACTGAAGGCCTGGATCGCCAAGCTGGAGAA GGCCCCCGTGTTCGCCTTTGATACAGAGACAGACAGCC TGGATAACATCTCCGCCAATCTGGTGGGCCTGTCTTTC GCCATCGAGCCTGGCGTGGCCGCCTATATCCCAGTGGC CCACGACTACCTGGATGCCCCCGACCAGATCAGCAGG GAGAGAGCCCTGGAGCTGCTGAAGCCTCTGCTGGAGG ATGAGAAGGCCCTGAAGGTCGGCCAGAACCTGAAGTA TGACAGGGGCATCCTGGCCAATTACGGCATCGAGCTGA GAGGCATCGCCTTTGACACCATGCTGGAGTCTTATATC CTGAATAGCGTGGCAGGCCGGCACGACATGGATTCCCT GGCCGAGAGGTGGCTGAAGCACAAGACAATCACCTTC GAGGAGATCGCCGGCAAGGGCAAGAACCAGCTGACCT TCAACCAGATCGCCCTGGAGGAGGCAGGCAGGTACGC AGCAGAGGACGCAGATGTGACCCTGCAGCTGCACCTG AAGATGTGGCCAGATCTGCAGAAGCACAAGGGCCCCC TGAACGTCTTTGAGAATATCGAGATGCCCCTGGTGCCT GTGCTGAGCCGGATCGAGCGCAACGGCGTGAAGATCG ACCCTAAGGTGCTGCACAATCACTCCGAGGAGCTGACC CTGAGACTGGCCGAGCTGGAGAAGAAGGCCCACGAGA TCGCCGGCGAGGAGTTCAACCTGTCCTCTACAAAGCAG CTGCAGACCATCCTGTTTGAGAAGCAGGGCATCAAGCC CCTGAAGAAAACCCCTGGAGGAGCACCATCTACCAGC GAGGAGGTGCTGGAGGAGCTGGCCCTGGATTATCCCCT GCCTAAAGTGATCCTGGAGTACCGGGGCCTGGCCAAG CTGAAGTCTACATATACCGACAAGCTGCCCCTGATGAT CAACCCTAAGACAGGCCGGGTGCACACCAGCTACCAC CAGGCAGTGACAGCAACCGGCCGCCTGAGCTCCACCG ATCCAAACCTGCAGAATATCCCCGTGAGGAATGAGGA GGGCAGGAGAATCAGACAGGCCTTCATCGCCCCCGAG GATTATGTGATCGTGTCCGCCGACTACTCTCAGATCGA GCTGAGGATCATGGCCCACCTGTCCAGAGATAAGGGC CTGCTGACAGCCTTCGCCGAGGGCAAGGACATCCACA GGGCAACCGCAGCAGAGGTGTTTGGCCTGCCTCTGGAG ACAGTGACCTCCGAGCAGCGGCGCTCTGCCAAGGCCAT CAACTTCGGCCTGATCTATGGCATGTCTGCCTTTGGCCT GGCCAGGCAGCTGAATATCCCTAGAAAGGAGGCCCAG AAGTACATGGACCTGTATTTCGAGAGATACCCCGGCGT GCTGGAGTACATGGAGAGGACACGCGCACAGGCAAAG GAGCAGGGCTATGTGGAGACACTGGATGGCAGGAGAC TGTACCTGCCAGACATCAAGTCTAGCAACGGAGCAAG GAGGGCAGCAGCAGAGAGGGCCGCCATCAATGCCCCC ATGCAGGGCACAGCCGCCGATATCATCAAGAGAGCCA TGATCGCAGTGGACGCCTGGCTGCAGGCAGAGCAGCC AAGGGTGAGAATGATCATGCAGGTGCACGATGAGCTG GTGTTTGAGGTGCACAAGGACGATGTGGACGCCGTGG CCAAGCAGATCCACCAGCTGATGGAGAACTGTACCCG CCTGGATGTGCCACTGCTGGTGGAAGTGGGAAGCGGA GAGAATTGGGACCAGGCCCACTGAG.

번호	명칭	목적	생성 전략
			폴리펩티드 서열은 SEQ ID NO: 43 이다: *MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPA ADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ QLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR QEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDL LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLK TYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI LDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIE MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQS FLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITK HVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKT EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE LALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLY ETRIDLSQLGGDKRPAATKKAGQAKKKK*GSSETPGTSESAT PEGG**APEVTATVISYDNYVTILDEETLKAWIAKLEKAP VFAFDTETDSLDNISANLVGLSFAIEPGVAAYIPVAHDY LDAPDQISRERALELLKPLLEDEKALKVGQNLKYDRG ILANYGIELRGIAFDTMLESYILNSVAGRHDMDSLAER WLKHKTITFEEIAGKGKNQLTFNQIALEEAGRYAAED ADVTLQLHLKMWPDLQKHKGPLNVFENIEMPLVPVL SRIERNGVKIDPKVLHNHSEELTLRLAELEKKAHEIAG EEFNLSSTKQLQTILFEKQGIKPLKKTPGGAPSTSEEVL EELALDYPLPKVILEYRGLAKLKSTYTDKLPLMINPKT GRVHTSYHQAVTATGRLSSTDPNLQNIPVRNEEGRRIR QAFIAPEDYVIVSADYSQIELRIMAHLSRDKGLLTAFAE GKDIHRATAAEVFGLPLETVTSEQRRSAKAINFGLIYG MSAFGLARQLNIPRKEAQKYMDLYFERYPGVLEYME RTRAQAKEQGYVETLDGRRLYLPDIKSSNGARRAAAE RAAINAPMQGTAADIIKRAMIAVDAWLQAEQPRVRMI MQVHDELVFEVHKDDVDAVAKQIHQLMENCTRLDVP LLVEVGSGENWDQAH** (키: *Cas9* (SEQ ID NO: 27)-링커 (SEQ ID NO: 32)-**클레노우** (SEQ ID NO: 28)

[0199]

번호	명칭	목적	생성 전략
4	pspCas9-3exo-테트라-com-CLCN5-sp-g1	spCas9 및 POLI 의 3' 엑소뉴클레아제 도메인의 융합 단백질, 및 인간 CLCN5 유전자를 표적화하는 sgRNA 를 발현시키기 위한 플라스미드.	pspCas9-exo-테트라-com-hCLCN5-sp-g1 을 주형으로 사용하고, 650 bp 단편을 증폭시키기 위해 클레노우-F (GTCTGCCACCCCTGAGGGAGGAGCACCAGAGGTGACA GCCAC) (SEQ ID NO: 44) 및 3exo-R (CtaggaattttctagactaGTGCTTCTGCAGATCTGGCCAC) (SEQ ID NO: 45)을 사용한다. 이 단편을 pspCas9-exo-테트라-com-hCLCN5-sp-g1 주입 반응의 Bsu36I 및 XbaI 부위 사이에 삽입하였다. 핵산 서열은 (SEQ ID NO: 46)이다: ATGGACTATAAGGACCACGACGGAGACTACAAGGATC ATGATATTGATTACAAAGACGATGACGATAAGATGGC CCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTC CCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACA TCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGAC GAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGG GCAACACCGACCGGCACAGCATCAAGAAGAACCTGAT CGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAG GCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACA CCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGAT CTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCT TCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGAT AAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCG TGGACGAGGTGGCCTACCACGAGAAGTACCCCACCAT CTACCACCTGAGAAAGAAACTGGTGGACAGCACCGAC AAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCA CATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCG ACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTT CATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGG AAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGC CATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTG GAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGA ATGGCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGC CTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGA GGATGCCAAACTGCAGCTGAGCAAGGACACCTACGAC GACGACCTGGACAACCTGCTGGCCCAGATCGGCGACC AGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCC GACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACAC CGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCA AGAGATACGACGAGCACCACCAGGACCTGACCCTGCT GAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTAC AAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACG CCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTT CTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACG GCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGA CCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGC ATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCAT TCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGG ACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCG CATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACA GCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAAC CATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAG GGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAA CTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCA AGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAAC GAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGA

[0200]

[0201]

번호	명칭	목적	생성 전략
			GAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGC CATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGA CCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAAT CGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAG ATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTG CTGAAAATTATCAAGGACAAGGACTTCCTGGACAATG AGGAAAACGAGGACATTCTGGAAGATATCGTGCTGAC CCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAA CGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGT GATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGG GGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGG ACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAG TCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGAT CCACGACGACAGCCTGACCTTTAAAGAGGACATCCAG AAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACG AGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAG AAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGC TCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACAT CGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAG AAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGG ATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCC TGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAA CGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGG GATATGTACGTGGACCAGGAACTGGACATCAACCGGC TGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGC TTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGAC CAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTG CCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACT GGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAG AAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGC CTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGAC AGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGC ACAGATCCTGGACTCCCGGATGAACACTAAGTACGAC GAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCA CCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGAT TTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCA CCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAA CCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGA GTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGA AGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGC TACCGCCAAGTACTTCTTCTACAGCAACATCATGAACT TTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATC CGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCG GGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCAC CGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATC GTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCA AAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCT GATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTAC GGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCT GGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAA CTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCA TGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTT CTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACC TGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTG GAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCG AACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAA ATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGA

번호	명칭	목적	생성 전략
			AGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACA GCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAG ATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGAT CCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCT ACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGC CGAGAATATCATCCACCTGTTTACCCTGACCAATCTGG GAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATC GACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGG ACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTAC GAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACA AAAGGCCGGCGGCCACGAAAAAGGCCGGccaggcaaaaaag aaaaaggGATCCTCTGAGACACCAGGCACCTCCGAGTCTG CCACCCCTGAGGGAGGAGCACCAGAGGTGACAGCCAC CGTGATCTCCTACGATAACTATGTGACAATCCTGGACG AGGAGACACTGAAGGCCTGGATCGCCAAGCTGGAGAA GGCCCCCGTGTTCGCCTTTGATACAGAGACAGACAGCC TGGATAACATCTCCGCCAATCTGGTGGGCCTGTCTTTC GCCATCGAGCCTGGCGTGGCCGCCTATATCCCAGTGGC CCACGACTACCTGGATGCCCCCGACCAGATCAGCAGG GAGAGAGCCCTGGAGCTGCTGAAGCCTCTGCTGGAGG ATGAGAAGGCCCTGAAGGTCGGCCAGAACCTGAAGTA TGACAGGGGCATCCTGGCCAATTACGGCATCGAGCTGA GAGGCATCGCCTTTGACACCATGCTGGAGTCTTATATC CTGAATAGCGTGGCAGGCCGGCACGACATGGATTCCCT GGCCGAGAGGTGGCTGAAGCACAAGACAATCACCTTC GAGGAGATCGCCGGCAAGGGCAAGAACCAGCTGACCT TCAACCAGATCGCCCTGGAGGAGGCAGGCAGGTACGC AGCAGAGGACGCAGATGTGACCCTGCAGCTGCACCTG AAGATGTGGCCAGATCTGCAGAAGCACtag. 폴리펩티드는 (SEQ ID NO: 47)이다: *MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPA* *ADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI* *KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS* *NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY* *HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE* *GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS* *ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF* *DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS* *DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ* *QLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM* *DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR* *QEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR* *KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP* *KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDL* *LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY* *HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLK* *TYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI* *LDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL* *HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIE* *MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT* *QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQS* *FLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL* *LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITK* *HVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ* *FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY* *KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE* *IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKT*

[0202]

번호	명칭	목적	생성 전략
			*EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE LALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLY ETRIDLSQLGGDKRPAATKKAGQAKKKKGSSETPGTSESAT PEGG**APEVTATVISYDNYVTILDEETLKAWIAKLEKAP VFAFDTETDSLDNISANLVGLSFAIEPGVAAYIPVAHDY LDAPDQISRERALELLKPLLEDEKALKVGQNLKYDRG ILANYGIELRGIAFDTMLESYILNSVAGRHDMDSLAER WLKHKTITFEEIAGKGKNQLTFNQIALEEAGRYAAED ADVTLQLHLKMWPDLQKH** (키: *Cas9* (SEQ ID NO: 27)-링커 (SEQ ID NO: 32)-**3' exo** (SEQ ID NO: 33)
5	pspCas9-Exo-테트라-com-hCLCN5-sp-g1	spCas9 및 POLI 의 5′ 및 3' 엑소뉴클레아제 도메인의 융합 단백질, 및 인간 CLCN5 유전자를 표적화하는 sgRNA 를 발현시키기 위한 플라스미드.	pspCas9-POLI-테트라-com-hCLCN5-sp-g1 을 AccI+NheI 로 절단하여 DNA 중합효소 활성을 코딩하는 DNA 를 제거한다. 평활 말단을 만들기 위해 DNA POLI 의 클레노우 단편으로 DNA 를 처리한다. 이후, T4 DNA 리가제를 사용하여 말단을 라이게이션시킨다. 핵산 서열은 (SEQ ID NO: 48)이다: ATGGACTATAAGGACCACGACGGAGACTACAAGGATC ATGATATTGATTACAAAGACGATGACGATAAGATGGC CCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTC CCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACA TCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGAC GAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGG GCAACACCGACCGGCACAGCATCAAGAAGAACCTGAT CGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAG GCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACA CCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGAT CTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCT TCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGAT AAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCG TGGACGAGGTGGCCTACCACGAGAAGTACCCCACCAT CTACCACCTGAGAAAGAAACTGGTGGACAGCACCGAC AAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCA CATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCG ACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTT CATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGG AAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGC CATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTG GAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGA ATGGCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGC CTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGA GGATGCCAAACTGCAGCTGAGCAAGGACACCTACGAC GACGACCTGGACAACCTGCTGGCCCAGATCGGCGACC AGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCC GACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACAC CGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCA AGAGATACGACGAGCACCACCAGGACCTGACCCTGCT GAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTAC AAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACG CCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTT

[0203]

[0204]

번호	명칭	목적	생성 전략
			CTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACG GCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGA CCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGC ATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCAT TCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGG ACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCG CATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACA GCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAAC CATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAG GGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAA CTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCA AGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAAC GAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGA GAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGC CATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGA CCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAAT CGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAG ATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTG CTGAAAATTATCAAGGACAAGGACTTCCTGGACAATG AGGAAAACGAGGACATTCTGGAAGATATCGTGCTGAC CCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAA CGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGT GATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGG GGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGG ACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAG TCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGAT CCACGACGACAGCCTGACCTTTAAAGAGGACATCCAG AAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACG AGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAG AAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGC TCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACAT CGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAG AAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGG ATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCC TGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAA CGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGG GATATGTACGTGGACCAGGAACTGGACATCAACCGGC TGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGC TTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGAC CAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTG CCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACT GGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAG AAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGC CTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGAC AGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGC ACAGATCCTGGACTCCCGGATGAACACTAAGTACGAC GAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCA CCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGAT TTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCA CCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAA CCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGA GTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGA AGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGC TACCGCCAAGTACTTCTTCTACAGCAACATCATGAACT TTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATC CGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCG GGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCAC

[0205]

번호	명칭	목적	생성 전략
			CGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATC GTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCA AAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCT GATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTAC GGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCT GGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAA CTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCA TGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTT CTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACC TGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTG GAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCG AACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAA ATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGA AGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACA GCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAG ATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGAT CCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCT ACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGC CGAGAATATCATCCACCTGTTTACCCTGACCAATCTGG GAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATC GACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGG ACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTAC GAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACA AAAGGCCGGCGGCCACGAAAAAGGCCGGccaggcaaaaaag aaaaaggGATCCTCTGAGACACCAGGCACCTCCGAGTCTG CCACCCCTGAGGGAGGAAGCGGAGGAAGCGGCTCCGT GCAGATCCCACAGAACCCCCTGATCCTGGTGGACGGCA GCTCCTACCTGTATCGGGCCTACCACGCCTTCCCACCTC TGACAAACTCCGCCGGAGAGCCAACCGGAGCCATGTA TGGCGTGCTGAATATGCTGAGGAGCCTGATCATGCAGT ACAAGCCTACACACGCCGCCGTGGTGTTTGATGCCAAG GGCAAGACCTTCCGCGACGAGCTGTTTGAGCACTACAA GAGCCACAGGCCACCAATGCCTGACGATCTGAGGGCA CAGATCGAGCCACTGCACGCAATGGTGAAGGCCATGG GCCTGCCTCTGCTGGCCGTGAGCGGAGTGGAGGCCGAC GATGTGATCGGCACACTGGCAAGGGAGGCAGAGAAGG CAGGCCGCCCAGTGCTGATCTCCACCGGCGACAAGGAT ATGGCCCAGCTGGTGACACCAAACATCACCCTGATCAA CACCATGACAAATACCATCCTGGGCCCCGAGGAGGTG GTGAATAAGTATGGCGTGCCTCCAGAGCTGATCATCGA TTTCCTGGCCCTGATGGGCGACTCTAGCGATAACATCC CTGGCGTGCCAGGAGTGGGAGAAAAGACCGCACAGGC CCTGCTGCAGGGCCTGGGAGGCCTGGACACCCTGTACG CCGAGCCAGAGAAGATCGCCGGCCTGTCCTTTAGGGGC GCCAAGACAATGGCCGCCAAGCTGGAGCAGAATAAGG AGGTGGCCTACCTGTCTTATCAGCTGGCCACAATCAAG ACCGACGTGGAGCTGGAGCTGACCTGCGAGCAGCTGG AGGTGCAGCAGCCTGCAGCAGAGGAGCTGCTGGGCCT GTTCAAGAAGTACGAGTTTAAGAGATGGACAGCCGAT GTGGAGGCCGGCAAGTGGCTGCAGGCAAAGGGAGCAA AGCCAGCAGCAAAGCCACAGGAGACAAGCGTGGCAGA CGAGGCACCAGAGGTGACAGCCACCGTGATCTCCTAC GATAACTATGTGACAATCCTGGACGAGGAGACACTGA AGGCCTGGATCGCCAAGCTGGAGAAGGCCCCCGTGTTC GCCTTTGATACAGAGACAGACAGCCTGGATAACATCTC CGCCAATCTGGTGGGCCTGTCTTTCGCCATCGAGCCTG GCGTGGCCGCCTATATCCCAGTGGCCCACGACTACCTG

번호	명칭	목적	생성 전략
			GATGCCCCCGACCAGATCAGCAGGGAGAGAGCCCTGG AGCTGCTGAAGCCTCTGCTGGAGGATGAGAAGGCCCT GAAGGTCGGCCAGAACCTGAAGTATGACAGGGGCATC CTGGCCAATTACGGCATCGAGCTGAGAGGCATCGCCTT TGACACCATGCTGGAGTCTTATATCCTGAATAGCGTGG CAGGCCGGCACGACATGGATTCCCTGGCCGAGAGGTG GCTGAAGCACAAGACAATCACCTTCGAGGAGATCGCC GGCAAGGGCAAGAACCAGCTGACCTTCAACCAGATCG CCCTGGAGGAGGCAGGCAGGTACGCAGCAGAGGACGC AGATGTGACCCTGCAGCTGCACCTGAAGATGTGGCCAG ATCTGCAGAAGCACAAGGGCCCCCTGAACGTCTTTGAG AATATCGAGATGCCCCTGGTGCCTGTGCTGAGCCGGAT CGAGCGCAACGGCGTGAAGATCGACCCTAAGGTGCTG CACAATCACTCCGAGGAGCTGACCCTGAGACTGGCCG AGCTGGAGAAGAAGGCCCACGAGATCGCCGGCGAGGA GTTCAACCTGTCCTCTACAAAGCAGCTGCAGACCATCC TGTTTGAGAAGCAGGGCATCAAGCCCCTGAAGAAAAC CCCTGGAGGAGCACCATCTACCAGCGAGGAGGTGCTG GAGGAGCTGGCCCTGGATTATCCCCTGCCTAAAGTGAT CCTGGAGTACCGGGGCCTGGCCAAGCTGAAGTCTctagct cgctttcttgctgtccaatttctattaaaggttcctttgttccctaagtccaactactaa. 폴리펩티드 서열은 (SEQ ID NO: 49)이다: *MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPA* *ADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI* *KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS* *NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY* *HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE* *GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS* *ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF* *DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS* *DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ* *QLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM* *DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR* *QEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR* *KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP* *KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDL* *LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY* *HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLK* *TYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI* *LDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL* *HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIE* *MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT* *QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQS* *FLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL* *LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITK* *HVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ* *FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY* *KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE* *IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKT* *EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA* *YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE* *AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE* *LALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD* *EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL* *FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLY*

[0206]

번호	명칭	목적	생성 전략
			*ETRIDLSQLGGDKRPAATKKAGQAKKKK*GSSETPGTSESAT PEGGSGGSGS**VQIPQNPLILVDGSSYLYRAYHAFPPLTN SAGEPTGAMYGVLNMLRSLIMQYKPTHAAVVFDAKG KTFRDELFEHYKSHRPPMPDDLRAQIEPLHAMVKAM GLPLLAVSGVEADDVIGTLAREAEKAGRPVLISTGDKD MAQLVTPNITLINTMTNTILGPEEVVNKYGVPPELIIDF LALMGDSSDNIPGVPGVGEKTAQALLQGLGGLDTLYA EPEKIAGLSFRGAKTMAAKLEQNKEVAYLSYQLATIK TDVELELTCEQLEVQQPAAEELLGLFKKYEFKRWTA DVEAGKWLQAKGAKPAAKPQETSVADEAPEVTATVIS YDNYVTILDEETLKAWIAKLEKAPVFAFDTETDSLDNI SANLVGLSFAIEPGVAAYIPVAHDYLDAPDQISRERALE LLKPLLEDEKALKVGQNLKYDRGILANYGIELRGIAFD TMLESYILNSVAGRHDMDSLAERWLKHKTITFEEIAG KGKNQLTFNQIALEEAGRYAAEDADVTLQLHLKMWP DLQKHKGPLNVFENIEMPLVPVLSRIERNGVKIDPKVL HNHSEELTLRLAELEKKAHEIAGEEFNLSSTKQLQTIL FEKQGIKPLKKTPGGAPSTSEEVLEELALDYPLPKVIL EYRGLAKLKSLARFLAVQFLLKVPLFPKSNY** (키: *Cas9* (SEQ ID NO: 27)-링커(SEQ ID NO: 32)-**5'-3' exo** (SEQ ID NO: 34)
6	pspCas9-5Exo-테트라-com-hCLCN5-sp-g1	spCas9 및 POLI 의 5' 엑소뉴클레아제 도메인의 융합 단백질, 및 인간 CLCN5 유전자를 표적화하는 sgRNA 를 발현시키기 위한 플라스미드.	pspCas9-POLI-테트라-com-hCLCN5-sp-g1 을 ClaI+NheI 로 절단하여 POLI 코딩 DNA 의 대부분을 제거한다. 이후, pspCas9-POLI-테트라-com-hCLCN5-sp-g1 으로부터 5' 엑소뉴클레아제 도메인을 코딩하는 400 bp 밴드를 증폭시키기 위해 5Exo-F (CAGAGCTGATCATCGATTTCCTGG) (SEQ ID NO: 50) 및 5Exo-R (AagaaagcgagctagccCACGGTGGCTGTCACCTCTGG) (SEQ ID NO: 51)을 사용한다. 이 DNA 를 주입 반응에 의해 ClaI+NheI 선형화된 벡터에 삽입하였다. 핵산 서열은 (SEQ ID NO: 52)이다: ATGGACTATAAGGACCACGACGGAGACTACAAGGATC ATGATATTGATTACAAAGACGATGACGATAAGATGGC CCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTC CCAGCAGCCGACAAGAAGTACAGCATCGGCCTGGACA TCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGAC GAGTACAAGGTGCCCAGCAAGAAATTCAAGGTGCTGG GCAACACCGACCGGCACAGCATCAAGAAGAACCTGAT CGGAGCCCTGCTGTTCGACAGCGGCGAAACAGCCGAG GCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACA CCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGAT CTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCT TCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGAT AAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCG TGGACGAGGTGGCCTACCACGAGAAGTACCCCACCAT CTACCACCTGAGAAAGAAACTGGTGGACAGCACCGAC AAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCA CATGATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCG ACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTT CATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGG

[0207]

[0208]

번호	명칭	목적	생성 전략
			AAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGC CATCCTGTCTGCCAGACTGAGCAAGAGCAGACGGCTG GAAAATCTGATCGCCCAGCTGCCCGGCGAGAAGAAGA ATGGCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGC CTGACCCCCAACTTCAAGAGCAACTTCGACCTGGCCGA GGATGCCAAACTGCAGCTGAGCAAGGACACCTACGAC GACGACCTGGACAACCTGCTGGCCCAGATCGGCGACC AGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCC GACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACAC CGAGATCACCAAGGCCCCCCTGAGCGCCTCTATGATCA AGAGATACGACGAGCACCACCAGGACCTGACCCTGCT GAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTAC AAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACG CCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTT CTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACG GCACCGAGGAACTGCTCGTGAAGCTGAACAGAGAGGA CCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGC ATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCAT TCTGCGGCGGCAGGAAGATTTTTACCCATTCCTGAAGG ACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCG CATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAACA GCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAAC CATCACCCCCTGGAACTTCGAGGAAGTGGTGGACAAG GGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAA CTTCGATAAGAACCTGCCCAACGAGAAGGTGCTGCCCA AGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAAC GAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGA GAAAGCCCGCCTTCCTGAGCGGCGAGCAGAAAAAGGC CATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGA CCGTGAAGCAGCTGAAAGAGGACTACTTCAAGAAAAT CGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAG ATCGGTTCAACGCCTCCCTGGGCACATACCACGATCTG CTGAAAATTATCAAGGACAAGGACTTCCTGGACAATG AGGAAAACGAGGACATTCTGGAAGATATCGTGCTGAC CCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAA CGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGT GATGAAGCAGCTGAAGCGGCGGAGATACACCGGCTGG GGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGG ACAAGCAGTCCGGCAAGACAATCCTGGATTTCCTGAAG TCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGAT CCACGACGACAGCCTGACCTTTAAAGAGGACATCCAG AAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACG AGCACATTGCCAATCTGGCCGGCAGCCCCGCCATTAAG AAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGC TCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACAT CGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAG AAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGG ATCGAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCC TGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAA CGAGAAGCTGTACCTGTACTACCTGCAGAATGGGCGG GATATGTACGTGGACCAGGAACTGGACATCAACCGGC TGTCCGACTACGATGTGGACCATATCGTGCCTCAGAGC TTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGAC CAGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTG CCCTCCGAAGAGGTCGTGAAGAAGATGAAGAACTACT GGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAG AAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGC

[0209]

번호	명칭	목적	생성 전략
			CTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGAC AGCTGGTGGAAACCCGGCAGATCACAAAGCACGTGGC ACAGATCCTGGACTCCCGGATGAACACTAAGTACGAC GAGAATGACAAGCTGATCCGGGAAGTGAAAGTGATCA CCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGAT TTCCAGTTTTACAAAGTGCGCGAGATCAACAACTACCA CCACGCCCACGACGCCTACCTGAACGCCGTCGTGGGAA CCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGA GTTCGTGTACGGCGACTACAAGGTGTACGACGTGCGGA AGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGC TACCGCCAAGTACTTCTTCTACAGCAACATCATGAACT TTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATC CGGAAGCGGCCTCTGATCGAGACAAACGGCGAAACCG GGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCAC CGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATC GTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCA AAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCT GATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTAC GGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCT GGTGGTGGCCAAAGTGGAAAAGGGCAAGTCCAAGAAA CTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCA TGGAAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTT CTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACC TGATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTG GAAAACGGCCGGAAGAGAATGCTGGCCTCTGCCGGCG AACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAA ATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAGA AGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACA GCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAG ATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGAT CCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCCT ACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGC CGAGAATATCATCCACCTGTTTACCCTGACCAATCTGG GAGCCCCTGCCGCCTTCAAGTACTTTGACACCACCATC GACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGG ACGCCACCCTGATCCACCAGAGCATCACCGGCCTGTAC GAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACA AAAGGCCGGCGGCCACGAAAAAGGCCGGccaggcaaaaaag aaaaaggGATCCTCTGAGACACCAGGCACCTCCGAGTCTG CCACCCCTGAGGGAGGAAGCGGAGGAAGCGGCTCCGT GCAGATCCCACAGAACCCCCTGATCCTGGTGGACGGCA GCTCCTACCTGTATCGGGCCTACCACGCCTTCCCACCTC TGACAAACTCCGCCGGAGAGCCAACCGGAGCCATGTA TGGCGTGCTGAATATGCTGAGGAGCCTGATCATGCAGT ACAAGCCTACACACGCCGCCGTGGTGTTTGATGCCAAG GGCAAGACCTTCCGCGACGAGCTGTTTGAGCACTACAA GAGCCACAGGCCACCAATGCCTGACGATCTGAGGGCA CAGATCGAGCCACTGCACGCAATGGTGAAGGCCATGG GCCTGCCTCTGCTGGCCGTGAGCGGAGTGGAGGCCGAC GATGTGATCGGCACACTGGCAAGGGAGGCAGAGAAGG CAGGCCGCCCAGTGCTGATCTCCACCGGCGACAAGGAT ATGGCCCAGCTGGTGACACCAAACATCACCCTGATCAA CACCATGACAAATACCATCCTGGGCCCCGAGGAGGTG GTGAATAAGTATGGCGTGCCTCCAGAGCTGATCATCGA TTTCCTGGCCCTGATGGGCGACTCTAGCGATAACATCC CTGGCGTGCCAGGAGTGGGAGAAAAGACCGCACAGGC CCTGCTGCAGGGCCTGGGAGGCCTGGACACCCTGTACG

번호	명칭	목적	생성 전략
			CCGAGCCAGAGAAGATCGCCGGCCTGTCCTTTAGGGGC GCCAAGACAATGGCCGCCAAGCTGGAGCAGAATAAGG AGGTGGCCTACCTGTCTTATCAGCTGGCCACAATCAAG ACCGACGTGGAGCTGGAGCTGACCTGCGAGCAGCTGG AGGTGCAGCAGCCTGCAGCAGAGGAGCTGCTGGGCCT GTTCAAGAAGTACGAGTTTAAGAGATGGACAGCCGAT GTGGAGGCCGGCAAGTGGCTGCAGGCAAAGGGAGCAA AGCCAGCAGCAAAGCCACAGGAGACAAGCGTGGCAGA CGAGGCACCAGAGGTGACAGCCACCGTGggctag. 폴리펩티드 서열은 (SEQ ID NO: 53)이다: *MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPA* *ADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI* *KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS* *NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY* *HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE* *GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS* *ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF* *DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS* *DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ* *QLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKM* *DGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR* *QEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR* *KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP* *KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDL* *LFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY* *HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLK* *TYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTI* *LDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL* *HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIE* *MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT* *QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQS* *FLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQL* *LNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITK* *HVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ* *FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY* *KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGE* *IRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKT* *EVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA* *YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE* *AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE* *LALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD* *EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHL* *FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLY* *ETRIDLSQLGGDKRPAATKKAGQAKKKK*GSSETPGTSESAT PEGGSGGSGS**VQIPQNPLILVDGSSYLYRAYHAFPPLTN** **SAGEPTGAMYGVLNMLRSLIMQYKPTHAAVVFDAKG** **KTFRDELFEHYKSHRPPMPDDLRAQIEPLHAMVKAM** **GLPLLAVSGVEADDVIGTLAREAEKAGRPVLISTGDKD** **MAQLVTPNITLINTMTNTILGPEEVVNKYGVPPELIIDF** **LALMGDSSDNIPGVPGVGEKTAQALLQGLGGLDTLYA** **EPEKIAGLSFRGAKTMAAKLEQNKEVAYLSYQLATIK** **TDVELELTCEQLEVQQPAAEELLGLFKKYEFKRWTA** **DVEAGKWLQAKGAKPAAKPQETSVADEAPEVTATVG** (키: (키: *Cas9* (SEQ ID NO: 27)-링커 (SEQ ID NO: 32)-**5' exo** (SEQ ID NO: 35))

[0210]

[0211]

번호	명칭	목적	생성 전략
7	pspCas9-POLI-ST2-com-53-sp-g2	spCas9-POLI 융합 단백질 및 인간 DMD 엑손 53 을 표적화하는 sgRNA 를 발현시키기 위한 벡터 플라스미드.	pspCas9-3′UTR-ST2-com-52-sp-g2 의 AflIII+Acc65I 단편을 사용하여 제한 효소 소화 및 T4 DNA 리가제에 의해 pspCas9-POLI-테트라-com-hCLCN5-sp-g1 의 단편을 대체하였다.
8	pspCas9-클레노우-ST2-com-53-sp-g2	spCas9-클레노우 융합 단백질 및 인간 DMD 엑손 53 을 표적화하는 sgRNA 를 발현시키기 위한 플라스미드.	pspCas9-3′UTR-ST2-com-52-sp-g2 의 AflIII+Acc65I 단편을 사용하여 제한 효소 소화 및 T4 DNA 리가제에 의해 pspCas9-클레노우-테트라-com-hCLCN5-sp-g1 의 단편을 대체하였다.
9	pspCas9-POLI-ST2-com 벡터	spCas9-POLI 융합 단백질 및 sgRNA 를 발현시키기 위한 벡터 플라스미드. 발현을 향상시키기 위해 spCas9-POLI 의 3'UTR 에 HBB 3' UTR 의 2 개 복사체가 존재한다. sgRNA 스캐폴드의 ST2 루프를 com-앱타머로 대체하여 com-COM 상호작용을 통해 바이러스 캡시드로의 Cas9 RNP 캡슐화를 가능하게 하였다.	pspCas9-3′UTR-ST2-com 벡터의 AflIII 및 Acc65I 단편을 사용하여 제한 효소 소화 및 T4 DNA 리가제에 의해 pspCas9-POLI-테트라-com 벡터의 AflIII 및 Acc65I 단편을 대체하였다.
10	pspCas9-3'UTR-ST2-com-HBB-W	spCas9 및 야생형 HBB sgRNA 를 발현시키기 위한 플라스미드. sgRNA 스캐폴드의 ST2	어닐링된 HBB-W 올리고를 T4 DNA 리가제에 의해 pspCas9-3'UTR-ST2-com-벡터의 BbsI 부위에 삽입하였다.

[0212]

번호	명칭	목적	생성 전략
		루프를 com-앱타머로 대체하여 com-COM 상호작용을 통해 바이러스 캡시드로의 Cas9 RNP 캡슐화를 가능하게 하였다.	
11	pspCas9-클레노우-ST2-com-벡터	spCas9-클레노우 융합 단백질 및 sgRNA 를 발현시키기 위한 벡터 플라스미드. 발현을 향상시키기 위해 spCas9-클레노우의 3'UTR 에 HBB 3' UTR 의 2 개 복사체가 존재한다. sgRNA 스캐폴드의 ST2 루프를 com-앱타머로 대체하여 com-Com 상호작용을 통해 바이러스 캡시드로의 Cas9 RNP 캡슐화를 가능하게 하였다.	U6-sgRNA 카세트를 함유하는 AflIII-Acc65I 단편을 pspCas9-poli-ST2-com 벡터로부터 절단하고, T4 DNA 리가제에 의해 pspCas9-클레노우-ST2-com-53-sp-g2 의 AflIII-Acc65I 부위에 삽입하였다.
12	pspCas9-클레노우-ST2-com-HBB-W	spCas9- 클레노우 융합 단백질 및 인간 HBB 를 표적화하는 sgRNA 를 발현시키기 위한 플라스미드.	어닐링된 HBB 올리고뉴클레오티드(HBB-sp-WF 및 HBB-sp-WR)를 T4 DNA 리가제에 의해 pspCas9-클레노우-ST2-com-벡터의 BbsI 부위에 삽입하였다.
13	pspCas9-3'UTR-ST2-com-DMD44	spCas9 및 인간 DMD 엑손 44 를 표적화하는 sgRNA 를	어닐링된 DMD44 올리고뉴클레오티드(DMD44-g1-F 및 DMD44-g1-R)를 T4 DNA 리가제에 의해 pspCas9-3'UTR-ST2-com-벡터의 BbsI 부위에 삽입하였다.

번호	명칭	목적	생성 전략
		발현시키기 위한 플라스미드.	
14	pspCas9-클레노우-ST2-com-DMD44	spCas9-클레노우 융합 단백질 및 인간 DMD 엑손 44 을 표적화하는 sgRNA 를 발현시키기 위한 플라스미드.	어닐링된 DMD44 올리고뉴클레오티드를 T4 DNA 리가제에 의해 pspCas9-클레노우-ST2-com-벡터의 BbsI 부위에 삽입하였다.
15	pspCas9-3'UTR-ST2-com-g5	spCas9 및 인간 유전자내 1 을 표적화하는 sgRNA를 발현시키기 위한 플라스미드(GRCh38.p13, 염색체 20, 32752960-32752979).	CRISPR J. 2021 Dec;4(6):914-928. doi: 10.1089/crispr.2020.0106. Epub 2021 Mar 16 을 참조한다.
16	pspCas9-클레노우-ST2-com-g5	spCas9-클레노우 융합 단백질 및 유전자내 1 을 표적화하는 sgRNA 를 발현시키기 위한 플라스미드.	어닐링된 올리고뉴클레오티드 ABE-g5-F 및 ABE-g5-R 을 T4 DNA 리가제에 의해 pspCas9-클레노우-ST2-com-벡터의 BbsI 부위에 삽입하였다.
17	pspCas9-3'UTR-ST2-com-P53	spCas9 및 인간 P53 을 표적화하는 sgRNA 를 발현시키기 위한 플라스미드.	CRISPR J. 2021 Dec;4(6):914-928. doi: 10.1089/crispr.2020.0106. Epub 2021 Mar 16 을 참조한다.
18	pspCas9-클레노우-ST2-com-P53	spCas9-클레노우 융합 단백질 및 인간 P53 을 표적화하는 sgRNA 를 발현시키기 위한 플라스미드.	어닐링된 P53 올리고뉴클레오티드(P53-g1F1 및 P53-g1R1)를 T4 DNA 리가제에 의해 pspCas9-클레노우-ST2-com-벡터의 BbsI 부위에 삽입하였다.

[0213]

번호	명칭	목적	생성 전략
19	pspCas9-3'UTR-ST2-com-MRE11	spCas9 및 인간 MRE11 을 표적화하는 sgRNA 를 발현시키기 위한 플라스미드.	어닐링된 MRE11 올리고뉴클레오티드(Mre11-gF 및 Mre11-gR)를 T4 DNA 리가제에 의해 pspCas9-3'UTR-ST2-com-벡터의 BbsI 부위에 삽입하였다.
20	pspCas9-클레노우-ST2-com-MRE11	spCas9-클레노우 융합 단백질 및 인간 MRE11 을 표적화하는 sgRNA 를 발현시키기 위한 플라스미드.	어닐링된 MRE11 올리고뉴클레오티드를 T4 DNA 리가제에 의해 pspCas9-클레노우-ST2-com-벡터의 BbsI 부위에 삽입하였다.
21	pspCas9-3'UTR-ST2-com-CtIP	spCas9 및 인간 RBBP8 을 표적화하는 sgRNA 를 발현시키기 위한 플라스미드.	어닐링된 CtIP 올리고뉴클레오티드(CtIP-gF 및 CtIP-gR)를 T4 DNA 리가제에 의해 pspCas9-3'UTR-ST2-com-벡터의 BbsI 부위에 삽입하였다.
22	pspCas9-클레노우-ST2-com-CtIP	spCas9-클레노우 융합 단백질 및 인간 TBBP8 을 표적화하는 sgRNA 를 발현시키기 위한 플라스미드.	어닐링된 CtIP 올리고뉴클레오티드를 T4 DNA 리가제에 의해 pspCas9-클레노우-ST2-com-벡터의 BbsI 부위에 삽입하였다.
23	Ppoli-CLCN5-g1	핵 표적화 신호를 갖는 E. 콜리 중합효소 I 및 *CLCN5* 에 대한 sgRNA 를 발현시키기 위한 플라스미드.	pspCas9-poli-테트라-com-clcn5-sp-g1 을 PpuMI 로 절단하여 Cas9 기능성 도메인을 인코딩하는 2.6 kb 를 결실시키고, 나머지 DNA 를 자가-라이게이션시켰다.
24	pspCas9-poli(D705A)-테트라-com-clcn5-sp-g1	중합효소 활성을 불활성화시키기 위한 D705A 돌연변이를 갖는 Cas9-pol 융합 단백질 및 *CLCN5* 에 대한 sgRNA 를	pspCas9-poli-테트라-com-clcn5-sp-g1 을 AccI 및 NheI 로 절단하고, 벡터를 회수하고(1 bk 밴드 제거); Pol-F1 (GCCAAGCTGAAGTCTACATATAC) (SEQ ID NO: 54) 및 Pol-MR (CTGAGAGTAGGCGGCGGACACG) (SEQ ID NO: 55)을 프라이머로서 사용하고 pspCas9-poli-테트라-com-clcn5-sp-g1 을 225 bp 의 단편(단편 1)을 증폭시키기 위한 주형으로 사용하고; Pol-MF

[0214]

번호	명칭	목적	생성 전략
		발현시키기 위한 플라스미드.	(CGTGTCCGCCGCCTACTCTCAG) (SEQ ID NO: 56) 및 pol-R1 (aagaaagcgagctagcaatgaa)을 프라이머로서 사용하고 pspCas9-poli-테트라-com-clcn5-sp-g1 을 840 bp 의 단편(단편 2)을 증폭시키기 위한 주형으로 사용한다. 주입 반응에 의해 단편 1 및 2 를 회수된 벡터에 삽입한다.
25	pspCas9-클레노우(D705A)-테트라-com-clcn5-sp-g1	중합효소 활성을 불활성화시키기 위한 D705A 돌연변이를 갖는 Cas9-클레노우 융합 단백질 및 *CLCN5* 에 대한 sgRNA 를 발현시키기 위한 플라스미드.	pspCas9-poli(D705A)-테트라-com-clcn5-sp-g1 의 AccI-FseI 단편을 T4 DNA 리가제에 의해 pspCas9-클레노우-테트라-com-clcn5-sp-g1 의 AccI-FseI 단편으로 대체하여 5' 엑소뉴클레아제 코딩 영역을 제거하였다.

[0215]

[0216] **표 3. 프라이머 서열 정보**

프라이머 명칭	서열	목적
Reporter-F	Tccatttcaggtgtcgtgag (SEQ ID NO: 57)	GFP 리포터 카세트에서 DNA 를 증폭시키기 위함.
Reporter-R2	TCCAGCTCGACCAGGATG (SEQ ID NO: 58)	
HBB-F1	AGCCAGGGCTGGGCATAAAAG (SEQ ID NO: 59)	HBB-R3 으로 NGS 에 대한 HBB 영역을 증폭시키는 데 사용됨
HBB-R3	TGGGAAAATAGACCAATAGGCAGAG (SEQ ID NO: 60)	HBB-R 또는 HBB-R 로 NGS 에 대한 HBB 영역을 증폭시키는 데 사용됨
DMD53-F	TCCTGTTGTTCATCATCCTAGC (SEQ ID NO: 61)	NGS 에 대한 DMD 53 엑손 표적 영역을 증폭시키기 위함
DMD53-R	TCCAGCCATTGTGTTGAATC (SEQ ID NO: 62)	
hCLCN5-F	GTTTAAGGGCCCGCCTTTTG (SEQ ID NO: 63)	NGS 에 대한 CLCN5 표적 영역을 증폭시키기 위함
hCLCN5-R	TGTCTTACCTCTCGGTGCCT (SEQ ID NO: 64)	
hCLCN5-F3	GACCCAGGTTTCCTGAGCTG (SEQ ID NO: 65)	SMRT 시퀀싱을 위해 CLCN5 표적 영역을 증폭시키기 위함
hCLCN5-R3	TTCAGAGCTTCCTCCCAAGC (SEQ ID NO: 66)	
OT4-F	CCCTATACCTGGGCTCCGTT (SEQ ID NO: 67)	NGS 에 대한 표적-외 4 영역을 증폭시키기 위함
OT4-R	GAAAGGGCCTCTCTCTTTGTAATG (SEQ ID NO: 68)	
OT5-F	AAGCTCTACAAGGGCAGAGAATG (SEQ ID NO: 69)	NGS 에 대한 표적-외 8 영역을 증폭시키기 위함
OT5-R	TCAAAGCTCCCAGATTCACGTT (SEQ ID NO: 70)	
OT8-F	CTGCTCTTTGCCTGTTGGAG (SEQ ID NO: 71)	NGS 에 대한 표적-외 8 영역을 증폭시키기 위함
OT8-R	GCTAAAGCTGGAAGGCTGTG (SEQ ID NO: 72)	
HBD-F2	AAAAGGCAGGGCAGAGTCGACTG (SEQ ID NO: 73)	NGS 에 대한 HBD 영역을 증폭시키기 위함
HBD-R2	GGTAGGAAAACAGCCCAAGGGAC (SEQ ID NO: 74)	
DMD44-F	CCATCACCCTTCAGAACCTGA (SEQ ID NO: 75)	NGS 에 대한 DMD 엑손 44 영역을 증폭시키기 위함
DMD44-R	Tcagtggctaacagaagctga (SEQ ID NO: 76)	
CLCN5 5' 서열	GTAGGATTCTAATCACTGCCTGCTC (SEQ ID NO: 77)	SMRT 시퀀싱 데이터를 분석할 때 결실을 찾는 데 사용되는 서열
CLCN5 3' 서열	CTTTCTGCACCTCCTGATAGCCTTG (SEQ ID NO: 78)	

[0217]

MRE11-onF	GCCAAGTGTGAATGTGCACA (SEQ ID NO: 79)	MRE11 NGS 에 대한 표적-내를 증폭시키기 위함
MRE11-onR	CCTCTTAGGCTATGACCAGGG (SEQ ID NO: 80)	
Ctip-onF	TCATTGGGAGGCCGAACATC (SEQ ID NO: 81)	NGS 에 대한 RBBP8 의 표적-내를 증폭시키기 위함
CtIP-onR	AAGGGCTGAAGGATGATGCA (SEQ ID NO: 82)	
P53-onF	CTGGCATTCTGGGAGCTTCA (SEQ ID NO: 83)	NGS 에 대한 p53 의 표적-내를 증폭시키기 위함
P53-onR	GAGACCTGTGGGAAGCGAAA (SEQ ID NO: 84)	
g5-onF	GTCTGAGGTCACACAGTGGG (SEQ ID NO: 85)	NGS 에 대한 유전자내 1 의 표적-내를 증폭시키기 위함
g5-onR	CTGAGAGCAGGGACCACATC (SEQ ID NO: 86)	
ABE-g5-F	ACCGGATGAGATAATGATGAGTCA (SEQ ID NO: 87)	유전자내 1 sgRNA 를 발현시키기 위한 작제물을 제조하기 위함
ABE-g5-R	aaacTGACTCATCATTATCTCATC (SEQ ID NO: 88)	
P53-g1F1	ACCGCCATTGTTCAATATCGTCCG (SEQ ID NO: 89)	P53 sgRNA 를 발현시키기 위한 작제물을 제조하기 위함
P53-g1R1	AAACCGGACGATATTGAACAATGG (SEQ ID NO: 90)	
DMD44-g1-F	CACCGtttagcatgttcccaattctc (SEQ ID NO: 91)	DMD 엑손 44 sgRNA 를 발현시키기 위한 작제물을 제조하기 위함
DMD44-g1-R	AAACgagaattgggaacatgctaaaC (SEQ ID NO: 92)	
HBB-sp-WF	ACCGGTAACGGCAGACTTCTCCTC (SEQ ID NO: 93)	HBB sgRNA 를 발현시키기 위한 작제물을 제조하기 위함
HBB-sp-WR	AAACGAGGAGAAGTCTGCCGTTAC (SEQ ID NO: 94)	
CtIP-gF	accgTTGCCCAAAGATTCCCCAG (SEQ ID NO: 95)	RBBP8 sgRNA 를 발현시키기 위한 작제물을 제조하기 위함
CtIP-gR	aaacCTGGGGAATCTTTGGGCAA (SEQ ID NO: 96)	
Mre11-gF	accgTGACTGAGATCTGAGTGCTC (SEQ ID NO: 97)	MRE11 sgRNA 를 발현시키기 위한 작제물을 제조하기 위함
Mre11-gR	aaacGAGCACTCAGATCTCAGTCA (SEQ ID NO: 98)	
53-sp-g2-F	ACCGactgttgcctccggttctga (SEQ ID NO: 99)	DMD 엑손 53 sgRNA 를 발현시키기 위한 작제물을 제조하기 위함
53-sp-g2-R	AAACtcagaaccggaggcaacagt (SEQ ID NO: 100)	
hCLCN5-sp-g1-F	ACCGGAGGACAAGTCGTACAATGG (SEQ ID NO: 101)	CLCN5 sgRNA 를 발현시키기 위한 작제물을 제조하기 위함
hCLCN5-sp-g1-R	AAACCCATTGTACGACTTGTCCTC (SEQ ID NO: 102)	

[0218]

[0219] **표 4. 단일 가이드 RNA 표적 서열**

명칭	프로토스페이서	PAM
CLCN5	GAGGACAAGTCGTACAATGG (SEQ ID NO: 103)	TGG
DMD53	ACTGTTGCCTCCGGTTCTGA (SEQ ID NO: 104)	AGG
DMD44	TTTAGCATGTTCCCAATTCTC (SEQ ID NO: 105)	AGG
HBB	GTAACGGCAGACTTCTCCTC (SEQ ID NO: 106)	AGG
유전자내 부위 1	GATGAGATAATGATGAGTCA (SEQ ID NO: 107)	GGG
P53	CCATTGTTCAATATCGTCCG (SEQ ID NO: 108)	GGG
MRE11	TGACTGAGATCTGAGTGCTC (SEQ ID NO: 109)	TGG
RBBP8	TTGCCCAAAGATTCCCCAG (SEQ ID NO: 110)	GGG

[0220]

[0221] **[0203] 세포 배양.** 최근에 보고된 HEK293T(ATCC CRL-3216™) 및 HEK293T-유래 *CLCN5* GFP-리포터 세포(Lu et al. Lentiviral Capsid-Mediated *Streptococcus pyogenes* Cas9 Ribonucleoprotein Delivery for Efficient and Safe Multiplex Genome Editing. *CRISPR J* 4(6):914-928 (2021))를 5% CO_2를 갖는 인큐베이터에서 37℃에서 10% FBS, 2 mM L-글루타민, 100 U/ml 페니실린 및 100 μg/ml 스트렙토마이신을 갖는 DMEM(ThermoFisher Scientific, Waltham, MA)에서 배양하였다. 인간 폐 섬유모세포 IMR90 세포(ATCC CCL-186)를 10% FBS, 2 mM L-글루타민, 100 U/ml 페니실린 및 100 μg/ml 스트렙토마이신이 보충된 이글 최소 필수 배지(Eagle's Minimum

Essential Medium)에서 배양하였다. 가동화 말초 혈액(Mobilized Peripheral Blood)(Lonza, 카탈로그 #: 4Y-101C, Basel, Switzerland)으로부터의 인간 CD34+ 전구 세포를 1x StemSpan™ CD34+ Expansion Supplement(Stemcell Technology, 카탈로그 #02691)로 보충된 혈청 비함유 배지(Stemcell Technology, 카탈로그 # 09605, Vancouver, CA)에서 배양하였다. 인간 골격근 근모세포(Lonza, 카탈로그 #: CC-2580)를 SkGMTM-2 골격근 세포 성장 배지-2 BulletKit™(Lonza, 카탈로그 #: CC-3245)에서 배양하였다.

[0222] **[0204] HEK293T 세포의 형질감염.** HEK293T 세포를 FuGENE HD(Promega, 카탈로그#: E2312, Madison, WI)를 사용하여 24-웰 플레이트에서 형질감염시켰다. 형질감염 전날, 1.25 x 10^5개 세포를 24-웰 플레이트에 시딩하였다. DNA 형질감염을 위해, 0.5 μg의 플라스미드 DNA를 50 μl의 OPTI-MEM에 첨가하였다. 다른 튜브에서, 1.5 μl의 FuGENE HD를 50 μl의 OPTI-MEM에 첨가하였다. 2개의 혼합물을 혼합하고 실온에서 15분 동안 인큐베이션한 후, 세포에 첨가하고, 이의 배지를 DNA 형질감염 직전에 OPTI-MEM으로 교체하였다. 형질감염 24시간 후에, 배지를 정상 성장 배지로 교체하고, 형질감염 72시간 후에 세포를 분석하였다.

[0223] **[0205] 인간 일차 세포의 뉴클레오펙션.** Nucleofector™ 2b 장치(Lonza)를 인간 일차 세포의 뉴클레오펙션에 사용하였다. IMR90 세포, 인간 근모세포 및 인간 CD34+ 조혈 세포를 Cell Line Nucleofector™ Kit R(Lonza, 카탈로그 #: VCA-1001, 프로그램 X-001), Human Dermal Fibroblast Nucleofector™ Kit(Lonza, 카탈로그 #: VPD-1001, 프로그램 P-022), 및 Human CD34+ Cell Nucleofector™ Kit(Lonza, 카탈로그 #: VPA-1003)로 각각 뉴클레오펙션시켰다. 각 뉴클레오펙션에 대한 세포 수는 2 x 10^5였다. 4.5 μg의 표적 플라스미드 DNA(sgRNA/Cas9 또는 sgRNA/Cas9-클레노우를 발현함) 및 0.5 μg의 GFP-발현 플라스미드 DNA(CmiR0001-MR03, GeneCopoeia, Inc., Rockville, MD)를 각각의 뉴클레오펙션에 사용하였고, 여기서 GFP-발현 플라스미드 DNA를 뉴클레오펙션 효율에 대한 지표로 사용하였다. 형질감염된 세포를 추가 실험 전에 형광 현미경하에서 유사한 GFP 양성 백분율에 대해 확인하였다.

[0224] **[0206] 인간 HEK293T 세포 및 IMR90 세포에서의 *MRE11* 또는 *RBBP8* 녹다운.** CRISPR/Cas9 뉴클레아제를 사용하여 *MRE11* 또는 *RBBP8*을 녹다운시켜 Cas9 또는 Cas9-클레노우에 의해 유도된 *CLCN5* 돌연변이 프로파일에서 이들의 역할을 관찰하였다. sgRNA 서열은 문헌[Shou et al., "Precise and Predictable CRISPR Chromosomal Rearrangements Reveal Principles of Cas9-Mediated Nucleotide Insertion. *Mol Cell*, 71, 498-509 e494 (2018)]에 설명된 바와 같이 검증되었으며, 이는 표 4에 열거되어 있다. *MRE11* 또는 *RBBP8* 녹다운의 기회를 증가시키기 위해, *CLCN5* sgRNA- 및 *MRE11/RBBP8* sgRNA-발현 플라스미드에 대한 DNA 비는 1:2였다. DNA 혼합물을 상기 설명된 바와 같이 FuGENE HD에 의해 HEK293T 세포 내로 및 뉴클레오펙션에 의해 IMR90 세포 내로 공동-형질감염시켰다.

[0225] **[0207] 렌티바이러스-유사 입자의 생산.** 렌티바이러스-유사 입자는 이전에 설명된 바와 같이 생산되었다(Lu et al. (2021)). 패키징 플라스미드 pspAX2-D64V-NC-COM은 뉴클레오캡시드 단백질에 삽입된 앱타머-결합 단백질 Com을 갖고, sgRNA의 ST2 루프는 com 앱타머로 대체되어 Cas9 RNP를 앱타머 com과 앱타머-결합 단백질 Com 사이의 상호작용을 통해 렌티바이러스 캡시드로 패키징할 수 있었다. 요약하면, 5백만개의 HEK293T 세포를 10-cm 조직 배양 디쉬에 시딩하였다. 세포 시딩 24시간 후에, 하기 DNA 혼합물을 500 μl의 OPTI-MEM: 7.5 μg의 pspAX2-D64V-NC-COM, 7.5 μg의 *CLCN5* sgRNA 및 Cas9(또는 Cas9-클레노우)를 발현하는 플라스미드 DNA, 및 3 μg의 pMD2.G에 첨가하였다. 한편, 500 μl의 OPTI-MEM을 54 μl의 1 mg/ml 폴리에틸렌이민(PEI, Polysciences Inc., Warrington, PA)과 혼합하였다. 혼합물을 실온에서 15분 동안 인큐베이션한 후, 혼합물을 전날 시딩된 세포에 첨가하고, 배지를 DNA 형질감염 직전에 OPTI-MEM으로 교체하였다. 형질감염 24시간 후에, 배지를 정상 성장 배지로 교체하고, 바이러스-유사 입자를 함유하는 배양 배지를 배지 교체 48시간 후에 수집하였다. 입자의 농도를 p24 기반 ELISA(Cell Biolabs, QuickTiter™ Lentivirus Titer Kit 카탈로그 번호 VPK-107, San Diego, CA)에 의해 정량하였다.

[0226] **[0208] 렌티바이러스-유사 입자의 형질도입.** 렌티바이러스-유사 입자를 HEK293T 세포 또는 *CLCN5* GFP 리포터 세포로 형질도입시키기 위해, 바이러스-유사 입자를 8 μg/ml 폴리브렌의 존재 하에 세포에 첨가하였다. 최대 200 ng의 p24 입자를 24-웰 플레이트에서 성장된 2.5 x 10^4개 세포에 첨가하였다. 형질도입 24시간 후에, 배지를 정상 성장 배지로 교체하였다. *CLCN5* GFP 리포터 세포에서 GFP 발현은 형질도입 36시간 후에 검출될 수 있었다. 형질도입 72시간 후에 DNA 분리를 위해 세포를 수집하였다.

[0227] **[0209] INDEL 프로파일 분석을 위한 표적 DNA의 PCR 증폭.** 유전체 DNA를 제조사의 지침에 따라 DNeasy Blood & Tissue Kit(Qiagen, Germantown, MD)로 분리하였다. 표적 DNA를 증폭시키기 위해 사용된 프라이머 서열은 표 3

에 제시되어 있다. PCR에 대한 유전체 DNA 주형 입력은 최대 0.5 μg이었다. 낮은 DNA 농도를 갖는 샘플의 경우, 0.2 μg의 DNA를 사용하였다. 미리 결정된 최소 사이클 수(25-30)를 사용하여 증폭 편향을 감소시켰다. 교정 CloneAmp HiFi PCR Premix(Takara, Mountain View, USA; 카탈로그 #639298)를 PCR에 사용하였다.

[0228] **[0210] 표적-외 분석.** *HBB* sgRNA1에 대한 4개의 잠재적인 표적-외를 분석하여 Cas9와 Cas9-클레노우 사이의 표적-외 활성을 비교하였다. 이들은 G1-OT4, G1-OT5, HBD 및 Off-8을 포함하였다. 예측된 표적-외의 영역을 이들의 각각의 특정 프라이머(표 3)로 증폭시키고, 차세대 시퀀싱(NGS) 분석을 수행하였다.

[0229] **[0211] NGS 및 데이터 분석.** NGS 분석은 Genewiz Inc.(Morrisville, NC)에 의해 "Amplicon EZ" 서비스를 사용하여 수행되었다. 샘플 당 대략 50,000개의 판독이 획득되었다. 3' 링커 및 5' 바코드 서열을 제거한 후, 생성된 판독을 돌연변이 분석을 위해 온라인 소프트웨어 Cas-Analyzer(Park et al., "Cas-analyzer: an online tool for assessing genome editing results using NGS data. *Bioinformatics*, 33, 286-288 (2017)), 및 CRISPResso2(Clement et al., CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nat Biotechnol*, 37, 224-226 (2019))에 제공하였다. 두 프로그램은 대부분의 경우 유사한 결과를 제공하였다. 그러나, CRISPResso2는 INDEL 비율이 5% 미만일 때 잘 수행되지 않았다. 달리 언급되지 않는 한 제시된 데이터를 Cas-Analyzer로 분석하였다.

[0230] **[0212] 단일-분자 실시간(SMRT) 시퀀싱(PacBio).** 인간 *CLCN5* 유전자를 표적으로 하는 큰 결실을 검출하기 위해 단일-분자 실시간(SMRT) 시퀀싱을 수행하였다. 4862 bp의 영역을 프라이머 hCLCN5-F3 및 hCLCN5-R3을 사용하여 LongAmp® Hot Start Taq 2X Master Mix(New England Biolab 카탈로그 번호: M0533, Ipswich, MA)에 의해 증폭시켰다. DNA를 SMRT 시퀀싱(Sequel I)을 위해 GCB Sequencing and Genomic Technologies Shared Resource(Duke University, Durham, NC)에 제출하였다. 8개의 바코딩된 샘플에 대해 하나의 SMRTcell을 사용하였다. Sequencing Center로부터 수신된 역다중화된 원형 컨센서스(CCS) 판독을 사용하여 R 프로그래밍을 통해 참조 서열과 비교함으로써 큰 결실을 검색하였다. 구체적으로, 모든 서열은 25-bp 5' 서열 및 25-bp 3' 서열의 존재에 대해 검색되었고(서열에 대해서는 표 3 참조), 각각은 가능한 시퀀싱 오류를 수용하기 위해 2개의 뉴클레오티드 차이(92% 동일성)를 허용하였다. 이러한 2개의 영역은 시퀀싱된 DNA의 5' 및 3' 말단에 있다. 이후, 각각의 판독에 대해 2개의 25-bp 영역 사이의 거리를 계산하였다. 참조 서열의 거리와 상이한 거리를 갖고 온전한 sgRNA 표적 부위가 없는 그러한 판독은 결실이 있는 판독으로 간주되었다. 결실이 있는 판독의 카운트 및 서열을 나열하였다.

[0231] **[0213] 통계 분석**. GraphPad Prism 소프트웨어(V5)를 t-검정, ANOVA 및 카이-제곱 검정에 사용하였다. $P<0.05$가 통계적으로 유의한 것으로 간주되었다. 평균 및 s.e.m(평균의 표준 오차)이 보고된다.

[0232] **B. 결과**

[0233] **[0214] DNA DSB에 대한 E. 콜리 DNA 중합효소 I의 표적화는 1-bp 결실 및 TIS를 증가시켰고, CRISPR/Cas9에 의해 유도된 큰 결실을 감소시켰다.** E. 콜리 DNA 중합효소 I(pol I)을 E. 콜리 및 효모에서 EvolvR에 사용된 링커 펩티드와 함께 스트렙토코커스 피오게네스 Cas9의 C-말단에 이를 융합시킴으로써 DNA 이중-가닥 파손(DSB)을 표적화하였다(도 1b). 인간 세포에서 융합 단백질 발현을 향상시키기 위해, pol I을 코딩하는 *polA* 유전자의 코돈을 최적화하였다. 융합 단백질을 사용하여 유전체 편집 활성을 민감하게 검출하기 위해 개발된 GFP-리포터 세포에서 돌연변이가 덴트병으로 언급되는 희귀한 신장병을 유발하는 인간 Chloride Voltage-Gated Channel 5(*CLCN5*)의 5' 비번역 영역(5' UTR)을 표적화하였다(Lu et al. (2021)). 이러한 GFP-리포터 세포는 시작 코돈과 GFP 코딩 서열의 제2 코돈 사이에 *CLCN5* sgRNA 표적 서열(*CLCN5* 5' UTR 영역으로부터)을 삽입함으로써 GFP 판독 프레임의 파괴로 인해 GFP를 발현하지 않는다. GFP는 유전체 편집이 인-프레임 INDEL에 의해 GFP 판독 프레임을 복원하는 경우에만 발현될 것이다(약 1/3 기회). 유사한 GFP-양성 세포가 *CLCN5* gRNA/Cas9 및 *CLCN5* gRNA/Cas9-pol로 처리된 세포에서 관찰되었다. 표적화된 딥 시퀀싱은 *CLCN5* gRNA/Cas9-pol에 의한 표적 부위에서의 유전체 편집을 확인하여(도 6), Cas9-pol 융합 단백질의 기능성을 입증하였다. *CLCN5* gRNA/Cas9 또는 *CLCN5* gRNA/Cas9-pol(인간 *CLCN5* 5' UTR을 표적화)을 발현하는 플라스미드 DNA를 HEK293T 세포로 형질감염시켰고, Cas9 및 Cas9-pol에 의해 생성된 유사한 INDEL 비율(Cas9의 경우 11.6% ± 1.6%, N=3, Cas9-pol의 경우 14.5% ± 0.7%, N=3, p=0.16)이 관찰되었으며, 이는 융합이 Cas9 활성을 손상시키지 않았음을 시사한다.

[0234] **[0215]** *CLCN5*를 표적화하는 Cas9 및 Cas9-pol에 의해 유도된 INDEL 프로파일을 분석하였고, Cas9-pol이 1-염기쌍(bp) 결실의 유의한 증가, >1-bp 결실의 감소(도 1c), 및 1-bp TIS 대 1-bp 비-TIS 비의 증가(도 1d)를 야기한 것으로 밝혀졌다. 데이터는 중합효소 I을 DSB로 표적화하는 것이 긴 결실보다 짧은 결실의 생성을 선호하고, 비-TIS보다 TIS를 선호한다는 가설을 뒷받침하였다.

[0235] **[0216] DSB에 대한 중합효소의 표적화는 TIS를 조절하는데 필요하지만 결실 크기는 아니다.** 중합효소를 DSB로 표적화하는 것이 관찰된 효과에 필요한지 여부를 시험하기 위해, 전장 pol I 및 핵 국소화 신호(NLS)를 함유하지만 Cas9 기능성 도메인의 대부분(REC 및 RuvC의 일부, 전체 HNH 및 Cas9의 PAM-상호작용 도메인의 일부)이 결여된 Cas9-pol의 결실 돌연변이체인 NLS-pol 단백질(도 2a)을 제조하였다. NLS-pol은 Cas9-pol의 NLS 및 링커 서열을 함유하였기 때문에, 적절하게 폴딩되어 핵에 표적화될 것으로 예상되었다. 그러나, 이는 다중 Cas9 도메인의 결실로 인해 Cas9의 DNA-결합 및 뉴클레아제 활성을 갖지 않아야 한다. Cas9, *CLCN5* gRNA, 및 NLS-pol은 HEK293T 세포에서 공동-발현되었고, Cas9 및 NLS-pol로 처리된 세포는 Cas9 처리된 세포와 Cas9-pol 처리된 세포의 것 사이에 1-bp 결실 백분율을 갖고, 공지되지 않은 메커니즘에 대한 예측된 절단 부위의 20bp 5' 및 20bp 3' 영역에서 DNA 치환 비율이 증가한 것으로 밝혀졌다(도 2b). 중요하게는, Cas9 및 NLS-pol의 공동-발현은 TIS를 Cas9-pol의 수준으로 증가시키지 않았다(도 2c).

[0236] **[0217]** 중합효소 활성이 관찰된 효과에 필요한지 여부를 결정하기 위해, Cas9를 중합효소 활성이 불활성화된 pol^{D705A}를 포함하는 다양한 돌연변이체 및 트렁케이션된 pol 단백질에 융합시켰다. 클레노우 단편, 불활성화된 중합효소 활성을 갖는 클레노우^{D705A}, 5' 엑소뉴클레아제 도메인(5Exo), 3' 엑소뉴클레아제 도메인(3Exo), 및 5' 엑소뉴클레아제 및 3' 엑소뉴클레아제 도메인 둘 모두(Exo)(도 2a). *CLCN5*는 이러한 Cas9 융합 단백질을 사용하여 HEK293T 세포에서 표적화되었고, DNA 중합효소 도메인이 없는 모든 융합 단백질은 1-bp 결실을 증가시키고 >1-bp 결실을 감소시키는 데 실패한 것으로 밝혀졌다. 중합효소 활성이 불활성화되었는지 여부에 관계없이, DNA 중합효소 도메인을 갖는 모든 융합 단백질은 1-bp 결실을 증가시키고 >1-bp 결실을 감소시켰다(도 2b). Cas9-5exo, Cas9-3exo 및 Cas9-exo는 서로 매우 유사한 돌연변이 프로파일을 나타내었으므로(도 7), 이들은 도 2b에서 하나의 그룹(Cas9-모든 Exo)으로 처리되었다. 따라서, 중합효소 활성보다는 중합효소 도메인은 1-bp 및 >1-bp 결실의 비를 교란시키기에 충분하였다.

[0237] **[0218]** TIS가 분석될 때, DNA pol I 및 클레노우만이 1-bp TIS를 증가시키고 1-bp 비-TIS를 감소시킬 수 있었고(도 2c), TIS를 생성하기 위한 중합효소-매개 채움의 필요성과 일치하였다. 2-bp 및 3-bp TIS를 분석할 때 동일한 결론이 내려졌다(도 8). pol I과 비교하여, 클레노우는 2-bp 및 3-bp TIS를 증가시키는데 더 강한 효과를 나타내었다. 이는 클레노우 단편에 pol I에 존재하는 5'>3' 엑소뉴클레아제 도메인이 결여되어 있고, 5'>3' 엑소뉴클레아제 도메인이 5' 오버행을 제거하고 결실의 생성을 선호할 수 있기 때문일 수 있다.

[0238] **[0219]** *CLCN5*는 다양한 융합 단백질로 인간 일차 IMR90 세포에서 유사하게 표적화되었다. 이러한 세포에서, 클레노우 도메인을 갖는 모든 융합 단백질(중합효소 활성을 갖거나 갖지 않음)은 1-bp 결실 빈도를 유의하게 증가시켰지만, 삽입 빈도는 오히려 >1-bp 결실 빈도를 감소시켰다(도 2d). 다시, DSB에 대한 중합효소 활성의 표적화는 TIS를 증가시키는 데 필요하였다(도 2e).

[0239] **[0220]** 돌연변이 프로파일과 Cas9에 의한 *CLCN5* 유전자좌를 표적화하는 전체 INDEL 비율 사이의 관계를 조사하였고, 전체 INDEL 비율은 돌연변이 프로파일에 영향을 미치지 않는 것으로 밝혀졌다(도 9). 따라서, DNA 돌연변이 프로파일에 대한 Cas9에 대한 pol I 또는 클레노우 단편 융합의 관찰된 효과는 Cas9 절단 활성에 대한 가능한 효과로 설명될 수 없었다. 가장 가능성 있는 메커니즘은 국소 세포 DNA 복구 기계를 방해하는 것일 수 있다.

[0240] **[0221]** 전체적으로, Cas9와의 융합 또는 중합효소 활성도 1-bp 결실을 증가시키기 위해 필요하지 않았지만, TIS를 증가시키기 위해 중합효소를 DSB로 표적화하는 것이 필요하였다. 그러나, Cas9에 중합효소를 융합시키는 것은 그렇게 하는 것이 국소 중합효소 농도를 증가시키고 다른 내인성 DSB에 대한 가능한 간섭을 감소시킬 것이기 때문에, 특히 많은 실험 설정에서 과발현된 것보다 제어된 양의 유전체 편집 효과기가 유전체 편집에 사용되기 때문에 여전히 유리하다.

[0241] **[0222] DNA 절제는 1-bp 대 >1-bp 결실의 교란에 관여하였다.** pol^{D705A} 및 클레노우^{D705A}가 또한 >1-bp 결실보다 1-bp 결실의 생성을 선호한다는 관찰은 본 발명자가 DNA 절제가 이러한 효과에 관여하는지 여부를 조사하도록 촉구하였다. 따라서, 본 발명자는 *RBBP8*(CtIP 단백질 발현) 녹다운이 있거나 없는 세포에서 *CLCN5* 유전자좌를 표적으로 하는 돌연변이 프로파일을 조사하였다. 하나의 대립유전자로부터의 발현만을 방해할 가능성을 고려하여, 본 발명자는 녹아웃(knockout)보다는 녹다운(knockdown)이라는 용어를 사용한다. 본 발명자는 이 유전자의 돌연변이를 매개하기 위해 이전에 검증된 *RBBP8* sgRNA(6)를 사용하였다. *RBBP8* sgRNA-발현 DNA를 *CLCN5* sgRNA-발현 DNA로 HEK293T 세포에 공동-형질감염시켰다. sgRNA-발현 작제물은 또한 Cas9 또는 Cas9-클레노우-발현 카세트를 함유하였다(사용된 DNA 작제물에 대해서는 보충 표 1 참조). 공동-형질감염된 유전자(*RBBP8* 및 *CLCN5*)의 INDEL 비율은 공동-형질감염과 일치하여 매우 유사하였다(보충 표 S4). 본 발명자는 공지되지 않은 이유로 Cas9

및 Cas9-클레노우-처리된 세포에서 상당히 상이한 전체 INDEL 비율을 관찰하였다. 그러나, DNA 돌연변이 프로파일은 전체 INDEL 비율과 독립적이기 때문에(보충 도면 S4), 이러한 INDEL 비율 차이는 본 발명의 돌연변이 프로파일 분석에 영향을 미치지 않을 것이다.

[0242] **[0223]** *RBBP8*을 녹다운시키는 것은 1-bp 결실의 백분율에 영향을 미치지 않았지만, >1-bp 결실의 백분율을 유의하게 감소시켰고, HEK293T 세포에서 *CLCN5*를 표적화할 때 삽입의 백분율을 증가시켰다(도 3a). 이러한 결과는 삽입 증가를 유발할 수 있는 DNA 절제의 억제와 일치한다. 이러한 관찰은 본 발명자의 논문이 개정되는 동안 출판된 Repair-Seq로부터의 관찰과 유사하다(54). 중요하게는, *RBBP8*이 녹다운된 세포에서, 클레노우 단편은 *RBBP8* 녹다운이 없는 세포에서보다 작은 정도로 >1-bp 결실을 감소시켰으며, 이는 *RBBP8*이 >1-bp 결실의 클레노우-매개 감소에 관여함을 시사한다. 1-bp의 TIS는 *RBBP8* 녹다운 HEK293T 세포에서 유의하게 변화하지 않았으며, 클레노우 단편은 1-bp TIS를 증가시켰다(도 3b).

[0243] **[0224]** *RBBP8* 녹다운이 있거나 없는 IMR90 세포에서 *CLCN5* 돌연변이 프로파일을 또한 조사하였다. IMR90 세포에서 *RBBP8*을 녹다운하면 Cas9-유도된 1-bp 결실을 증가시켰고, 이는 정상 IMR90 세포에서 클레노우 단편을 Cas9에 융합시키는 것과 유사한 효과였다(도 10). 이렇게 하면 또한 1-bp TIS를 증가시켰고 1-bp 비-TIS를 감소시켰다. *MRE11* 또는 *RBBP8* 녹다운 세포에서, 클레노우는 1-bp 결실, >1-bp 결실 또는 TIS에 영향을 미치지 않았다. 이러한 관찰은 *MRE11* 또는 *RBBP8*을 녹다운시키고 클레노우를 Cas9에 융합시키는 것이 IMR90 세포에서 유사한 경로를 방해할 수 있음을 시사한다. 전체적으로, 이러한 데이터는 DNA 절제를 방해하는 것이 Cas9 DNA 돌연변이 프로파일에 대한 pol I 또는 클레노우 단편의 효과에 대한 메카니즘 중 적어도 하나임을 시사한다.

[0244] **[0225]** HEK293T 세포에서 *CLCN5*를 표적화하는 1 내지 30 bp의 결실의 경우, 1-bp 결실을 제외하고, 11-bp 및 8-bp 결실이 가장 빈번하게 관찰되었다(도 3c). 11-bp 결실의 빈도는 *RBBP8* 녹다운에 의해 영향을 받지 않았으며(도 3c), 이는 다른 메커니즘이 또한 이러한 결실의 생성에 관여할 수 있음을 시사한다. 이러한 결실은 클레노우 단편을 Cas9에 융합시킴으로써 크게 억제되었다. 8-bp 결실의 빈도는 *RBBP8* 녹다운에 의해 감소되었고, 클레노우 단편에 의해 추가로 감소되지 않았다. 이러한 관찰은 클레노우 단편이 Cas9 DNA 돌연변이 프로파일에 영향을 미치는 한 가지 방식이 DNA 절제를 방해하는 것을 통하는 것이라는 개념을 추가로 뒷받침한다.

[0245] **[0226]** Cas9-처리된 세포에서 가장 빈번하게 결실된 서열을 조사할 때 결실 주위에 미세상동성이 주목되었다(도 3d, 밑줄). 가장 빈번하게 관찰된 11-bp 결실은 예측된 절단 부위의 맨 끝(점선)에서 미세상동성(녹색 밑줄)을 가졌다. 이 경우, DNA 합성은 MMEJ 동안 일치하지 않는 3' 플랩을 제거하기 위해 XPF-ERCC1(55,56)과 같은 3'-플랩 엔도뉴클레아제에 대한 필요 없이 개시될 수 있다. 이는 아마도 11-bp 결실이 Cas9-처리된 세포에서 가장 빈번하게 관찰된 이유를 설명할 수 있다. *RBBP8* 녹다운은 이러한 결실의 빈도에 영향을 미치지 않았으며, 절제 억제 하에 유사한 결실을 생성하는 다른 공지되지 않은 단백질이 있을 수 있다. 일반적으로, 공통 결실은 결실 주위에 미세상동성을 갖거나, 예측된 절단 부위에 한쪽 말단을 갖거나, 둘 모두를 갖는다.

[0246] **[0227] Cas9-클레노우는 다수의 인간 세포 유형에서 다수의 유전자좌에서 1-bp 결실을 증가시켰다.** DNA 돌연변이 프로파일에 대한 pol I 또는 클레노우의 효과가 표적 서열- 또는 세포 유형-특이적인지 여부를 조사하였다. Cas9-클레노우 융합 단백질은 이의 더 작은 크기 및 증가하는 1-bp 결실 및 TIS에 대한 현저한 효과를 고려하여 후속 실험에 사용되었다. 뒤시엔느 근이영양증(*DMD*) 엑손 53 및 *DMD* 엑손 44(서로로부터 537 kb 떨어져 있음), *HBB*의 5' 코딩 영역, 및 유전자간 유전자좌 유전자내 1(GRCh38.p13, 염색체 20, 32752960-32752979)을 포함하는 4개 이상의 유전자좌를 조사하였다. 단일-절단 sgRNA로 이러한 엑손을 표적화하는 것이 엑손 결실에 의해 야기된 DMD 환자에서 디스트로핀을 회복시킬 수 있기 때문에 *DMD* 엑손 53 및 44를 선택하였다. 겸상 적혈구 질환을 치료하는데 유전체 편집의 가능한 적용을 위해 *HBB* 5' 코딩 영역을 선택하였다. 또한, 이 영역은 인간 체세포에서 Cas9-유도된 유전자 전환을 조사하기 위해 CRISPR/Cas9로 표적화되었다(Parsijani et al. "CRISPR/Cas9 increases mitotic gene conversion in human cells," *Gene Ther*, **27**, 281-296 (2020)). 유전자내 유전자내 1을 선택하여 관찰된 효과에 대한 표적 유전자 생성물의 가능한 기여를 배제하였다. HEK293T 세포에 추가하여, 다양한 유전자좌가 인간 일차 섬유모세포 IMR90 세포, 인간 CD34+ 조혈 줄기 세포 또는 인간 일차 근모세포에서 표적화되었다. 총 8개의 유전자좌/세포를 조사하였다(표 5).

[0247] **표 5. 결실, 총 삽입 및 1-bp TIS에 대한 Cas9-클레노우의 효과**

유전자좌/세포	1-bp 결실 (모든 INEL 의 %)		>1-bp 결실 (모든 INEL 의 %)		삽입 (모든 INEL 의 %)		1-bp TIS (모든 삽입의 %)	
	Cas9	Cas9-Kle	Cas9	Cas9-Kle	Cas9	Cas9-Kle	Cas9	Cas9-Kle
CLCN5/293T	9.9±0.92	23.1±0.53***	55.3±0.66	39.8±1.2***	34.8±0.45	37.1±1.1*	45.0±1.39	65.4±1.44***
CLCN5/IMR90	18.1±2.00	39.3±3.59**	36.7±3.90	39.3±3.72	45.1±5.16	22.7±3.60*	41.2±8.57	70.4±5.33*
DMD53/293T	8.4±0.14	19.8±0.45***	64.4±1.3	51.6±0.37***	27.2±1.40	28.5±0.31	34.0±4.17	52.6±0.68*
DMD44/근모세포	11.2±1.95	28.8±2.71**	75.8±2.06	57.4±2.65**	13.0±1.41	13.7±0.67	32.6±9.27	66.46±6.63*
HBB/조혈 세포	37.4±4.11	59.6±3.05*	56.7±4.59	36.4±2.90*	4.2±1.10	4.0±0.42	21.2±6.02	46.7±5.90*
HBB/IMR90	40.8±9.99	81.2±5.47*	55.9±10.64	14.7±6.32*	2.3±1.02	4.5±2.34	너무 적음	너무 적음
HBB/293T	13.3±0.84	46.7±1.34***	65.1±0.59	41.1±1.29***	19.6±1.03	10.3±0.16**	41.9±1.14	39.1±3.33
유전자내 1/IMR90	14.3±2.1	33.7±1.22**	47.4±2.1	48.0±1.41	38.3±0.60	18.2±0.47***	23.3±1.93	50.3±0.96**

*, **, ***는 양측 t-검정에서 각각 p<0.05, 0.01 및 0.0001 을 나타낸다. 청색, 적색 및 흑색은 각각 Cas9-클레노우 그룹에서 증가, 감소 및 불변을 나타낸다. Cas9-Kle: Cas9-클레노우. 클레노우 융합은 5 개 유전자좌/세포에서 총 삽입 백분율을 유의하게 변화시켰기 때문에, 1-bp TIS 는 모든 삽입의 %로 표현되었다.

[0248]

[0249] **[0228]** 모든 경우에, DSB에 대한 클레노우 단편의 표적화는 1-bp 결실을 유의하게 증가시켰으며(표 5), 이는 8개의 분석된 유전자좌/세포에서 평균 18.80% ± 4.62%(N=8)에서 41.29% ± 7.39%(N=8)로 증가하였다. *HBB*/IMR90에서, Cas9-클레노우에 의해 생성된 모든 INDEL의 81.2% ± 5.47%는 1-bp 결실이었다. 1-bp 결실의 증가는 4가지 가능한 현상을 동반하였다: 1) >1-bp 결실의 감소(*DMD44*/근모세포, *HBB*/조혈 세포 및 *HBB*/IMR90); 2) 삽입의 감소(*CLCN5*/IMR90, 유전자내 1/IMR90), 3) >1-bp 결실 및 삽입 둘 모두의 감소(*HBB*/293T); 및 4) >1-bp 결실의 감소 및 총 삽입의 증가(*CLCN5*/293T, *DMD53*/293T). 삽입에 대한 클레노우의 효과는 다양했지만, 이는 8개 유전자좌/세포 중 6개에서 >1-bp 결실의 유의한 감소를 야기하였다. IMR90, HEK293T 및 조혈 세포에서 *HBB* 유전자좌를 표적화하는 것은 1-bp 결실의 증가를 설명하는 상이한 방식을 나타내었다.

[0250] **[0229]** 명백한 결실 피크를 갖지 않은 *CLCN5*/IMR90을 제외한 모든 유전자좌/세포에서(도 11), Cas9-클레노우는 예측된 절단 부위에서 미세상동성을 갖는 >1-bp 결실의 백분율을 가장 유의하게 감소시켰고(녹색 선으로 밑줄쳐짐), 일반적으로 Cas9에 의해 생성된 가장 높은 >1-bp 결실 피크였다(도 4, 도 11). 더 작은 >1-bp 결실 피크는 미세상동성을 갖지 않았지만 예측된 절단 부위에 위치하거나(점선 수직선), 절단 부위로부터 떨어져 있는 미세상동성을 가졌다(적색 선으로 밑줄). 이러한 관찰은 3' 플랩을 제거할 필요가 없는 이러한 MMEJ 사건이 클레노우에 의한 억제에 가장 민감하였음을 시사한다.

[0251] **[0230] Cas9-클레노우는 다수의 인간 세포 유형에서 다수의 유전자좌에서 TIS를 증가시켰다.** 이러한 유전자좌/세포에서 TIS에 대한 클레노우 단편의 효과를 조사하였다. 2- 및 3-bp TIS는 본 발명자들이 대부분의 유전자좌/세포에서 분석한 약 50,000개 판독에서 적었기 때문에 초점은 1-bp TIS에 있었다. 여러 경우에 클레노우는 전체 삽입의 백분율을 크게 변화시켰으므로, 모든 INDEL보다는 모든 삽입에서 TIS의 백분율을 비교하였다. 삽입이 너무 적고 분석되지 않은 하나의 유전자좌/세포(*HBB*/IMR90)를 제외하고, 7개의 유전자좌/세포 중 6개는 모든 삽입에서 TIS 점유율의 유의한 증가를 나타내었다(표 5). *HBB*/293T만이 Cas9와 Cas9-클레노우 사이에서 유사한 TIS 백분율을 나타내었고, 약 80%의 1-bp 삽입이 TIS였다. 이러한 관찰은 *HBB*/293T의 경우에 내인성 중합효소가 5' 오버행을 채우는데 매우 효율적이었다는 것을 시사하며, 이는 Cas9-클레노우가 TIS 점유율을 추가로 증가시키지 않은 이유를 설명한다.

[0252] **[0231]** 클레노우 단편을 Cas9에 융합시키면 3개의 유전자좌/세포(*CLCN5*/IMR90, *HBB*/293T 및 유전자내 1/IMR90)에서 총 삽입을 유의하게 감소시켰다. 3개의 유전자좌/세포 중 2개(*CLCN5*/IMR90 및 유전자내 1/IMR90)에서, 비-TIS 삽입은 삽입 감소의 100%에 기여하였다.

[0253] **[0232] 클레노우를 Cas9에 융합시키면 CRISPR/Cas9-유도된 큰 DNA 결실이 감소하였다.** >1-bp 결실보다 1-bp 결실을 선호하는 DSB에 대한 DNA 중합효소 표적화의 관찰은 그렇게 하는 것이 큰(>500 bp) 표적-내 결실의 생성을 감소시킬 수 있는지의 조사를 촉발시켰다. HEK293T 세포에서 *CLCN5*를 표적화하는 큰 결실을 분석하였다. 앰플리콘의 중간에 sgRNA 표적 서열을 갖는 4862 bp의 영역을 증폭시키는 한 쌍의 프라이머를 설계하였다(도 4). Cas9

RNP 또는 Cas9-클레노우 RNP를 함유하는 렌티바이러스-유사 입자가 생성되었고, *CLCN5* GFP-리포터 세포를 2가지 유형의 RNP로 처리한 후 유사한 백분율의 GFP-양성 세포가 관찰되었는데, 이는 동등한 유전체 편집 활성을 시사한다.

[0254] **[0233]** 이후, HEK293T 세포를 Cas9 RNP 또는 Cas9-클레노우 RNP로 처리하였다. DNA를 처리 72시간 후에 증폭시키고 단일-분자 실시간(SMRT) 시퀀싱(PacBio)을 수행하였다. Cas9-클레노우 처리된 세포에서보다 Cas9 처리된 세포에서 >0.5 kb, >1 kb 및 >2 kb 결실을 갖는 더 많은 판독이 발견되었다(표 6 및 도 5).

[0255] **표 6. *CLCN5*를 표적화하는 Cas9 및 Cas9-클레노우에 의해 생성된 큰 결실.**

	HEK293T 세포			IMR90 세포		
	Cas9	Cas9-클레노우	sgRNA 없음	Cas9	Cas9-클레노우	sgRNA 없음
총 CCS[a]	23477	11497	12040	19147	29793	11417
>0.5 kb 결실을 갖는 CCS	100	10	0	151	13	1
>1 kb 결실을 갖는 CCS	76	1	0	132	2	1
>2 kb 결실을 갖는 CCS	28	1	0	50	2	0
카이-제곱 검정	P<0.0001			P<0.0001		

[0256] [a]5' 및 3' 인덱스 서열 둘 모두를 함유하는 CCS(도 5의 박스)만을 분석하였다. CCS: 원형 컨센서스

[0257] **[0234]** *CLCN5*는 *CLCN5* sgRNA/Cas9 또는 *CLCN5* sgRNA/Cas9-클레노우를 발현하는 플라스미드 DNA의 뉴클레오펙션에 의해 IMR90 세포에서 표적화되었다. Cas9-처리된 세포가 Cas9-클레노우-처리된 세포보다 더 낮은 짧은 INDEL 비율을 가졌음에도 불구하고(NGS 분석을 위해 동일한 DNA 샘플을 사용한 표 7 참조), 이들은 Cas9-클레노우 처리된 세포와 비교하여 더 많은 >0.5kb, >1kb 및 >2kb 결실을 가졌다.

[0258] **표 7. 인간 일차 세포에서 Cas9 및 Cas9-클레노우의 INDEL 비율(%).**

유전자좌/세포	Cas9	Cas9-클레노우	배경
HBB/IMR90	2.6 ± 0.50	6.2 ± 0.87*	0.02
HBB/CD34+ 세포	0.77 ± 0.03	1.6 ± 0.09***	0.01
CLCN5/IMR90	4.9 ± 0.50	7.8 ± 0.79*	0.09
DMD44/근모세포	11.9 ± 1.8	12.8 ± 0.4	0
유전자내 1/IMR90	0.77 ± 0.12	5.8 ± 0.46 ***	N/A

[0259] * 및 ***는 t-검정에서 각각 p<0.05 및 p<0.0001 을 나타낸다(n=3).

[0260] **[0235]** >0.5 kb 결실을 갖는 CCS 판독을 조사하였고, 결실된 영역이 sgRNA 표적 부위에 걸쳐 있거나 관여하는 것으로 밝혀져(도 5), 이들이 표적-내 결실임을 확인하였다. 제한된 유형의 결실만이 관찰되었다. Cas9는 Cas9-클레노우가 한 것보다 더 많은 유형의 결실 및 각 유형의 결실에 대한 더 흔한 CCS 판독을 생성하였다. 둘 모두의 현상은 Cas9가 Cas9-클레노우보다 큰 결실을 생성하기 쉬운 것으로 설명될 수 있다. 동일한 유형의 결실에 대한 다수의 CCS 판독은 하나의 세포에서 단일 결실 사건 또는 다수의 세포에서 다수의 결실 사건의 결과일 수 있다. Cas9-처리된 및 Cas9-클레노우-처리된 IMR90 세포에서 동일한 결실 유형에 대한 본 발명자의 관찰(도 5b에서 *로 표시됨)은 동일한 결실 유형에 대한 다중 결실 사건의 가능성을 뒷받침하였다. 대안적으로, Cas9는 Cas9-클레노우보다 처리 후 더 빨리 큰 결실을 유도할 수 있었고, 이는 결실의 표현을 증가시켰다.

[0261] **[0236] Cas9-클레노우는 인간 일차 세포에서 INDEL 비율을 증가시켰다.** HEK293T 세포에서 Cas9 및 Cas9-클레노우(또는 Cas9-pol)는 유사한 수준의 INDEL 비율을 생성한 반면, 인간 일차 세포에서 Cas9-클레노우는 6개 유전자좌/세포 중 5개에서 유의하게 더 높은 INDEL 비율을 생성하였다(표 7). 일반적으로, INDEL 비율은 일차 세포에서 비교적 낮았고 그 이유는 불분명하였다. 비-표적화 sgRNA로 처리된 세포의 배경 INDEL이 매우 낮았고, 모든 INDEL이 예측된 절단 부위 주위에 있었기 때문에, 관찰된 INDEL은 진정한 INDEL로 확인되었다. 1/10개의

GFP-발현 플라스미드 DNA가 뉴클레오펙션 실험에 포함되었고, >50%의 GFP-양성 세포가 관찰되었다. 따라서, 낮은 INDEL 비율은 낮은 뉴클레오펙션 효율 때문이 아니었다. 뉴클레오펙션 후 72 배양 시간 동안 세포 사멸이 관찰되었다. 인간 ES 세포에서 관찰된 바와 같이, 지속적으로 생성되는 DSB의 독성 효과로 인해 많은 Cas9 또는 Cas9-클레노우 양성 세포가 손실되었을 수 있다(Happaniemie et al., "CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response," *Nat Med*, **24**, 927-930 (2018)).

[0262] **[0237] Cas9-클레노우는 DNA 치환 비율 또는 표적-외를 증가시키지 않았다.** DNA 중합효소를 DSB로 표적화하는 것이 DSB 주위의 DNA 돌연변이 비율을 증가시킬 수 있는지 여부를 조사하였다. 예측된 절단 부위 주위의 40 bps 영역(각 면에서 20 bp)에서의 DNA 치환 비율을 하기 이유로 조사하였다: 1) DSB로부터 15-20 nt 떨어진 MRE11 닉(18), 2) pol I이 15-20개의 뉴클레오티드의 진행성을 나타내었음(61), 및 3) EvolR(Cas9 닉카제와 오류가 발생하기 쉬운 DNA pol I 사이의 융합 단백질)이 15-20 bp의 돌연변이 윈도우를 나타내었음(47). Cas9-클레노우는 DNA 치환 비율을 증가시키지 않았다(표 4). 또한, 비-표적화 sgRNA를 갖는 음성 대조군에서도 유사한 DNA 치환 비율이 관찰되었으며, 이는 관찰된 DNA 치환이 주로 세포 이질성, PCR 및 시퀀싱 오류로부터 비롯되었음을 시사한다. 따라서, DSB에 대한 DNA 중합효소 표적화는 DNA 치환 비율을 증가시키지 않았다.

[0263] **[0238]** Cas9의 가능한 표적-외에 대한 클레노우 단편을 Cas9에 융합시키는 효과를 또한 시험하였다. 인간 IMR90 세포에서 *HBB* 5' 코딩 영역이 표적화되었고, 서열 유사성에 기반하여 예측된 4개의 잠재적인 표적-외에서 INDEL 비율이 검출되었다(표 8; PAM은 밑줄이 그어져 있다; 이탤릭체 뉴클레오티드는 미스매치를 나타낸다). 현재 가장 민감한 표적외 검출 방법 중 하나인 표적화된 NGS(<0.5%, NGS의 검출 한계)를 사용하여 Cas9 및 Cas9-클레노우 처리된 세포에서 표적-외는 관찰되지 않았다. 따라서, 데이터는 클레노우를 Cas9에 융합시키는 것이 NGS에 의해 검출 가능한 수준으로 표적-외를 증가시키지 않았음을 보여주었다. 클레노우 단편은 진정한 표적과 유사성을 갖는 서열에서 검출 가능한 표적-외를 야기하지 않았기 때문에, 진정한 표적과 유사성을 갖지 않는 서열에서 무작위 표적-외를 야기할 가능성은 적었다.

[0264] **표 8. 표적-외 HBB, HBD, OT-4, OT-5, 및 OT-8의 서열.**

표적	SEQ ID NO	서열
HBB	148	GTAACGGCAGACTTCTCCTCAGG
HBD	149	*TT*GAC*A*GCAG*T*CTTCTCCTCAGG
OT-4	150	*A*TAC*TTAAG*GACTTCTCCTCCAT
OT-5	151	G*GAGG*GGCAG*G*CTTCTCCTCTGG
OT-8	152	*TCACA*GGCAGACTTCTCC*A*CGGG

[0265]

[0266] **[0239]** 표적-외에 대한 클레노우의 효과는 세 번째 방식으로 조사되었다. HEK293T 유래된 GFP-리포터 세포는 *HBB* sgRNA와 하나의 뉴클레오티드 차이를 갖고 *HBB* sgRNA에 대해 "표적-외"인 *HBB* 겸상 돌연변이체 서열에서 CRISPR/Cas9 유도된 INDEL을 검출하기 위해 사용되었다. 이러한 세포는 내인성 *HBB* 유전자에 *HBB* sgRNA 진정한 표적 및 통합 GFP-리포터 카세트에 표적-외를 함유한다. 완벽하게 일치하는 *HBB* sgRNA를 갖는 내인성 *HBB* 유전자를 표적화하고, 표적-외로서 겸상 돌연변이체 서열에서 INDEL 비율을 조사하였다. 분석은 Cas9 및 Cas9-클레노우가 내인성 *HBB* 표적에 대해(Cas9의 경우 21.47% ± 0.80%, N=3; Cas9-클레노우의 경우 21.70% ± 1.10%, N=3, p=0.8723), 또한 통합된 겸상 돌연변이체 서열(Cas9의 경우 10.80% ± 0.80%, N=3; Cas9-클레노우의 경우 12.73% ± 1.20%, N=3, p=0.2509)에 대해 유사한 INDEL 비율을 가졌다. 전체적으로 이러한 실험은 클레노우 단편을 DSB로 표적화하는 것이 DSB에서 표적-외 또는 DNA 돌연변이를 증가시키지 않았음을 입증한다.

[0267] **[0240]** 본원에 제시된 바와 같이, Cas9 융합 단백질을 갖는 DSB에 대한 E. 콜리 DNA pol I 또는 클레노우 단편의 표적화는 작은 결실 대 큰 결실의 비, 및 TIS 대 비-TIS의 비를 증가시켰다. 중요하게는, 그렇게 함으로써 500 bp 초과의 표적-내 결실의 생성이 억제되었다. 이러한 효과는 4개의 세포 유형(1개 세포주 및 3개의 일차 세포 유형) 및 5개의 표적 부위를 포함하는 분석된 모든 유전자좌(8개 유전자좌/세포)에서 관찰되었다. 비-TIS에 비해 결실 크기 감소 및 TIS 증가의 효과는 세포 유형 또는 표적 부위 특이적이지 않다. 일차 세포에서, 클레노우를 Cas9에 융합시키면 5개 사례 중 4개 사례에서 전체 INDEL 비율이 유의하게 증가하였다.

[0268] **[0241]** DNA 절제는 HDR, MMEJ 및 SSA에 필요하다. 후자의 2개의 대안적인 NHEJ DNA 복구 경로는 짧은 결실 및 긴 결실을 생성할 것이다. MRE11-RAD50-NBS1 복합체는 DNA 절제 개시를 담당하고, EXO1, BLM 및 DNA2는 광범위한 절제를 담당한다. 대응하는 DNA 절제를 통해 큰 결실의 생성을 억제하려는 시도가 있었다. 본원에 제공된 데

이터는 DNA 절제를 방해하는 것이 결실 크기에 대한 클레노우의 효과에 대한 메커니즘 중 하나임을 시사한다. 먼저, DNA 절제에 관여하는 단백질인 MRE11 또는 CtIP를 녹다운시키는 것이 1-bp 결실을 증가시키고 >1-bp 결실을 감소시키는 것으로 결정되었다. 둘째, 결실 크기에 대한 클레노우의 효과는 MRE11 또는 CtIP 녹다운 하에서 손실되었다. 이러한 관찰은 MRE11 복합체를 억제하는 것이 MMEJ의 억제를 유발한다는 보고와 일치한다(Hussmann et al. "Mapping the genetic landscape of DNA double-strand break repair. *Cell*, 184, 5653-5669 e5625 (2021)). *CLCN5*에서 흔히 관찰되는 11-bp 결실은 MRE11 또는 CtIP 녹다운에 의해 영향을 받지 않았지만, Cas9-클레노우 융합에 의해 억제되었으며, 이는 클레노우 융합이 또한 MRE11-독립적 결실을 감소시킨다는 것을 시사한다. 클레노우를 Cas9에 융합시키는 것은 HEK293T 세포에서 *DMD* 엑손 53을 표적화할 때 9-bp 결실의 빈도를 감소시키지 않았다. 이러한 결실은 클레노우가 방해할 수 없는 프로세스에 의해 생성될 수 있다.

[0269] **[0242]** 불활성화된 중합효소 활성을 갖는 pol^{D705A} 및 클레노우^{D705A}는 pol I 및 클레노우 단편과 마찬가지로 결실 크기에 유사한 효과를 갖는 것으로 결정되었다. 이러한 단백질은 하기 2개의 비-배타적인 메커니즘을 통해 DNA 절제를 방해할 수 있다: 1) Cas9의 C-말단에 벌크 펩티드(≥629 AA)의 첨가는 DNA 절제 복합체 또는 조절 단백질의 동원을 방지할 수 있고, 2) pol^{D705A} 또는 클레노우^{D705A}의 잔류 DNA 결합 활성은 DNA 절제를 방해할 수 있다. pol^{D705A} 및 클레노우^{D705A}는 TIS의 백분율에 영향을 미치지 않기 때문에, 이들은 단지 작은 결실을 증가시키고 큰 결실을 감소시킬 필요가 있을 때 유용할 수 있다. DNA 절제에 대항하는 것이 관찰의 기초가 되는 메커니즘인 것으로 보이지만, 다른 세포 DNA 손상 복구 기계를 완전히 배제할 수는 없다.

[0270] **[0243]** TIS/비-TIS 비를 증가시키는 Cas9-pol 및 Cas9-클레노우의 능력은 중합효소 활성의 국소 이용가능성에 의존적이었다. 이러한 의존성은 TIS가 중합효소에 의해 Cas9-생성된 5' 오버행 말단을 채운 결과라는 관찰과 일치한다. 본원에 제공된 연구에서, 클레노우 단편은 2-bp 및 3-bp TIS를 증가시키는데 pol I보다 더 활성이었다(도 8). 이는 pol I의 5' 엑소뉴클레아제 도메인(클레노우 단편에 없음)이 5' 오버행에 대해 중합효소 도메인과 경쟁할 수 있음을 시사한다. 전자는 5' 오버행을 제거하고 결실을 선호하는 반면, 후자는 5' 오버행을 채워 TIS를 생산한다. 1-nt 5' 오버행을 제거하는 것은 또한 pol^{D705A}가 1-bp 결실을 증가시키는 메카니즘 중 하나일 수 있다.

[0271] **[0244]** Cas9-pol 및 Cas9-클레노우는 TIS만을 증가시켰지만 비-TIS는 증가시키지 않았다. IMR90 세포에서 *CLCN5* 또는 유전자내 부위 1을 표적화하는 경우, 클레노우를 Cas9에 융합시키면 전체 삽입이 크게 감소하였지만, TIS가 아닌 비-TIS만이 감소하였다. 인간 세포에서, DNA 중합효소 μ는 TIS 및 비-TIS 둘 모두를 생성하는데 필요하지만, 비-TIS를 생성하기 위해 다른 단백질이 필요한지 여부는 불분명하다. 클레노우 단편을 Cas9로 표적화하는 것은 TIS만을 증가시켰지만 비-TIS는 증가시키지 않았고(또는 일부 경우에, 비-TIS만 감소시켰지만 TIS는 아님), 이는 TIS 및 비-TIS가 상이한 메카니즘을 통해 생성되었음을 시사한다. pol 또는 클레노우 단편을 Cas9에 융합시키는 것은 5' 오버행의 채움을 촉진하지만, 비-TIS를 생성하는 원인이 되는 단백질 동원에 억제 효과를 가질 수 있다.

[0272] **[0245]** Cas9-클레노우는 인간 일차 세포에서 5개 사례 중 4개 사례에서 전체 INDEL 비율을 유의하게 증가시켰지만 HEK293T 세포에서는 그렇지 않았다. 이는 몇 가지 비배타적인 메카니즘에 의해 설명될 수 있다: 1) DNA 절제에 대항하고 DNA를 완벽하게 복구하는 상동성 재조합의 억제; 2) DNA 리가제가 INDEL을 생성하지 않으면서 상보적인 5' 오버행을 라이게이션하기 전에 5' 오버행을 채움; 및 3) Cas9는 일차 세포에서 P53-매개된 DNA 손상 스트레스를 유도하지만 HEK293T 세포에서는 유도하지 않으며, 클레노우를 Cas9에 융합시키는 것은 반복된 무익한 편집을 방지함으로써 스트레스를 감소시킴. HDR에 대한 클레노우를 Cas9에 융합시키는 효과는 조사되지 않았다. 그렇게 하는 것은 HDR에 필요한 DNA 절제에 대한 효과를 고려할 때 HDR에 대한 억제 효과를 가질 수 있다. 클레노우 도메인을 DSB로 표적화하는 것은 HDR에 독립적인 유전체 편집 적용에 유용할 수 있다.

[0273] **[0246]** 이들 연구는 클레노우 도메인을 RNA-가이드된 엔도뉴클레아제, 예를 들어, Cas9 엔도뉴클레아제에 융합시키는 것이 시험관내 및 생체내 적용에서 유전체 편집 안전성 및 효율을 개선시킬 수 있음을 보여준다. 클레노우 도메인을 Cas9에 융합시키면 예측할 수 없는 큰 표적-내 DNA 결실의 생성을 감소시켰고, 이는 다수의 그룹에 의해 관찰되었다(Owens et al. "Microhomologies are prevalent at Cas9-induced larger deletions. *Nucleic Acids Res,* 47, 7402-7417 (2019); Adikusuma et al. "Large deletions induced by Cas9 cleavage. *Nature*, 560, E8-E9 (2018); 및 Kosicki et al. "Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol*, 36, 765-771 (2018). 이는 일차 세포에서 유전체 편집 효율을 증가시켰고, DNA 치환 비율 또는 표적-외 비율을 증가시키지 않았다. 또한, 1-bp 결실 및 TIS

에 대한 효과는 질환-유발 유전자를 파괴하거나 리프레이밍(reframing)에 의해 파괴된 유전자를 복원하는 효율을 개선시키기 위해 바람직한 유형의 돌연변이의 백분율을 증가시키는데 사용될 수 있다.

[0274] **서열**

SEQ ID NO:1	MS2 코트 단백질(MCP) DNA 서열	ATGGCTTCTAACTTTACTCAGTTCGTTCTCGTCGACAATGGCGGAACTGG CGACGTGACTGTCGCCCCAAGCAACTTCGCTAACGGGATCGCTGAATG GATCAGCTCTAACTCGCGTTCACAGGCTTACAAAGTAACCTGTAGCGTT CGTCAGAGCTCTGCGCAGAATCGCAAATACACCATCAAAGTCGAGGTG CCTAAAGGCGCCTGGCGTTCGTACTTAAATATGGAACTAACCATTCCAA TTTTCGCCACGAATTCCGACTGCGAGCTTATTGTTAAGGCAATGCAAGG TCTCCTAAAAGATGGAAACCCGATTCCCTCAGCAATCGCAGCAAACTCC GGCATCTAC
SEQ ID NO:2	MS2 코트 단백질(MCP) 아미노산 서열	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKD GNPIPSAIAANSGIY
SEQ ID NO:3	PP7 코트 단백질(PCP) DNA 서열	tccaaaacaatagtcctctccgtaggggaggcaacacggactttgaccgaaatccagtcaacc gctgaccgacaaatctttgaagagaaagtagggcctcttgtgggccgactgcgcttgactgca agcttgcgacaaaacggcgcaaagactgcctatagggtcaaccttaaactcgaccaagccga cgtggtcgatagcggtctccctaaggttcggtatacgcaggtctggagtcatgacgtaacaat cgtagcaaacagcacagaagcctcccgaaaaagcctctacgatctgacgaaatccttggtgg ctacgtcacaggtggaagacctcgttgtcaaccttgtacctctgggtcga
SEQ ID NO:3	PP7 코트 단백질(PCP) 아미노산 서열	SKTIVLSVGEATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYR VNLKLDQADVVDSGLPKVRYTQVWSHDVTIVANSTEASRKSLYDLTKSLVA TSQVEDLVVNLVPLGR
SEQ ID NO:5	람다 N RNA-결합 도메인(위치(1-22) DNA 서열	ATGGATGCACAAACACGCCGCCGCGAACGTCGCGCAGAGAAACAGGC TCAATGGAAAGCAGCAAAT
SEQ ID NO:6	람다 N RNA-결합 도메인(위치(1-22) 아미노산 서열	MDAQTRRRERRAEKQAQWKAAN

[0275]

SEQ ID NO:7	Com 단백질 DNA 서열	atgaaatcaattcgctgtaaaaactgcaacaaactgttatttaaggcggattcctttgatcacatt gaaatcaggtgtccgcgttgcaaacgtcacatcataatgctgaatgcctgcgagcatcccacg gagaaacattgtgggaaaagagaaaaaatcacgcattctgacgaaaccgtgcgttattgagt at
SEQ ID NO:8	Com 단백질 아미노산 서열(GenBank AAF01130.1)	MKSIRCKNCNKLLFKADSFDHIEIRCPRCKRHIIMLNACEHPTEKHCGKREKI THSDETVRY
SEQ ID NO:9	MS2 앱타머 서열(RNA)	ACAUGAGGAUCACCCAUGU
SEQ ID NO:10	MS2 앱타머 서열(DNA)	ACATGAGGATCACCCATGT
SEQ ID NO:11	PP7 앱타머 서열(RNA)	GGAGCAGACGAUAUGGCGUCGCUCC
SEQ ID NO:12	PP7 앱타머 서열(DNA)	GGAGCAGACGATATGGCGTCGCTCC
SEQ ID NO:13	박스-B; 람다 N RNA-결합 도메인 앱타머 서열(RNA)	GGGCCCUGAAGAAGGGCCC
SEQ ID NO:14	박스-B; 람다 N RNA-결합 도메인 앱타머 서열(DNA)	GGGCCCTGAAGAAGGGCCC
SEQ ID NO:15	com 앱타머 RNA 서열	CUGAAUGCCUGCGAGCAUC

[0276]

SEQ ID NO:16	com 앱타머 DNA 서열	CTGAATGCCTGCGAGCAT
SEQ ID NO:17	인간 베타 헤모글로빈(HBB) 3' UTR(DNA)	gctcgctttcttgctgtccaatttctattaaaggttcctttgttccctaagtccaactactaaactgg gggatattatgaagggccttgagcatctggattctgcctaataaaaaacatttattttcattgc
SEQ ID NO:18	인간 베타 헤모글로빈(HBB) 3' UTR(RNA)	gcucgcuuucuugcuguccaauuucuauuaaagguuccuuuguucccuaaguccaac uacuaaacugggggauauuaugaagggccuugagcaucuggauucugccuaauaaa aaacauuuauuuucauugc
SEQ ID NO:19	HIV-1 뉴클레오캡시드(NC) DNA 서열	ATACAGAAAGGCAATTTTAGGAACCAAAGAAAGACTGTTAAGTGTTTCA ATTGTGGCAAAGAAGGGCACATAGCCAAAAATTGCAGGGCCCCTAGGA AAAAGGGCTGTTGGAAATGTGGAAAGGAAGGACACCAAATGAAAGATT GTACTGAGAGACAGGCTAAT
SEQ ID NO:20	HIV-1 뉴클레오캡시드(NC) 아미노산 서열	IQKGNFRNQRKTVKCFNCGKEGHIAKNCRAPRKKGCWKCGKEGHQMKD CTERQAN
SEQ ID NO:21	HIV-1 매트릭스 단백질(MA) DNA 서열	atgggtgcgagagcgtcagtattaagcgggggagaattagatcgatgggaaaaaattcggt taaggccagggggaaagaaaaatataaattaaaacatatagtatgggcaagcagggagct agaacgattcgcagttaatcctggcctgttagaaacatcagaaggctgtagacaaatactggg acagctacaaccatcccttcagacaggatcagaagaacttagatcattatataatacagtagca accctctattgtgtgcatcaaaggatagagataaaagacaccaaggaagctttagacaagata gaggaagagcaaaacaaaagtaagaaaaaagcacagcaagcagcagctgacacaggaca cagcaatcaggtcagccaaaattac
SEQ ID NO:22	HIV-1 매트릭스 단백질(MA) 아미노산 서열	GARASVLSGGELDRWEKIRLRPGGKKKYKLKHIVWASRELERFAVNPGLLET SEGCRQILGQLQPSLQTGSEELRSLYNTVATLYCVHQRIEIKDTKEALDKIEEE QNKSKKKAQQAAADTGHSNQVSQNY
SEQ ID NO:23	HIV-1 바이러스 단백질(VPR) DNA 서열	ATGGAACAAGCCCCAGAAGACCAGGGACCGCAGAGGGAACCATACAA TGAATGGACACTAGAACTTTTAGAGGAACTCAAGCGGGAAGCAGTCAG ACACTTTCCTAGACCATGGCTTCATGGCTTAGGACAACATATCTATGAAA CCTATGGAGATACTTGGACGGGGGTGGAAGCTATAATAAGAATTCTGC

[0277]

		AACGACTACTGTTTGTCCATTTCAGAATTGGGTGCCAGCATAGCCGAAT AGGCATTCTAAGACAGAGAAGAGCAAGAAATGGAGCCAGTAGATCCTA A
SEQ ID NO:24	HIV-1 바이러스 단백질(VPR) 아미노산 서열	MEQAPEDQGPQREPYNEWTLELLEELKREAVRHFPRPWLHGLGQHIYETY GDTWTGVEAIIRILQRLLFVHFRIGCQHSRIGILRQRRARNGASRS
SEQ ID NO:25	바이러스 코어에서 패키징을 향상시키기 위한 코돈 변화를 갖는 HIV-1 음성 조절 인자(NEF) DNA 서열(G3C, V153L, 및 E177G 돌연변이; 밑줄)	atgggtTgcaagtggtcaaaaagtagtgtgattggatggcctgctgtaagggaaagaatga gacgagctgagccagcagcagatggggtgggagcagtatctcgagacctagaaaaacatg gagcaatcacaagtagcaatacagcagctaacaatgctgcttgtgcctggctagaagcacaa gaggaggaagaggtgggttttccagtcacacctcaggtacctttaagaccaatgacttacaag gcagctgtagatcttagccactttttaaaagaaaaggggggactggaagggctaattcactcc caaagaagacaagatatccttgatctgtggatctaccacacacaaggctacttccctgattggc agaactacacaccagggccaggggtcagatatccactgacctttggatggtgctacaagcta gtaccagttgagccagataagCtGgaagaggccaataaaggagagaacaccagcttgttac accctgtgagcctgcatggaatggatgaccctgGAagagaagtgttagagtggaggtttga cagccgcctagcatttcatcacgtggcccgagagctgcatccggagtacttcaagaactgc (황색 위치는 seq ID. 8에 설명된 변화를 코딩하기 위해 변화된다.
SEQ ID NO:26	바이러스 코어에서 패키징을 향상시키기 위한 돌연변이를 갖는 HIV-1 음성 조절 인자(NEF) 아미노산	MGCKWSKSSVIGWPAVRERMRRAEPAADGVGAVSRDLEKHGAITSSNTA ANNAACAWLEAQEEEEVGFPVTPQVPLRPMTYKAAVDLSHFLKEKGGLEG LIHSQRRQDILDLWIYHTQGYFPDWQNYTPGPGVRYPLTFGWCYKLVPVEP DKLEEANKGENTSLLHPVSLHGMDDPGREVLEWRFDSRLAFHHVARELHP EYFKNC

[0278]

	서열(G3C, V153L, 및 E177G 돌연변이; 밑줄)	
SEQ ID NO:27	DNA 중합효소 I(DNA pol I)	VQIPQNPLILVDGSSYLYRAYHAFPPLTNSAGEPTGAMYGVLNMLRSLIMQYKPTH AAVVFDAKGKTFRDELFEHYKSHRPPMPDDLRAQIEPLHAMVKAMGLPLLAVSGV EADDVIGTLAREAEKAGRPVLISTGDKDMAQLVTPNITLINTMTNTILGPEEVVNKY GVPPELIIDFLALMGDSSDNIPGVPGVGEKTAQALLQGLGGLDTLYAEPEKIAGLSFR GAKTMAAKLEQNKEVAYLSYQLATIKTDVELELTCEQLEVQQPAAEELLGLFKKYEFK RWTADVEAGKWLQAKGAKPAAKPQETSVADEAPEVTATVISYDNYVTILDEETLKA WIAKLEKAPVFAFDTETDSLDNISANLVGLSFAIEPGVAAYIPVAHDYLDAPDQISRE RALELLKPLLEDEKALKVGQNLKYDRGILANYGIELRGIAFDTMLESYILNSVAGRHD MDSLAERWLKHKTITFEEIAGKGKNQLTFNQIALEEAGRYAAEDADVTLQLHLKMW PDLQKHKGPLNVFENIEMPLVPVLSRIERNGVKIDPKVLHNHSEELTLRLAELEKKAH EIAGEEFNLSSTKQLQTILFEKQGIKPLKKTPGGAPSTSEEVLEELALDYPLPKVILEYRG LAKLKSTYTDKLPLMINPKTGRVHTSYHQAVTATGRLSSTDPNLQNIPVRNEEGRRI RQAFIAPEDYVIVSADYSQIELRIMAHLSRDKGLLTAFAEGKDIHRATAAEVFGLPLET VTSEQRRSAKAINFGLIYGMSAFGLARQLNIPRKEAQKYMDLYFERYPGVLEYMERT RAQAKEQGYVETLDGRRLYLPDIKSSNGARRAAAERAAINAPMQGTAADIIKRAMI AVDAWLQAEQPRVRMIMQVHDELVFEVHKDDVDAVAKQIHQLMENCTRLDVPL LVEVGSGENWDQAH
SEQ ID NO:28	DNA Pol I 의 클레노우 단편	APEVTATVISYDNYVTILDEETLKAWIAKLEKAPVFAFDTETDSLDNISANLVGLSFAIE PGVAAYIPVAHDYLDAPDQISRERALELLKPLLEDEKALKVGQNLKYDRGILANYGIE LRGIAFDTMLESYILNSVAGRHDMDSLAERWLKHKTITFEEIAGKGKNQLTFNQIALE EAGRYAAEDADVTLQLHLKMWPDLQKHKGPLNVFENIEMPLVPVLSRIERNGVKID PKVLHNHSEELTLRLAELEKKAHEIAGEEFNLSSTKQLQTILFEKQGIKPLKKTPGGAPS TSEEVLEELALDYPLPKVILEYRGLAKLKSTYTDKLPLMINPKTGRVHTSYHQAVTATG RLSSTDPNLQNIPVRNEEGRRIRQAFIAPEDYVIVSADYSQIELRIMAHLSRDKGLLTA FAEGKDIHRATAAEVFGLPLETVTSEQRRSAKAINFGLIYGMSAFGLARQLNIPRKEA QKYMDLYFERYPGVLEYMERTRAQAKEQGYVETLDGRRLYLPDIKSSNGARRAAAE RAAINAPMQGTAADIIKRAMIAVDAWLQAEQPRVRMIMQVHDELVFEVHKDDVD AVAKQIHQLMENCTRLDVPLLVEVGSGENWDQAH
SEQ ID NO:29	Cas9	MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPAADKKYSIGLDIGTNS VGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT RRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEK YPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLV QTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL

[0279]

		TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYI DGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAI LRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEV VDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPA FLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL LKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYT GWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTT QKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRM NTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA LIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANG EIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPK RNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMER SSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPS KYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDK VLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDKRPAATKKAGQAKKKK
SEQ ID NO: 30	Cpf1	MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQ FFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFK NLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYF KGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIKKDL AEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGENTK RKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQS FYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIG TAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCR FEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQ TNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYITQKPYS DEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFDDKAIKE NKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKG YEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLT FENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVY KLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKRFTEDKFFFH CPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRGERHLAYYTLVDGKGNIIKQDT FNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQVVHEIAKLVI EYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRA YQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFD

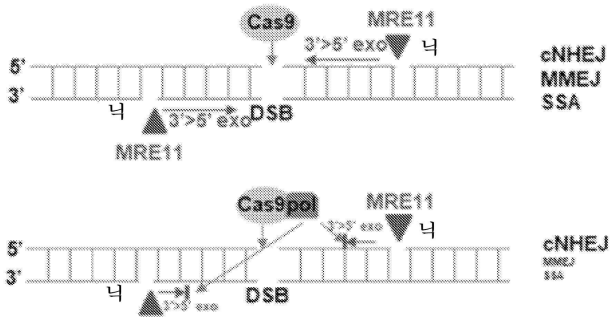
[0280]

		KICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKNHNWDTREVY PTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQMRNSKTGTELDYL ISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKGLMLLGRIKNNQEGKKLN LVIKNEEYFEFVQNRNN

[0281]

도면

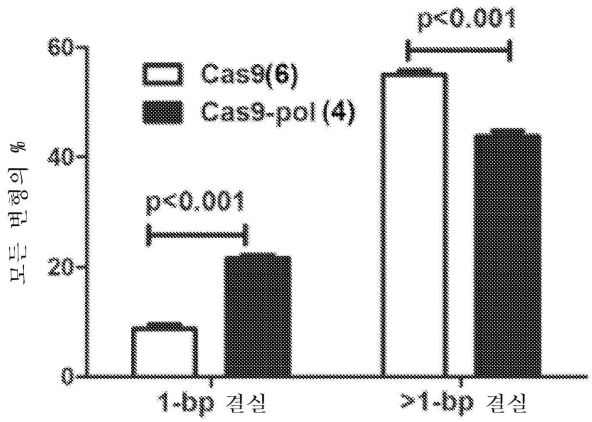
도면1a



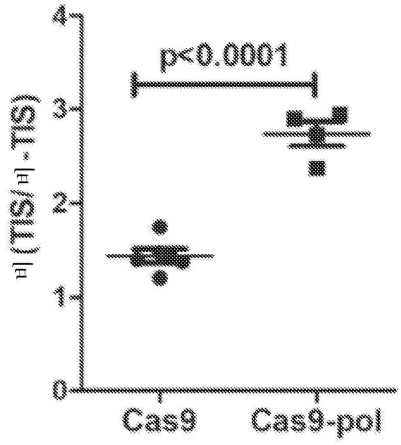
도면1b



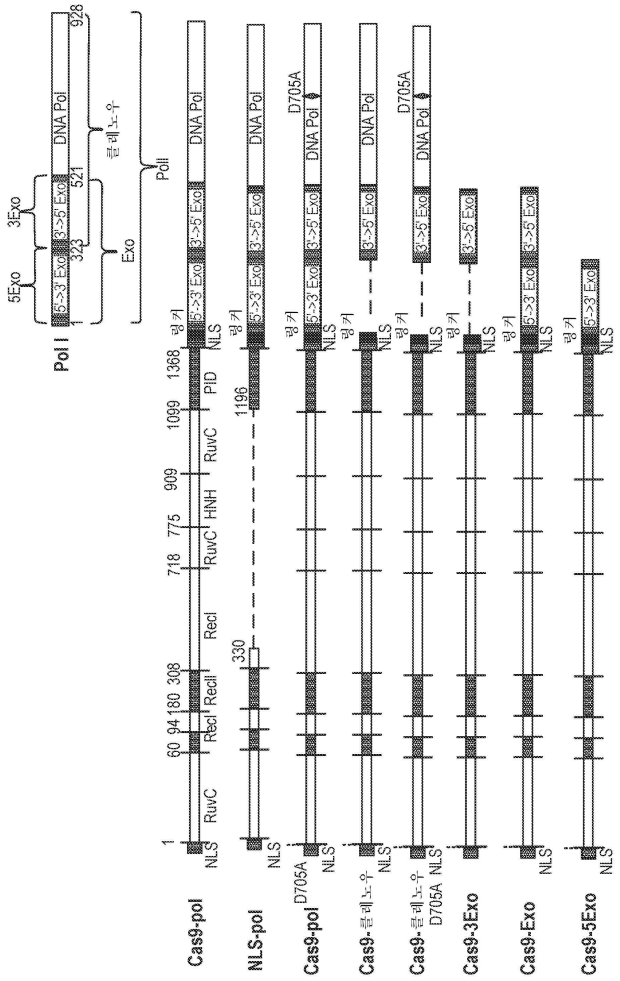
도면1c



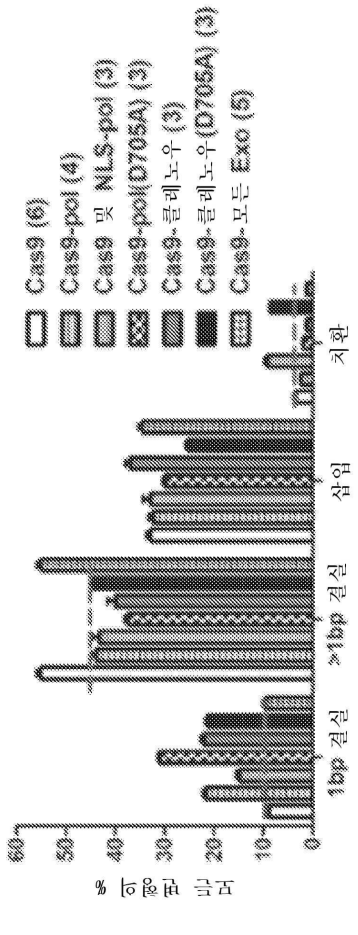
도면1d



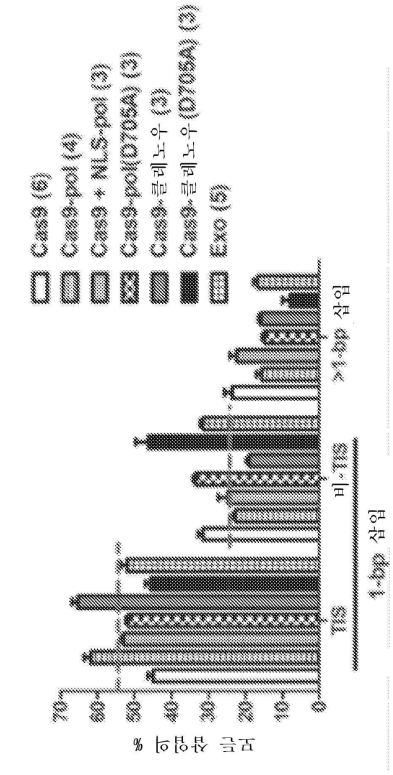
도면2a


Pol I
5Exo
3Exo
Exo
클레노우
PolI
DNA Pol
3'->5' Exo
5'->3' Exo
링커
NLS
RuvC
RecI
RecII
RecI
HNH
RuvC
PID
RuvC
D705A
Cas9-pol
NLS-pol
Cas9-pol
Cas9-클레노우
Cas9-클레노우
Cas9-3Exo
Cas9-Exo
Cas9-5Exo

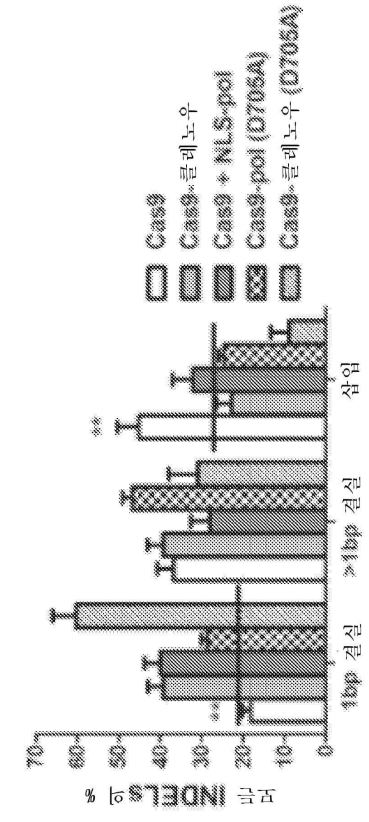
도면2b


Cas9 (6)
Cas9-pol (4)
Cas9 및 NLS-pol (3)
Cas9-pol(D705A) (3)
Cas9-클레노우 (3)
Cas9-클레노우(D705A) (3)
Cas9-모든 Exo (5)
모든 변형의 %
0
10
20
30
40
50
60
1bp 결실
>1bp 결실
삽입
치환

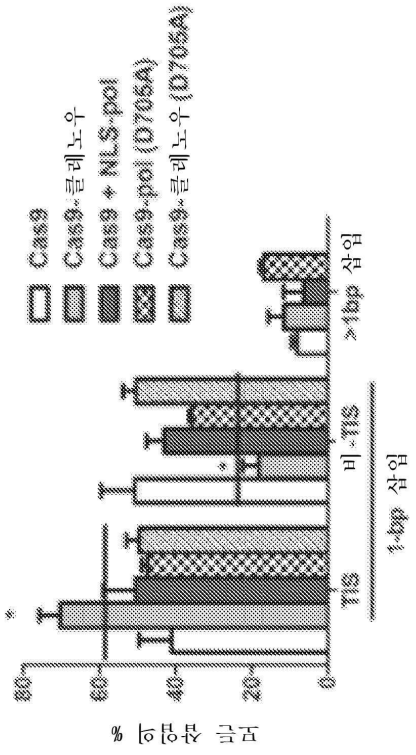
도면2c


Cas9 (6)
Cas9-pol (4)
Cas9 + NLS-pol (3)
Cas9-pol(D705A) (3)
Cas9-클레노우 (3)
Cas9-클레노우(D705A) (3)
Exo (5)
모든 삽입의 %
70
60
50
40
30
20
10
0
TIS
1-bp 삽입
비-TIS
>1-bp 삽입

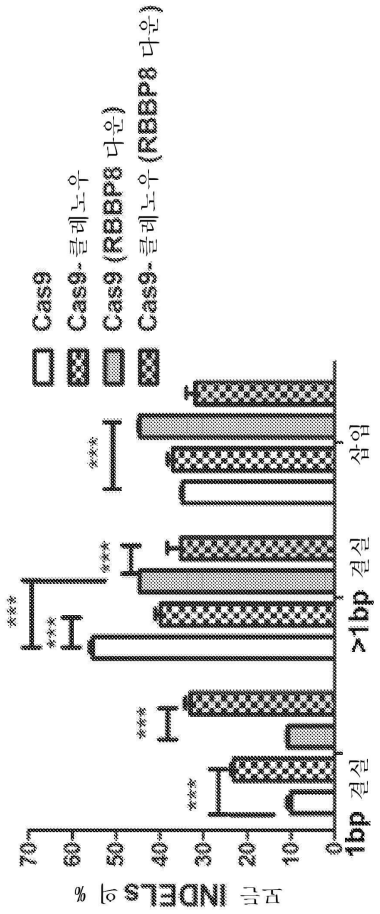
도면2d


Cas9
Cas9-클레노우
Cas9 + NLS-pol
Cas9-pol (D705A)
Cas9-클레노우 (D705A)
모든 INDELs의 %
70
60
50
40
30
20
10
0
1bp 결실
>1bp 결실
삽입

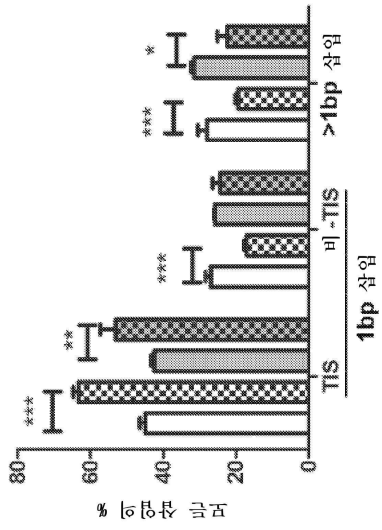
도면2e


Cas9
Cas9-클레노우
Cas9 + NLS-pol
Cas9-pol (D705A)
Cas9-클레노우 (D705A)
모든 삽입의 %
60
40
20
0
*
YIS
1-bp 삽입
비-YIS
>1bp 삽입

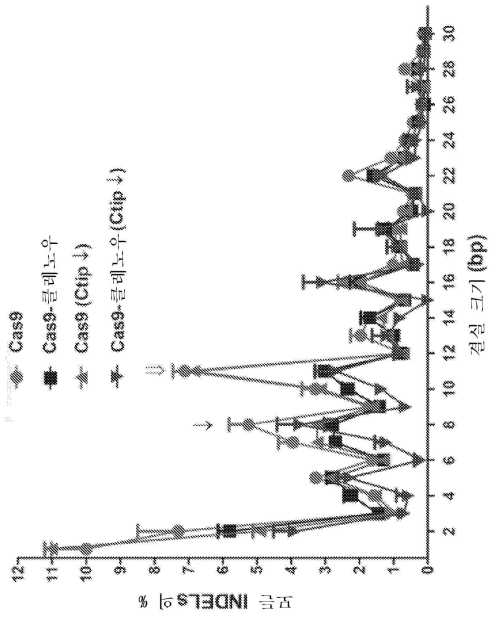
도면3a


Cas9
Cas9-클레노우
Cas9 (RBBP8 다운)
Cas9-클레노우 (RBBP8 다운)
모든 INDELs의 %
0
10
20
30
40
50
60
70
1bp 결실
>1bp 결실
삽입
***
**

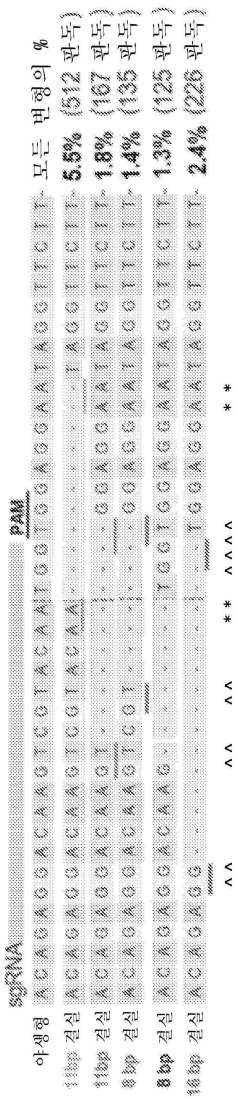
도면3b


모든 삽입의 %
0
20
40
60
80
TIS
1bp 삽입
비-TIS
>1bp 삽입
***
**
*

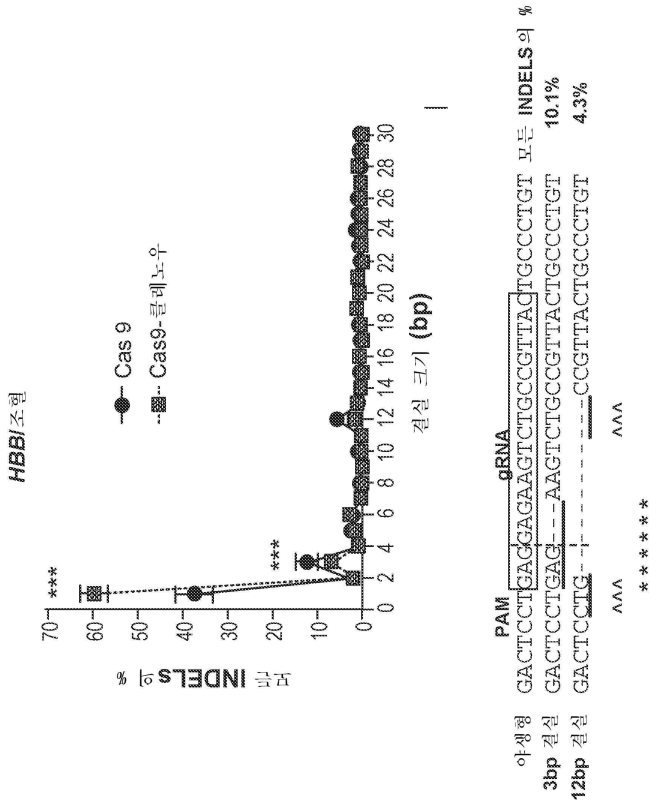
도면3c



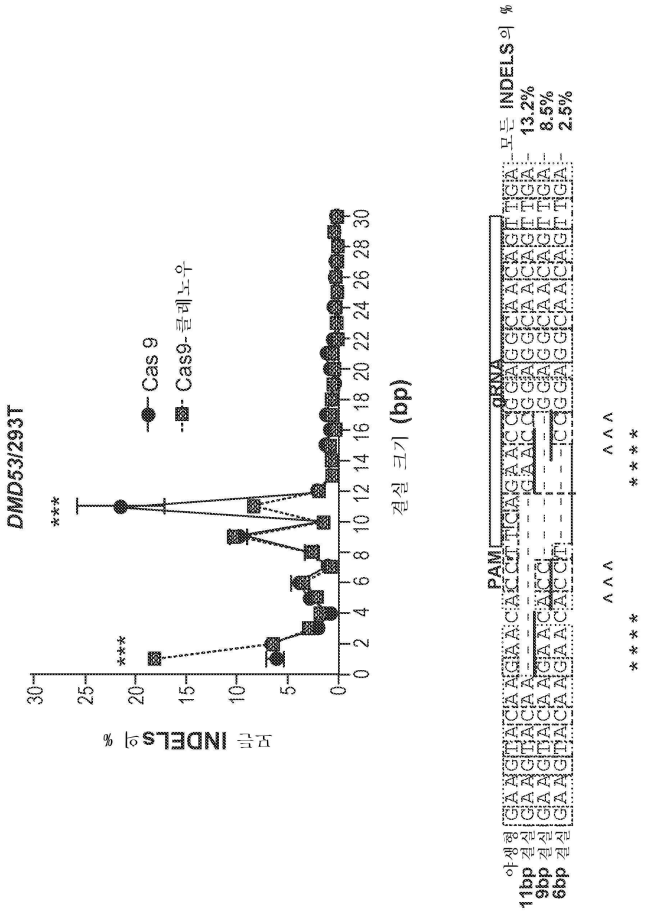
도면3d


sgRNA
PAM
야생형 ACAGAGGACGAAGTCGTACAATGGTGGAGGAATAGGTTCTT
11bp 결실 ACAGAGGACGAAGTCGTACAA---------TAGGTTCTT 5.5% (512 판독)
11bp 결실 ACAGAGGACGAAGT----------GGAGGAATAGGTTCTT 1.8% (167 판독)
8bp 결실 ACAGAGGACGAAGT--------GGAGGAATAGGTTCTT 1.4% (135 판독)
8 bp 결실 ACAGAGGACGAAG----------TGGTGGAGGAATAGGTTCTT 1.3% (125 판독)
16 bp 결실 ACAGAGG----------TGGTGGAGGAATAGGTTCTT 2.4% (226 판독)

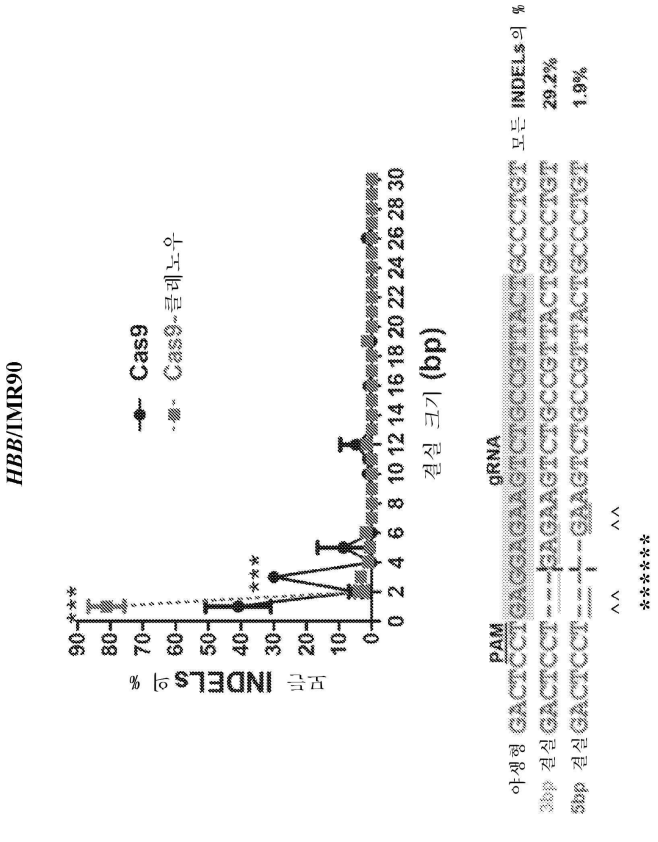
도면4a


HBB1조절
Cas 9
Cas9-클레노우
모든 INDELS의 %
결실 크기 (bp)
***
***
PAM
gRNA
야생형 GACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGT 모든 INDELS의 %
3bp 결실 GACTCCTGAG---AAGTCTGCCGTTACTGCCCTGT 10.1%
12bp 결실 GACTCCTG----------CCGTTACTGCCCTGT 4.3%

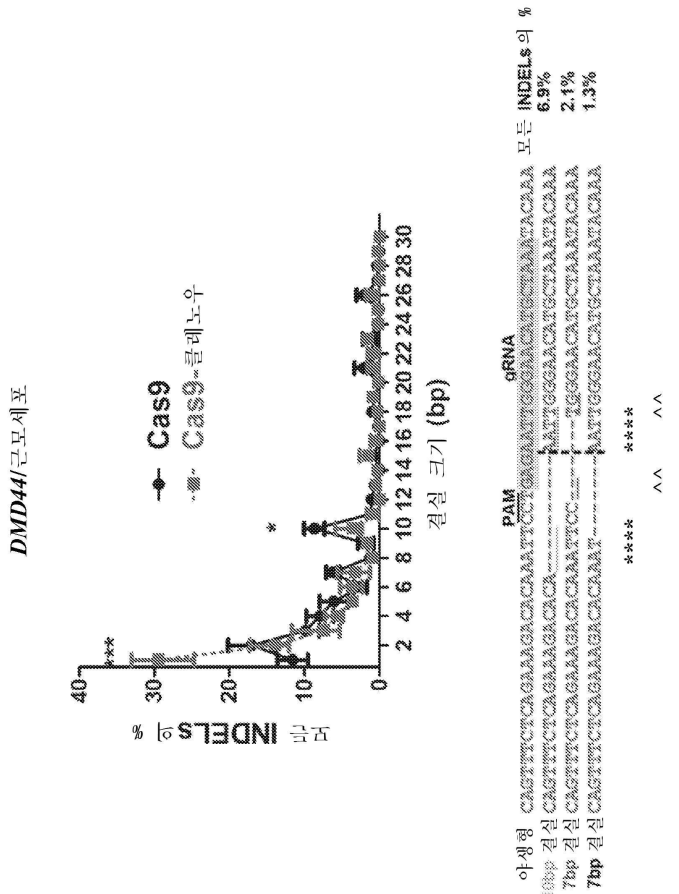
도면4b


DMD53/293T
***
***
Cas 9
Cas9-클레노우
모든 INDELs의 %
결실 크기 (bp)
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
0 5 10 15 20 25 30
야생형 GAAGTTACACAAGAAACCCTTCAGAACGGAGGCACACAGTTGA
11bp 결실 GAAGTTACACAA----------GAACGGAGGCACACAGTTGA
9bp 결실 GAAGTTACACAAGAA---------GAACGGAGGCACACAGTTGA
6bp 결실 GAAGTTACACAAGAAAC------CGGAGGCACACAGTTGA
PAM
gRNA
모든 INDELs의 %
13.2%
8.5%
2.5%
* * * *
* * * *

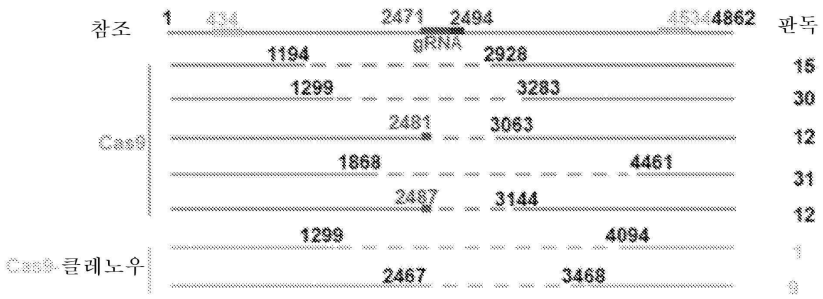
도면4c


HBB/IMR90
Cas9
Cas9-클레노-유
모든 INDELs의 %
***
***
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30
결실 크기 (bp)
gRNA
PAM
야생형 GACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGT 모든 INDELs의 %
3bp 결실 GACTCCT---AGAAGTCTGCCGTTACTGCCCTGT 29.2%
5bp 결실 GACTCCT-----GAAGTCTGCCGTTACTGCCCTGT 1.9%
^^
^^
****

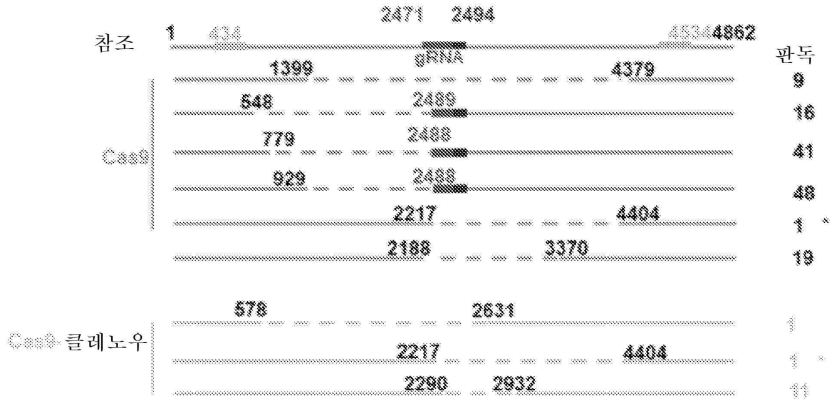
도면4d


DMD44/근모세포
Cas9
Cas9-클레노우
결실 크기 (bp)
모든 INDELs의 %
PAM
gRNA
야생형 CAGTTTCTCAGAAAGACACAAATTCCTGGCAATTGGGAACATGCTAAATACAAA
10bp 결실 CAGTTTCTCAGAAAGACACA----------ATTGGGAACATGCTAAATACAAA
7bp 결실 CAGTTTCTCAGAAAGACACAAATTCC-----GGGAACATGCTAAATACAAA
7bp 결실 CAGTTTCTCAGAAAGACACAAAT----------ATTGGGAACATGCTAAATACAAA
모든 INDELs의 %
6.9%
2.1%
1.3%

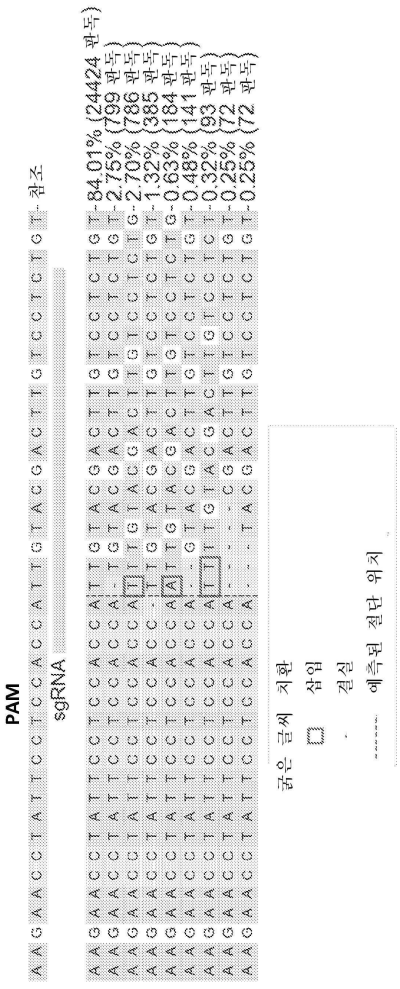
도면5a


참조
1
434
2471
2494
4534
4862
gRNA
판독
Cas9
Cas9-클레노우
1194
2928
15
1299
3283
30
2481
3063
12
1868
4461
31
2487
3144
12
1299
4094
1
2467
3468
9

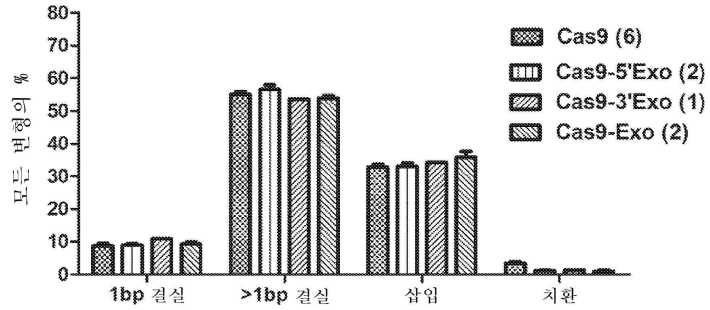
도면5b


참조
2471 2494
1 434
4534 4962
gRNA
판독
1399
4379
9
548
2488
16
Cas9
779
2488
41
929
2488
48
2217
4404
1
2188
3370
19
578
2631
Cas9 클레노우
2217
4404
2290
2932

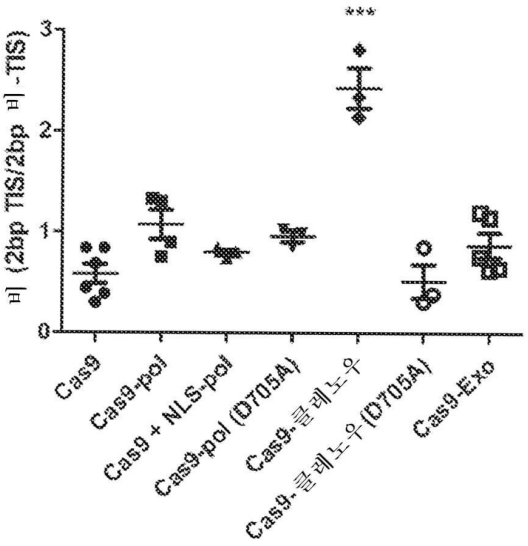
도면6


PAM
sgRNA
참조
-84.01% (24424 판독)
-2.75% (799 판독)
-2.70% (786 판독)
-2.70% (385 판독)
-1.32% (184 판독)
-0.63% (141 판독)
-0.48% (93 판독)
-0.32% (72 판독)
-0.25% (72 판독)
-0.25% (72 판독)
굵은 글씨 치환
삽입
결실
예측된 절단 위치

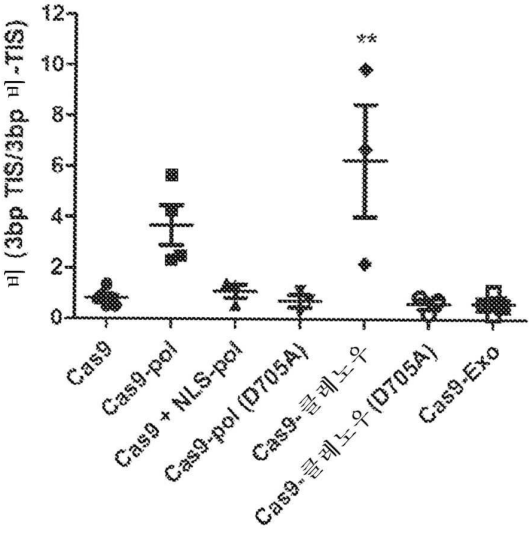
도면7



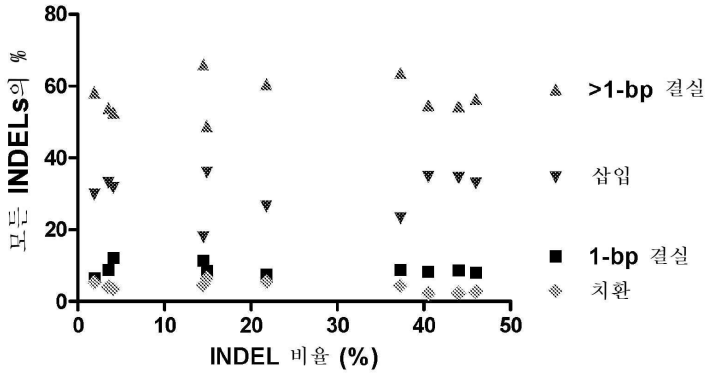
도면8a



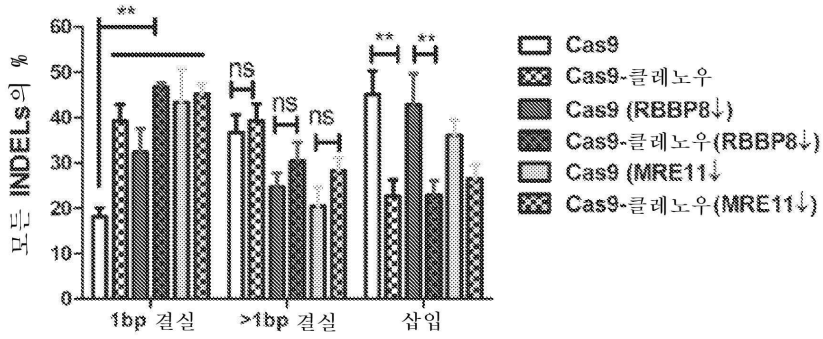
도면8b



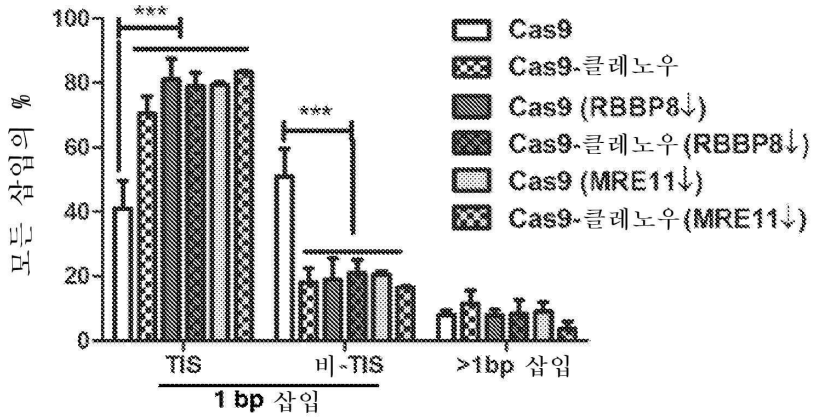
도면9



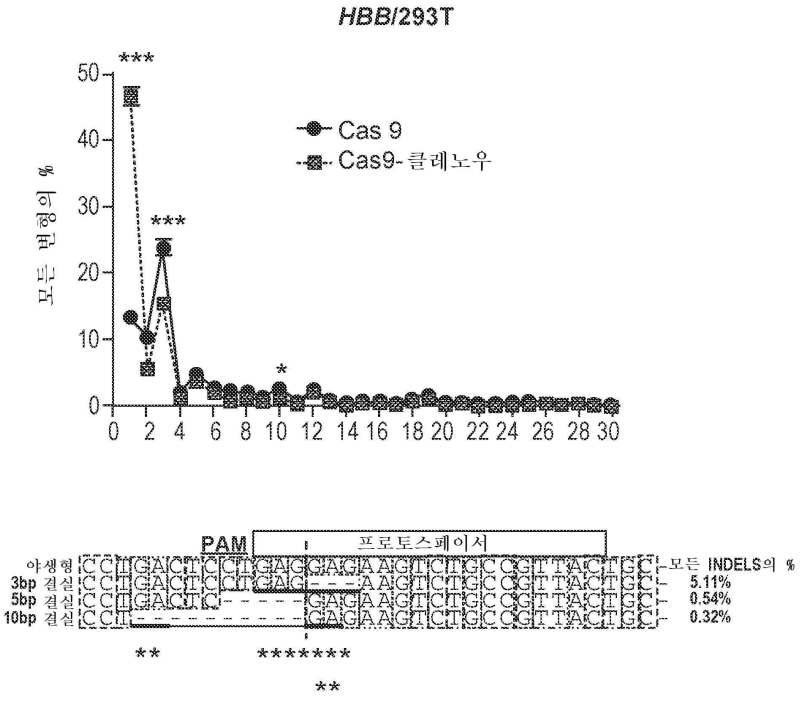
도면10a



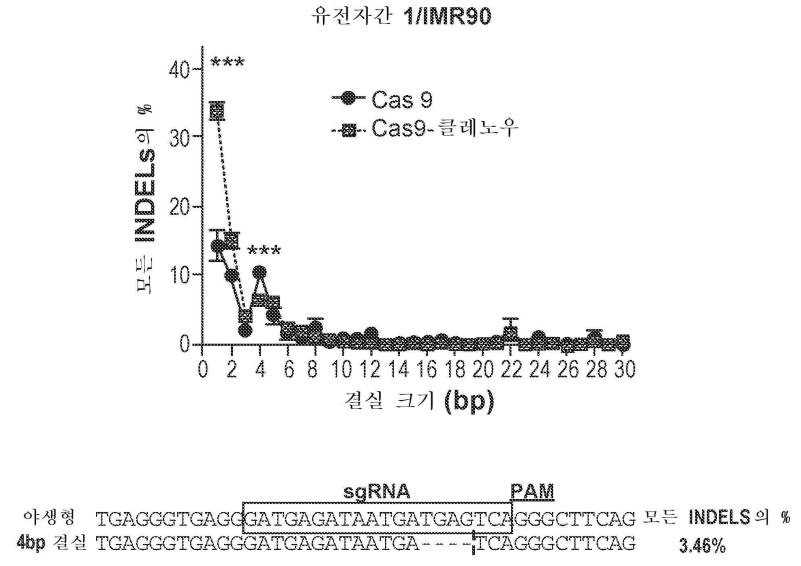
도면10b



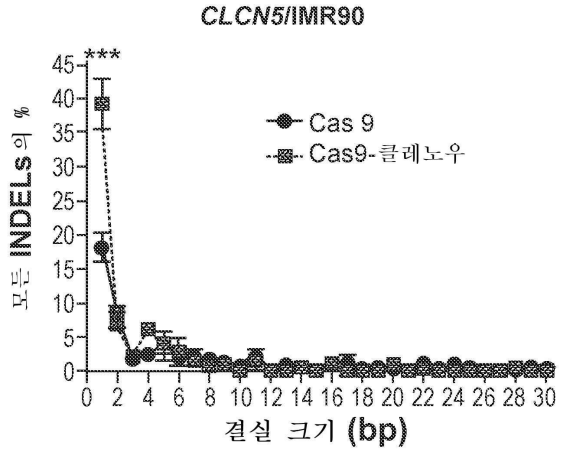
도면11a

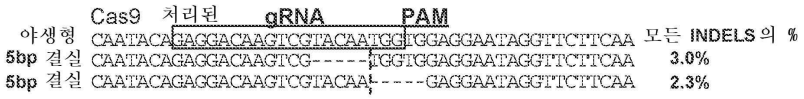

HBB/293T
Cas 9
Cas9-클레노우
모든 변형의 %
PAM
프로토스페이서
야생형
3bp 결실
5bp 결실
10bp 결실
모든 INDELS의 %
5.11%
0.54%
0.32%

도면11b


유전자간 1/IMR90
Cas 9
Cas9-클레노우
모든 INDELs의 %
결실 크기 (bp)
sgRNA
PAM
야생형 TGAGGGTGAGGGATGAGATAATGATGAGTCAGGGCTTCAG
4bp 결실 TGAGGGTGAGGGATGAGATAATGA----TCAGGGCTTCAG
모든 INDELS의 %
3.46%

도면11c


CLCN5/IMR90
***
모든 INDELs의 %
Cas 9
Cas9-클레노우
결실 크기 (bp)


Cas9 처리된
gRNA
PAM
야생형 CAATACAGAGGACAAGTCGTACAATGGTGGAGGAATAGGTTCTTCAA 모든 INDELS의 %
5bp 결실 CAATACAGAGGACAAGTCG-----TGGTGGAGGAATAGGTTCTTCAA 3.0%
5bp 결실 CAATACAGAGGACAAGTCGTACAA----GAGGAATAGGTTCTTCAA 2.3%

서 열 목 록 (첨부)



아이콘을 클릭하시면 서열목록 파일이 열립니다.

본 공보 PDF는 첨부파일을 가지고 있습니다. Acrobat Reader PDF뷰어를 제공하지 않는 브라우저(크롬, 파이어폭스, 사파리 등)의 경우 첨부파일 열기가 제한되어 있으므로 Acrobat Reader PDF뷰어 설치 후 공보 PDF를 다운로드 받아 해당 뷰어에서 조회해주시기 바랍니다.