



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2022년03월17일

(11) 등록번호 10-2374379

(24) 등록일자 2022년03월10일

(51) 국제특허분류(Int. Cl.)

C12N 15/90 (2006.01) A01K 67/027 (2006.01)

C07K 14/725 (2006.01) C12N 15/64 (2006.01)

C12N 15/85 (2006.01)

(52) CPC특허분류

C12N 15/907 (2013.01)

A01K 67/0278 (2013.01)

(21) 출원번호 10-2016-7037057

(22) 출원일자(국제) 2015년06월05일

심사청구일자 2020년05월28일

(85) 번역문제출일자 2016년12월30일

(65) 공개번호 10-2017-0027743

(43) 공개일자 2017년03월10일

(86) 국제출원번호 PCT/US2015/034503

(87) 국제공개번호 WO 2015/188109

국제공개일자 2015년12월10일

(30) 우선권주장

62/008,832 2014년06월06일 미국(US)

62/017,916 2014년06월27일 미국(US)

(56) 선행기술조사문헌

WO2013063361 A1\*

WO2013163394 A1\*

Nat Protoc. 8(11):2281-2308 (2013.11.)\*

\*는 심사관에 의하여 인용된 문헌

(73) 특허권자

리제너론 파마슈티칼스 인코포레이티드

미합중국 뉴욕주 10591 타리타운 올드 소우 밀 리버 로드 777

(72) 발명자

아우어바흐 보이테크

미국 뉴욕 10591 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내

프렌듀이 데이비드

미국 뉴욕 10591 테리타운 올드 소우 밀 리버 로드 777 리제너론 파마슈티칼스 인코포레이티드 내

(뒷면에 계속)

(74) 대리인

특허법인와이에스장

전체 청구항 수 : 총 37 항

심사관 : 최성호

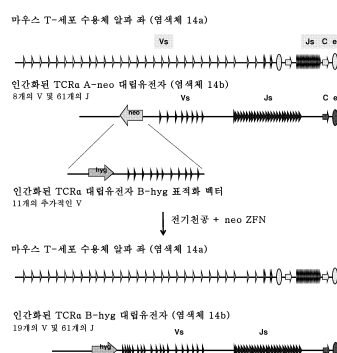
(54) 발명의 명칭 표적화된 좌를 변형시키는 방법 및 조성물

## (57) 요약

세포에서 하나 이상의 표적 좌(locus)를 변형시키는 방법 및 조성물이 제공된다. 그러한 방법은 세포를 제공하는 단계를 포함하고, 세포는 세포에서 활성인 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 제1 폴리뉴클레오티드를 포함하고, 제1 폴리뉴클레오티드는 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 추가

(뒷면에 계속)

## 대표도 - 도1



로 포함한다. 제1 뉴클레아제 작용제가 세포 내로 도입되고, 제1 뉴클레아제 작용제는 제1 인식 부위에 닉(nick) 또는 이중-가닥 브레이크(break)를 유도한다. 제1 인식 부위에 충분히 가깝게 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암(arm)에 의해 플랭킹된(flanked) 제1 삽입 폴리뉴클레오티드를 포함하는 제1 표적화 벡터가 추가로 세포 내로 도입된다. 이어서, 표적 좌에 통합된 제1 삽입 폴리뉴클레오티드를 게놈 내에 포함하는 적어도 하나의 세포가 확인된다.

(52) CPC특허분류

**C07K 14/7051** (2013.01)

**C12N 15/64** (2013.01)

**C12N 15/8509** (2013.01)

**A01K 2217/072** (2013.01)

**A01K 2267/0387** (2013.01)

**C12N 2015/8527** (2013.01)

(72) 발명자

**드로젯 구스타보**

미국 뉴욕 10591 테리타운 올드 소우 밀 리버 로드  
777 리제너론 파마슈티칼스 인코포레이티드 내

**가글리아르디 앤써니**

미국 뉴욕 10591 테리타운 올드 소우 밀 리버 로드  
777 리제너론 파마슈티칼스 인코포레이티드 내

**쿠노 준코**

미국 뉴욕 10591 테리타운 올드 소우 밀 리버 로드  
777 리제너론 파마슈티칼스 인코포레이티드 내

**발렌주엘라 데이비드 엠.**

미국 뉴욕 10591 테리타운 올드 소우 밀 리버 로드  
777 리제너론 파마슈티칼스 인코포레이티드 내

## 명세서

### 청구범위

#### 청구항 1

세포에서 표적 좌를 연속적으로 변형시키는 방법으로서,

(a) (1) 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 핵산; 및 (2) 제1 뉴클레아제 작용제에 대한 제1 뉴클레아제 인식 부위를 포함하는 제1 선택 카세트를 포함하는 표적 좌를 포함하는 세포를 제공하는 단계로서, 여기서 상기 제1 뉴클레아제 작용제는 징크 핑거 뉴클레아제 (ZFN), 전사 활성화인자-유사 이펙터 뉴클레아제 (TALEN), 또는 클러스터화된 규칙적으로 사이 공간을 둔 짧은 회문성 반복 (CRISPR)-관련 (Cas) 단백질 및 가이드 RNA (gRNA)이고, 상기 제1 뉴클레아제 인식 부위는 제1 선택 마커의 코딩 영역 또는 제1 선택 마커의 임의의 비-단백질 코딩 영역에 위치하는, 단계;

(b) (i) 제1 뉴클레아제 인식 부위에 닉 (nick) 또는 이중-가닥 브레이크 (break)를 유도하는 제1 뉴클레아제 작용제 또는 제1 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드; 및

(ii) 표적 좌에 위치한 제1 표적 부위에 상응하는 제1 상동성 아암 및 표적 좌에 위치한 제2 표적 부위에 상응하는 제2 상동성 아암에 의해 플랭킹되는 제1 삽입 폴리뉴클레오티드를 포함하며, 제1 삽입 폴리뉴클레오티드는 (1) 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 핵산; 및 (2) 제2 뉴클레아제 작용제에 대한 제2 뉴클레아제 인식 부위를 포함하는 제2 선택 카세트를 포함하는, 제1 표적화 벡터를 세포 내로 도입하는 단계로서,

여기서, 제2 뉴클레아제 작용제는 ZFN, TALEN, 또는 Cas 단백질 및 gRNA이고,

제1 선택 마커 및 제2 선택 마커는 상이하고,

제1 뉴클레아제 작용제는 제2 뉴클레아제 작용제와 상이하고,

제2 뉴클레아제 인식 부위는 제2 선택 마커의 코딩 영역 또는 제2 선택 마커의 임의의 비-단백질-코딩 영역에 위치하는, 단계;

(c) 표적 좌에서 제1 삽입 폴리뉴클레오티드를 포함하는 변형된 세포를 확인하는 단계로서, 여기서 변형된 세포는 제2 선택 마커의 활성화는 갖지만, 제1 선택 마커의 활성화는 갖지 않는, 단계; 및

(d) 삽입 폴리뉴클레오티드의 하나 이상의 추가적인 통합 라운드를 수행하는 단계로서, 여기서 방법에 사용된 표적화 벡터는 제1 뉴클레아제 작용제에 대한 제1 뉴클레아제 인식 부위를 포함하는 제1 선택 마커 및 제2 뉴클레아제 작용제에 대한 제2 뉴클레아제 인식 부위를 포함하는 제2 선택 마커의 사용을 대체하는, 단계

를 포함하는, 방법.

#### 청구항 2

제1 항에 있어서, 단계 (d)는

(I) (i) 제2 뉴클레아제 인식 부위에 닉 (nick) 또는 이중-가닥 브레이크 (break)를 유도하는 제2 뉴클레아제 작용제 또는 제2 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드; 및

(ii) 표적 좌에 위치한 제3 표적 부위에 상응하는 제3 상동성 아암 및 표적 좌에 위치한 제4 표적 부위에 상응하는 제4 상동성 아암에 의해 플랭킹되는 제2 삽입 폴리뉴클레오티드를 포함하는 제2 표적화 벡터

를 변형된 세포로 도입하는 단계로서,

여기서, 제2 삽입 폴리뉴클레오티드는 (1) 제3 선택 마커 및 (2) 제3 뉴클레아제 작용제에 대한 제3 뉴클레아제 인식 부위를 포함하는 제3 선택 카세트를 포함하고, 여기서 제1 선택 마커와 제3 선택 마커는 동일하고, 제3 뉴클레아제 인식 부위는 제1 뉴클레아제 인식 부위와 동일하나 제2 뉴클레아제 인식 부위와는 상이하고, 제1 뉴클레아제 작용제와 제3 뉴클레아제 작용제는 서로 동일하나 제2 뉴클레아제 작용제와는 상이한, 단계; 및

(II) 표적 좌에 통합된 제2 삽입 폴리뉴클레오티드를 포함하는 적어도 하나의 세포를 확인하는 단계로서, 여기서 상기 적어도 하나의 세포는 제3 선택 마커의 활성을 갖지만 제2 선택 마커의 활성은 갖지 않는, 단계를 포함하는 것을 특징으로 하는 방법.

### 청구항 3

제1 항 또는 제2 항에 있어서, 확인 단계 (c)는

- (i) 제1 선택 마커의 활성을 갖지 않는 세포의 확인을 가능하게 하는 조건 하에서 세포를 배양하는 단계; 또는
- (ii) 제1 및 제2 표적 부위에 통합된 제1 삽입 폴리뉴클레오티드를 포함하는 적어도 하나의 세포를 확인하는 단계

를 포함하는 것을 특징으로 하는 방법.

### 청구항 4

제1 항 또는 제2 항에 있어서, 확인 단계 (c)는 대립유전자의 변형 (MOA) 검정을 통해 수행되는 것을 특징으로 하는 방법.

### 청구항 5

제1 항 또는 제2 항에 있어서, 단계 (a)의 제1 선택 카세트는 제1 표적 부위 및 제2 표적 부위에 의해 플랜킹되는 것을 특징으로 하는 방법.

### 청구항 6

제1 항 또는 제2 항에 있어서, 제1 선택 마커 또는 제2 선택 마커는 항생제에 대한 저항성을 부여하고,

선택적으로 항생제는 G418, 하이그로마이신, 블라스티시딘, 네오마이신, 또는 퓨로마이신을 포함하고,

선택적으로 제1 선택 마커는 네오마이신에 대한 저항성을 부여하고 제2 선택 마커는 하이그로마이신에 대한 저항성을 부여하거나, 또는 제1 선택 마커는 하이그로마이신에 대한 저항성을 부여하고 제2 선택 마커는 네오마이신에 대한 저항성을 부여하는 것을 특징으로 하는 방법.

### 청구항 7

제1 항 또는 제2 항에 있어서, 제1 선택 마커 또는 제2 선택 마커는 유도성 프로모터에 작동가능하게 연결되고, 선택 마커의 발현은 세포에 독성이고,

선택적으로 제1 선택 마커 또는 제2 선택 마커는 하이포잔틴-구아닌 포스포리보실트랜스퍼라제 (HGPRT) 또는 단순 포진 바이러스 (HSV-TK)의 티미딘 키나제를 포함하는 것을 특징으로 하는 방법.

### 청구항 8

제1 항 또는 제2 항에 있어서, 제1 표적화 벡터와 제1 뉴클레아제 작용제의 조합 사용은 제1 표적화 벡터의 단독 사용과 비교하여 증가된 표적화 효율을 가져오고,

선택적으로 제1 표적화 벡터의 표적화 효율은 제1 표적화 벡터의 단독 사용과 비교하여 2배 이상 증가되는 것을 특징으로 하는 방법.

### 청구항 9

제1 항 또는 제2 항에 있어서, 제1 뉴클레아제 인식 부위는 제1 선택 마커의 인트론, 엑손, 프로모터, 프로모터 조절 영역, 또는 인핸서 영역에 위치하고, 제2 뉴클레아제 인식 부위는 제2 선택 마커의 인트론, 엑손, 프로모터, 프로모터 조절 영역, 또는 인핸서 영역에 위치하는 것을 특징으로 하는 방법.

### 청구항 10

제1 항 또는 제2 항에 있어서, 단계 (b)는 제1 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드를 세포 내로 도입하는 단계를 포함하며,

(I) 제1 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드는 제1 뉴클레아제 작용제를 인코딩하는 핵산 서열을 포함하는 발현 작제물을 포함하고, 뉴클레아제 작용제를 인코딩하는 핵산 서열은 세포에서 활성인 프로모터에 작동가능하게 연결되거나; 또는

(II) 제1 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드는 뉴클레아제 작용제를 인코딩하는 mRNA를 포함하는 것을 특징으로 하는 방법.

#### 청구항 11

제1 항 또는 제2 항에 있어서, 제1 뉴클레아제 작용제 또는 제2 뉴클레아제 작용제는 ZFN 또는 Cas 단백질 및 gRNA인 것을 특징으로 하는 방법.

#### 청구항 12

제11 항에 있어서, 제1 뉴클레아제 작용제 또는 제2 뉴클레아제 작용제는 ZFN이고, 제1 뉴클레아제 인식 부위 또는 제2 뉴클레아제 인식 부위는 서열 번호 9-12 중 어느 하나를 포함하는 것을 특징으로 하는 방법.

#### 청구항 13

제11 항에 있어서, 제1 뉴클레아제 작용제 또는 제2 뉴클레아제 작용제는 Cas 단백질 및 gRNA이고, Cas 단백질은 Cas9이고, gRNA는

(I) 프로토스페이서 인접 모티프 (PAM) 서열에 의해 바로 플랭킹된 제1 뉴클레아제 인식 부위 또는 제2 뉴클레아제 인식 부위를 표적화하는 클러스터화된 규칙적으로 사이 공간을 둔 짧은 회문성 반복 (CRISPR) RNA (crRNA); 및

(II) 트랜스-활성화 CRISPR RNA (tracrRNA)를 포함하며,

선택적으로 표적 좌는 서열 번호 1의 뉴클레오티드 서열을 포함하는 것을 특징으로 하는 방법.

#### 청구항 14

제13 항에 있어서, gRNA는 서열 번호 13-20 중 어느 하나를 포함하는 것을 특징으로 하는 방법.

#### 청구항 15

제13 항에 있어서, gRNA는 서열 번호 2, 3, 4, 5, 6, 7, 또는 8을 포함하며,

선택적으로 가이드 RNA는 서열 번호 3 또는 7을 포함하는 것을 특징으로 하는 방법.

#### 청구항 16

제13 항에 있어서, 제1 뉴클레아제 작용제는 Cas 단백질 및 제1 gRNA를 포함하고, 제2 뉴클레아제 작용제는 Cas 단백질 및 제2 gRNA를 포함하고, 제1 gRNA 및 제2 gRNA는 상이한 것을 특징으로 하는 방법.

#### 청구항 17

제12 항에 있어서, 제1 선택 마커는 네오마이신에 대한 저항성을 부여하고, 제1 뉴클레아제 작용제는 ZFN이고, 제1 뉴클레아제 인식 부위는 서열 번호 9 또는 10을 포함하며, 제2 선택 마커는 하이그로마이신에 대한 저항성을 부여하고, 제2 뉴클레아제 작용제는 ZFN이고, 제2 뉴클레아제 인식 부위는 서열 번호 11 또는 12를 포함하는 것을 특징으로 하는 방법.

#### 청구항 18

제13 항에 있어서, 제1 선택 마커는 네오마이신에 대한 저항성을 부여하고, 제1 뉴클레아제 작용제는 Cas9 단백질 및 서열 번호 13-16 중 어느 하나를 포함하는 가이드 RNA이며, 제2 선택 마커 하이그로마이신에 대한 저항성을 부여하고, 제2 뉴클레아제 작용제는 Cas9 단백질 및 서열 번호 17-20 중 어느 하나를 포함하는 가이드 RNA이며,

선택적으로 제1 선택 마커는 네오마이신에 대한 저항성을 부여하고, 제1 뉴클레아제 작용제는 Cas9 단백질 및 서열 번호 13을 포함하는 가이드 RNA이며, 제2 선택 마커는 하이그로마이신에 대한 저항성을 부여하고, 제2 뉴

클레아제 작용제는 Cas9 단백질 및 서열 번호 17을 포함하는 가이드 RNA인 것을 특징으로 하는 방법.

#### 청구항 19

제1 항 또는 제2 항에 있어서,

(I) 제1 표적 부위 및 제2 표적 부위는 제1 뉴클레아제 인식 부위에 바로 인접하거나; 또는

(II) 제1 표적 부위 및 제2 표적 부위는 제1 뉴클레아제 인식 부위로부터 10개의 뉴클레오티드 내지 14 kb에 위치하는 것을 특징으로 하는 방법.

#### 청구항 20

제1 항 또는 제2 항에 있어서,

(I) 제1 상동성 아암 및 상기 제2 상동성 아암의 총 합계는 10 kb 이상이고; 및/또는

(II) 제1 상동성 아암 및 제2 상동성 아암은 각각 5 kb 내지 100 kb 범위이고; 및/또는

(III) 제1 표적화 벡터는 10 kb 이상이거나 20 kb 내지 300 kb이고; 및/또는

(IV) 제1 삽입 폴리뉴클레오티드는 길이가 5 kb 내지 300 kb 범위인 것을 특징으로 하는 방법.

#### 청구항 21

제1 항 또는 제2 항에 있어서, 제1 삽입 폴리뉴클레오티드는 제1 관심 폴리뉴클레오티드를 더 포함하며,

선택적으로 제1 관심 폴리뉴클레오티드는 인간 폴리뉴클레오티드를 포함하는 것을 특징으로 하는 방법.

#### 청구항 22

제21 항에 있어서, 제1 관심 폴리뉴클레오티드는 T 세포 수용체의 영역을 인코딩하는 폴리뉴클레오티드를 포함하고,

선택적으로 T 세포 수용체는 T 세포 수용체 알파이고,

선택적으로 제1 관심 폴리뉴클레오티드는 T 세포 수용체 알파 좌의 적어도 하나의 가변 영역 유전자 분절 및/또는 결합 영역 유전자 분절을 포함하는 것을 특징으로 하는 방법.

#### 청구항 23

제21 항에 있어서, 제1 관심 폴리뉴클레오티드는 인간 면역글로불린 중쇄 가변 영역 아미노산 서열을 인코딩하는 게놈 핵산 서열을 포함하고,

선택적으로 제1 관심 폴리뉴클레오티드는 비-인간 면역글로불린 중쇄 불변 영역 핵산 서열에 작동가능하게 연결된 비재배열된 인간 면역글로불린 중쇄 가변 영역 핵산 서열을 포함하는 것을 특징으로 하는 방법.

#### 청구항 24

제21 항에 있어서, 제1 관심 폴리뉴클레오티드는 인간 면역글로불린 경쇄 가변 영역 아미노산 서열을 인코딩하는 게놈 핵산 서열을 포함하고,

선택적으로 게놈 핵산 서열은 비재배열된 인간  $\lambda$  및/또는  $\kappa$  경쇄 가변 영역 핵산 서열을 포함하거나 또는 게놈 핵산 서열은 재배열된 인간  $\lambda$  및/또는  $\kappa$  경쇄 가변 영역 핵산 서열을 포함하는 것을 특징으로 하는 방법.

#### 청구항 25

제21 항에 있어서, 제1 관심 폴리뉴클레오티드는 적어도 하나의 질환 대립유전자를 포함하는 것을 특징으로 하는 방법.

#### 청구항 26

제21 항에 있어서, 제1 관심 폴리뉴클레오티드는

(I) 세포의 게놈에서 핵산 서열에 대해 상동성 또는 이중상동성인 핵산 서열; 또는  
(II) 외인성 핵산 서열  
을 포함하는 것을 특징으로 하는 방법.

#### 청구항 27

제1 항 또는 제2 항에 있어서, 제1 표적화 벡터는 표적 좌에서 5 kb 내지 3 Mb의 서열이 결실되도록 설계된 것을 특징으로 하는 방법.

#### 청구항 28

제1 항 또는 제2 항에 있어서, 제1 삽입 폴리뉴클레오티드의 표적 좌 내로의 통합은 녹아웃, 녹-인, 점 돌연변이, 도메인 교체, 엑손 교체, 인트론 교체, 조절 서열 교체, 유전자 교체, 또는 이들의 조합을 야기하는 것을 특징으로 하는 방법.

#### 청구항 29

제1 항 또는 제2 항에 있어서, 방법에 의해 생성된 세포는 표적 좌에서 내인성 핵산 서열의 외인성 관심 폴리뉴클레오티드로의 대체를 포함하는 유전자 변형을 포함하는 것을 특징으로 하는 방법.

#### 청구항 30

제1 항 또는 제2 항에 있어서, 표적 좌는 세포의 게놈에 위치하거나 세포 내의 벡터에 위치하는 것을 특징으로 하는 방법.

#### 청구항 31

제30 항에 있어서, 표적 좌는 면역글로불린 좌를 포함하는 것을 특징으로 하는 방법.

#### 청구항 32

제30 항에 있어서, 표적 게놈 좌는 T 세포 수용체 좌를 포함하고, 선택적으로 T 세포 수용체 좌는 T 세포 수용체 알파 좌인 것을 특징으로 하는 방법.

#### 청구항 33

제1 항 또는 제2 항에 있어서, 제1 뉴클레아제 작용제에 의해 유도된 Nick 또는 이중-가닥 브레이크는 제1 선택 마커의 활성을 방해하는 것을 특징으로 하는 방법.

#### 청구항 34

제33 항에 있어서, 표적 좌에서 제1 삽입 폴리뉴클레오티드의 삽입은 제1 선택 마커의 활성을 방해하는 것을 특징으로 하는 방법.

#### 청구항 35

제1 항 또는 제2 항에 있어서, 세포는 진핵 세포이고, 선택적으로 세포는 포유류 세포이고, 선택적으로 세포는

- (I) 비-인간 포유류 세포;
- (II) 다능성 세포;
- (III) 인간 유도 다능성 줄기 세포;
- (IV) 인간 섬유아세포;
- (V) 비-인간 배아 줄기 (ES) 세포;
- (VI) 마우스 배아 줄기 (ES) 세포 또는 래트 배아 줄기 (ES) 세포;
- (VII) 조혈 줄기 세포;

(VIII) 신경 줄기 세포; 또는  
(IX) 설치류 세포  
인 것을 특징으로 하는 방법.

**청구항 36**

제35 항에 있어서, 세포는 마우스 ES 세포 또는 래트 ES 세포인 것을 특징으로 하는 방법.

**청구항 37**

제36 항에 있어서,

(e) 단계 (d)에서 생성된 세포를 전-상실기의 숙주 배아 내로 도입하여 변형된 숙주 배아를 생성하는 단계; 및  
(f) 변형된 숙주 배아를 대리모에 이식하여 단계 (d)에서 생성된 세포로부터 유래된 F0 세대를 생성하는 단계  
를 더 포함하는 것을 특징으로 하는 방법.

**청구항 38**

삭제

**청구항 39**

삭제

**청구항 40**

삭제

**청구항 41**

삭제

**청구항 42**

삭제

**청구항 43**

삭제

**청구항 44**

삭제

**청구항 45**

삭제

**청구항 46**

삭제

**청구항 47**

삭제

**청구항 48**

삭제

**청구항 49**



삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

## 발명의 설명

### 기술 분야

[0001] 관련 출원과의 상호 참조

[0002] 본 출원은 2014년 6월 6일자로 출원된 미국 가출원 제62/008,832호, 및 2014년 6월 27일자로 출원된 미국 가출원 제62/017,916호에 대한 이익을 주장하며, 이들 각각은 전체적으로 본 명세서에 참고로 포함된다.

[0003] 기술분야

[0004] 본 방법 및 조성물은 분자 생물학의 분야에 관한 것이다. 특히, 세포에서 표적화된 좌 (locus)를 변형시키는 방법 및 조성물이 제공된다.

[0005] EFS 웹을 통한 텍스트 파일에 대하여

[0006] 서열 목록의 공식적인 사본은, 2015년 6월 5일자로 제작되고 크기가 5 KB이고 파일명이 461003SEQLIST.TXT인 ASCII 포맷 서열 목록으로서 EFS-웹을 통해 전자적으로 제출되며, 본 명세서와 동시에 제출된다. 이러한 ASCII 포맷 문서에 포함된 서열 목록은 본 명세서의 일부이며 전체적으로 본 명세서에 참고로 포함된다.

### 배경 기술

[0007] 게놈 좌에 특정 핵산 서열을 첨가, 결실, 또는 대체하도록 특이적으로 설계된 표적화 벡터를 사용하는 상동성 (homologous) 재조합은 세포에서 원하는 게놈 변형을 달성하기 위한 대중적인 접근법이다. 표적 좌에 또는 표적 좌 부근에 닉 (nick) 또는 이중-가닥 브레이크 (break)를 도입하도록 특이적으로 조작된 뉴클레아제가 표적 좌에서의 상동성 재조합의 효율을 향상시키기 위해 표적화 벡터와 조합하여 사용될 수 있다.

[0008] 상동성 재조합을 통한 표적화된 변형의 기술이 지난 20년에 걸쳐 상당히 발전해왔음에도 불구하고, 표적화 벡터를 사용하여 허용가능한 표적화 효율을 달성하기에는 여전히 어려움이 남아 있다. 효능 및 효율을 개선하는, 표적화된 변형을 생성하는 방법이 필요하다.

### 발명의 내용

[0009] 세포에서 하나 이상의 표적 좌를 변형시키는 방법 및 조성물이 제공된다.

[0010] 일부 실시 형태에서, 세포에서 표적 좌를 변형시키는 방법이 제공되고, 본 방법은 (a) 세포를 제공하는 단계로서, 세포는 세포에서 활성인 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 제1 폴리뉴클레

오티드를 포함하는 표적 좌를 포함하고, 제1 폴리뉴클레오티드는 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 추가로 포함하는, 상기 세포를 제공하는 단계, (b) (i) 제1 인식 부위에 닉 또는 이중-가닥 브레이크를 유도하는 제1 뉴클레아제 작용제; 및 (ii) 제1 인식 부위에 충분히 가깝게 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암 (arm)에 의해 플랭킹된 제1 삽입 폴리뉴클레오티드를 포함하는 제1 표적화 벡터를 세포 내로 도입하는 단계; 및 (c) 표적 좌에 통합된 제1 삽입 폴리뉴클레오티드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함한다.

[0011] 일부 실시 형태에서, 세포에서 표적 좌를 변형시키는 방법은 (a) 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 제1 폴리뉴클레오티드를 포함하는 제1 표적 좌를 포함하는 세포를 제공하는 단계로서, 제1 폴리뉴클레오티드는 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 추가로 포함하는, 상기 세포를 제공하는 단계, (b) (i) 세포에서 활성인 프로모터에 작동가능하게 연결된 제1 뉴클레아제 작용제를 인코딩하는 하나 이상의 발현 작제물로서, 제1 뉴클레아제 작용제는 제1 폴리뉴클레오티드에서의 제1 인식 부위에 닉 또는 이중-가닥 브레이크를 유도하여 제1 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 발현 작제물; 및 (ii) 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 제2 폴리뉴클레오티드를 포함하는 제1 삽입 폴리뉴클레오티드를 포함하는 제1 표적화 벡터로서, 제1 삽입 핵산은 제1 표적 좌에 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹되는, 상기 제1 표적화 벡터를 세포 내로 도입하는 단계; 및 (c) 제1 표적 좌에 제1 삽입 핵산을 포함하는 변형된 세포를 확인하는 단계로서, 변형된 세포는 제2 선택 마커의 활성은 갖지만 제1 선택 마커의 활성은 갖지 않는, 상기 변형된 세포를 확인하는 단계를 포함하며, 제1 및 제2 선택 마커는 상이하다.

[0012] 일 실시 형태에서, 표적 좌는 세포의 게놈 내에 있다. 다른 실시 형태에서, 표적 좌는 세포 내의 벡터에 위치한다. 일 실시 형태에서, 제1 인식 부위에서의 닉 또는 이중 가닥 브레이크는 제1 선택 마커의 활성을 방해한다. 또 다른 실시 형태에서, 확인 단계 (c)는 제1 선택 마커의 활성을 갖지 않는 세포의 확인을 가능하게 하는 조건 하에서 세포를 배양하는 단계를 포함한다. 일 실시 형태에서, 제1 선택 마커를 포함하는 제1 폴리뉴클레오티드는 제1 표적 부위 및 제2 표적 부위에 의해 플랭킹된다. 일 실시 형태에서, 확인 단계 (c)는 제1 및 제2 표적 부위에 통합된 제1 삽입 폴리뉴클레오티드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함한다. 일 실시 형태에서, 제1 삽입 폴리뉴클레오티드는 (a) 제1 관심 폴리뉴클레오티드; 및 (b) 세포에서 활성인 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 제2 폴리뉴클레오티드로서, 제2 폴리뉴클레오티드는 제2 뉴클레아제 작용제에 대한 제2 인식 부위를 포함하는, 상기 제2 폴리뉴클레오티드를 포함한다.

[0013] 일 실시 형태에서, 본 방법은 (a) 표적 좌에 통합된 제1 삽입 폴리뉴클레오티드를 포함하는 세포 내로 (i) 제2 인식 부위에 닉 또는 이중-가닥 브레이크를 유도하는 제2 뉴클레아제 작용제; 및 (ii) 제2 인식 부위에 충분히 가깝게 위치한 제3 및 제4 표적 부위에 상응하는 제3 및 제4 상동성 아암에 의해 플랭킹된 제2 삽입 폴리뉴클레오티드를 포함하는 제2 표적화 벡터를 도입하는 단계; 및 (b) 표적 좌에 통합된 제2 삽입 폴리뉴클레오티드를 포함하는 적어도 하나의 세포를 확인하는 단계를 추가로 포함한다. 일 실시 형태에서, 제2 인식 부위에서의 닉 또는 이중 가닥 브레이크는 제2 선택 마커의 활성을 방해한다. 일 실시 형태에서, 확인 단계 (b)는 제2 선택 마커의 활성을 갖지 않는 세포의 확인을 가능하게 하는 조건 하에서 세포를 배양하는 단계를 포함한다. 일 실시 형태에서, 제2 선택가능 마커를 포함하는 제2 폴리뉴클레오티드는 제3 표적 부위 및 제4 표적 부위에 의해 플랭킹된다. 일 실시 형태에서, 확인 단계 (b)는 제3 및 제4 표적 부위에 통합된 제2 삽입 폴리뉴클레오티드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함한다.

[0014] 일 실시 형태에서, 제2 삽입 폴리뉴클레오티드는 (a) 제2 관심 폴리뉴클레오티드; 및 (b) 세포에서 활성인 제3 프로모터에 작동가능하게 연결된 제3 선택 마커를 인코딩하는 제3 폴리뉴클레오티드로서, 제3 폴리뉴클레오티드는 제3 뉴클레아제 작용제에 대한 제3 인식 부위를 포함하는, 상기 제3 폴리뉴클레오티드를 포함한다. 일 실시 형태에서, 제1 뉴클레아제 작용제는 제2 뉴클레아제 작용제와 상이하다. 일 실시 형태에서, 제1 선택 마커는 제2 선택 마커와 상이하다. 일 실시 형태에서, 제1 및 제3 뉴클레아제 인식 부위는 서로 동일하고 제2 뉴클레아제 인식 부위와 상이하고; 제1 및 제3 뉴클레아제 작용제는 서로 동일하고 제2 뉴클레아제 작용제와 상이하다. 일 실시 형태에서, 제1 및 제3 선택 마커는 동일하다. 일 실시 형태에서, 제1, 제2 또는 제3 선택 마커 중 하나는 항생제에 대한 저항성을 부여한다. 일 실시 형태에서, 항생제는 G418, 하이그로마이신, 블라스토시딘, 네오마이신, 또는 퓨로마이신을 포함한다. 일 실시 형태에서, 제1, 제2 또는 제3 선택 마커 중 하나는 유도성 프로모터에 작동가능하게 연결되고, 선택가능 마커의 발현은 세포에 독성이다. 일 실시 형태에서, 제1, 제2 또는 제3 선택 마커는 하이포잔틴-구아닌 포스포리보실트랜스퍼라제 (hypoxanthine-guanine phosphoribosyltransferase) (HGPRT) 또는 단순 포진 바이러스 (Herpes simplex virus) (HSV-TK)의 티미딘 키

나제를 포함한다. 일 실시 형태에서, 상기 세포는 원핵 세포이다. 일 실시 형태에서, 세포는 진핵 세포이다. 일 실시 형태에서, 진핵 세포는 포유류 세포이다. 일 실시 형태에서, 포유류 세포는 비-인간 포유류 세포이다. 일 실시 형태에서, 포유류 세포는 설치류 유래이다. 일 실시 형태에서, 설치류는 래트 또는 마우스이다.

[0015] 일 실시 형태에서, 세포는 다능성 (pluripotent) 세포이다. 일 실시 형태에서, 포유류 세포는 인간 유도 다능성 줄기 (induced pluripotent stem) (iPS) 세포이다. 일 실시 형태에서, 다능성 세포는 비-인간 배아 줄기 (embryonic stem) (ES) 세포이다. 일 실시 형태에서, 다능성 세포는 마우스 배아 줄기 (ES) 세포 또는 래트 배아 줄기 (ES) 세포이다. 일 실시 형태에서, 다능성 세포는 조혈 줄기 세포이다. 일 실시 형태에서, 다능성 세포는 신경 줄기 세포이다. 일 실시 형태에서, 포유류 세포는 인간 섬유아세포이다.

[0016] 일 실시 형태에서, 제1 표적화 벡터와 제1 뉴클레아제 작용제의 조합 사용은 제1 표적화 벡터의 단독 사용과 비교하여 증가된 표적화 효율을 가져온다. 일 실시 형태에서, 제1 표적화 벡터의 표적화 효율은 제1 표적화 벡터의 단독 사용과 비교하여 2배 이상 증가된다.

[0017] 일 실시 형태에서, 제1 또는 제2 뉴클레아제 작용제는 뉴클레아제 작용제를 인코딩하는 핵산 서열을 포함하는 발현 작제물을 포함하고, 핵산은 세포에서 활성인 제4 프로모터에 작동가능하게 연결된다. 일 실시 형태에서, 제1 또는 제2 뉴클레아제 작용제는 뉴클레아제를 인코딩하는 mRNA이다. 일 실시 형태에서, 제1 또는 제2 뉴클레아제 작용제는 징크 핑거 뉴클레아제 (zinc finger nuclease) (ZFN)이다. 일 실시 형태에서, 제1 또는 제2 뉴클레아제 작용제는 전사 활성화인자-유사 이펙터 뉴클레아제 (Transcription Activator-Like Effector Nuclease) (TALEN)이다. 일 실시 형태에서, 제1 또는 제2 뉴클레아제 작용제는 메가뉴클레아제이다.

[0018] 일 실시 형태에서, 제1 또는 제2 뉴클레아제 작용제는 클러스터화된 규칙적으로 사이 공간을 둔 짧은 회문성 반복 (CRISPR)-관련 (Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated) (Cas) 단백질 및 가이드 RNA (guide RNA) (gRNA)를 포함한다. 일 실시 형태에서, 가이드 RNA (gRNA)는 (a) 제1, 제2, 또는 제3 인식 부위를 표적으로 하는 클러스터화된 규칙적으로 사이 공간을 둔 짧은 회문성 반복 (CRISPR) RNA (crRNA); 및 (b) 트랜스-활성화 (trans-activating) CRISPR RNA (tracrRNA)를 포함한다. 일 실시 형태에서, 제1 또는 제2 인식 부위는 프로토스페이스 인접 모티프 (Protospacer Adjacent Motif) (PAM) 서열에 의해 바로 플랭킹된다. 일 실시 형태에서, 관심 게놈 좌는 서열 번호 1의 뉴클레오티드 서열을 포함한다. 일 실시 형태에서, Cas 단백질은 Cas9이다. 일 실시 형태에서, gRNA는 (a) 서열 번호 2의 핵산 서열의 키메라 (chimeric) RNA; 또는 (b) 서열 번호 3의 핵산 서열의 키메라 RNA를 포함한다. 일 실시 형태에서, crRNA는 서열 번호 4; 서열 번호 5; 또는 서열 번호 6을 포함한다. 일 실시 형태에서, tracrRNA는 서열 번호 7 또는 서열 번호 8을 포함한다.

[0019] 일 실시 형태에서, 제1, 제2, 및/또는 제3 인식 부위는 제1, 제2, 또는 제3 선택 마커의 인트론, 엑손, 프로모터, 프로모터 조절 영역, 또는 인핸서 영역에 위치한다. 일 실시 형태에서, 제1 표적 부위 및 제2 표적 부위는 제1 인식 부위에 바로 인접한다. 일 실시 형태에서, 제1 표적 부위 및 제2 표적 부위는 제1 인식 부위로부터 약 10개 뉴클레오티드 내지 약 14 kb에 위치한다. 일 실시 형태에서, 제3 표적 부위 및 제4 표적 부위는 제2 인식 부위에 바로 인접한다. 일 실시 형태에서, 제3 표적 부위 및 제4 표적 부위는 제2 인식 부위로부터 약 10개 뉴클레오티드 내지 약 14 kb에 위치한다.

[0020] 일 실시 형태에서, 제1 상동성 아암 및 제2 상동성 아암의 총 합계는 약 10 kb 이상이다. 일 실시 형태에서, 제3 상동성 아암 및 제4 상동성 아암의 총 합계는 약 10 kb 이상이다. 일 실시 형태에서, 제1 삽입 폴리뉴클레오티드의 길이는 약 5 kb 내지 약 300 kb 범위이다. 일 실시 형태에서, 제2 삽입 폴리뉴클레오티드의 길이는 약 5 kb 내지 약 300 kb 범위이다.

[0021] 일 실시 형태에서, 제1 삽입 폴리뉴클레오티드의 표적 좌 내로의 통합은 녹아웃 (knockout), 녹-인 (knock-in), 점 돌연변이, 도메인 교체 (swap), 엑손 교체, 인트론 교체, 조절 서열 교체, 유전자 교체, 또는 이들의 조합을 야기한다. 일 실시 형태에서, 제2 삽입 폴리뉴클레오티드의 표적 좌 내로의 통합은 녹아웃, 녹-인, 점 돌연변이, 도메인 교체, 엑손 교체, 인트론 교체, 조절 서열 교체, 유전자 교체, 또는 이들의 조합을 야기한다.

[0022] 일 실시 형태에서, 제1 삽입 폴리뉴클레오티드는 인간 폴리뉴클레오티드를 포함하는 관심 폴리뉴클레오티드를 포함한다. 일 실시 형태에서, 제2 삽입 폴리뉴클레오티드는 인간 폴리뉴클레오티드를 포함하는 관심 폴리뉴클레오티드를 포함한다. 일 실시 형태에서, 제1 삽입 폴리뉴클레오티드는 T 세포 수용체 알파 좌의 영역을 포함하는 관심 폴리뉴클레오티드를 포함한다.

[0023] 일 실시 형태에서, 제2 삽입 폴리뉴클레오티드는 T 세포 수용체 알파 좌의 영역을 포함하는 관심 폴리뉴클레오

티드를 포함한다. 일 실시 형태에서, 제1 또는 제2 삽입 폴리뉴클레오티드는 T 세포 수용체 알파 좌의 적어도 하나의 가변 영역 유전자 분절 및/또는 결합 (joining) 영역 유전자 분절을 포함하는 관심 폴리뉴클레오티드를 포함한다. 일 실시 형태에서, T 세포 수용체 알파 좌의 영역은 인간 유래의 것이다.

[0024] 일 실시 형태에서, 제1 삽입 폴리뉴클레오티드는 비-인간 면역글로불린 중쇄 불변 영역 핵산 서열에 작동가능하게 연결된 비재배열된 인간 면역글로불린 중쇄 가변 영역 핵산 서열을 포함하는 관심 폴리뉴클레오티드를 포함한다.

[0025] 일 실시 형태에서, 확인 단계는 대립유전자의 변형 (modification of allele) (MOA) 검정을 통해 수행된다. 일 실시 형태에서, 제1 삽입 폴리뉴클레오티드는 세포의 게놈 내의 핵산 서열에 상동성 또는 이종상동성 (orthologous)인 핵산 서열을 포함하는 관심 폴리뉴클레오티드를 포함한다. 일 실시 형태에서, 제2 삽입 폴리뉴클레오티드는 세포의 게놈 내의 핵산 서열에 상동성 또는 이종상동성인 핵산 서열을 포함한다. 일 실시 형태에서, 제1 삽입 폴리뉴클레오티드는 외인성 핵산 서열을 포함하는 관심 폴리뉴클레오티드를 포함한다. 일 실시 형태에서, 제2 삽입 폴리뉴클레오티드는 외인성 핵산 서열을 포함하는 관심 폴리뉴클레오티드를 포함한다.

[0026] 일부 실시 형태에서, 세포에서 표적 좌를 변형시키는 방법은 (a) 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 핵산을 포함하는 제1 표적 좌를 포함하는 세포를 제공하는 단계; (b) (i) Cas 단백질 및 제1 가이드 RNA (gRNA)를 인코딩하는 하나 이상의 발현 작제물로서, 이들 각각은 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제1 핵산에서의 제1 gRNA 표적 부위에 Nick 또는 이중-가닥 브레이크를 유도하여 제1 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 발현 작제물, 및 (ii) 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 제2 핵산을 포함하는 제1 삽입 핵산을 포함하는 제1 표적화 벡터로서, 제1 삽입 핵산은 제1 표적 좌에 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹되는, 상기 제1 표적화 벡터를 세포 내로 도입하는 단계; 및 (c) 제1 표적 좌에 제1 삽입 핵산을 포함하는 변형된 세포를 확인하는 단계로서, 변형된 세포는 제2 선택 마커의 활성은 갖지만 제1 선택 마커의 활성은 갖지 않는, 상기 변형된 세포를 확인하는 단계를 포함하며, 제1 및 제2 선택 마커는 상이하다. 일 실시 형태에서, 제1 gRNA는 제1 삽입 핵산에 혼성화되지 않는다. 일 실시 형태에서, 관심 표적 좌는 세포의 게놈에 위치한다. 다른 실시 형태에서, 관심 표적 좌는 세포 내의 벡터에 위치한다. 일 실시 형태에서, 확인 단계 (c)는 제2 선택 마커의 활성은 갖지만 제1 선택 마커의 활성은 갖지 않는 변형된 세포의 확인을 가능하게 하는 조건 하에서 세포를 배양하는 단계를 포함한다.

[0027] 일 실시 형태에서, 본 방법은 (d) 제1 표적 좌에 제1 삽입 핵산을 포함하는 변형된 세포 내로 (i) Cas 단백질 및 제2 gRNA를 인코딩하는 하나 이상의 핵산으로서, 이들 각각은 변형된 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제2 핵산을 포함하는 제1 삽입 핵산에서 제2 gRNA 표적 부위에 Nick 또는 이중-가닥 브레이크를 유도하여 제2 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 핵산, 및 (ii) 제3 프로모터에 작동가능하게 연결된 제3 선택 마커를 인코딩하는 제3 핵산을 포함하는 제2 삽입 핵산을 포함하는 제2 표적화 벡터로서, 제2 삽입 핵산은 제2 표적 좌에 위치한 제3 및 제4 표적 부위에 상응하는 제3 및 제4 상동성 아암에 의해 플랭킹되는, 상기 제2 표적화 벡터를 도입하는 단계; 및 (e) 제2 표적 좌에 제2 삽입 핵산을 포함하는 제2 변형된 세포를 확인하는 단계로서, 제2 변형된 세포는 제3 선택 마커의 활성은 갖지만 제2 선택 마커의 활성은 갖지 않는, 상기 제2 변형된 세포를 확인하는 단계를 추가로 포함하며, 제2 및 제3 선택 마커는 상이하다. 일 실시 형태에서, 제1 및 제2 표적 좌는 서로 바로 인접하게 위치한다. 다른 실시 형태에서, 제1 또는 제2 표적 좌는 제1 또는 제2 gRNA 표적 부위로부터 약 10개 뉴클레오티드 내지 약 14 kb, 약 10개 뉴클레오티드 내지 약 100개 뉴클레오티드, 약 100개 뉴클레오티드 내지 약 500개 뉴클레오티드, 약 500개 뉴클레오티드 내지 약 1000개 뉴클레오티드, 약 1 kb 내지 약 5 kb, 약 5 kb 내지 약 10 kb, 또는 약 10 kb 내지 약 14 kb에 위치한다. 일 실시 형태에서, 제2 gRNA는 제2 삽입 핵산에 혼성화되지 않는다. 일 실시 형태에서, 확인 단계 (e)는 제3 선택 마커의 활성은 갖지만 제2 선택 마커의 활성은 갖지 않는 제2 변형된 세포의 확인을 가능하게 하는 조건 하에서 변형된 세포를 배양하는 단계를 포함한다.

[0028] 일 실시 형태에서, 본 방법은 (f) 제2 표적 좌에 제2 삽입 핵산을 포함하는 제2 변형된 세포 내로 (i) Cas 단백질 및 제3 gRNA를 인코딩하는 하나 이상의 발현 작제물로서, 이들 각각은 제2 변형된 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제3 핵산을 포함하는 제2 삽입 핵산에서 제3 gRNA 표적 부위에 Nick 또는 이중-가닥 브레이크를 유도하여 제3 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 발현 작제물, 및 (ii) 제4 프로모터에 작동가능하게 연결된 제4 선택 마커를 인코딩하는 제4 핵산을 포함하는 제3 삽입 핵산을 포함하는 제3 표적화 벡터로서, 제3 삽입 핵산은 제3 표적 좌에 위치한 제5 및 제6 표적 부위에 상응하는 제5 및 제6 상동성 아암에 의해 플랭킹되는, 상기 제3 표적화 벡터를 도입하는 단계; 및 (g) 제3 표적 좌



에 제3 삽입 핵산을 포함하는 제3 변형된 세포를 확인하는 단계로서, 제3 변형된 세포는 제4 선택 마커의 활성은 갖지만 제3 선택 마커의 활성은 갖지 않는, 상기 제3 변형된 세포를 확인하는 단계를 추가로 포함하며, 제3 및 제4 선택 마커는 상이하다. 일 실시 형태에서, 제2 및 제3 표적 좌는 서로 바로 인접하게 위치한다. 다른 실시 형태에서, 제2 또는 제3 표적 좌는 제1 또는 제2 gRNA 표적 부위로부터 약 10개 뉴클레오티드 내지 약 14 kb에 위치한다.

[0029] 일 실시 형태에서, 제1, 제2, 제3, 또는 제4 마커는 항생제에 대한 저항성을 부여한다. 일 실시 형태에서, 항생제는 G418, 하이그로마이신, 블라스토시딘, 네오마이신, 또는 푸로마이신을 포함한다. 일 실시 형태에서, 제1, 제2, 제3, 또는 제4 선택 마커는 하이포잔틴-구아닌 포스포리보실트랜스퍼라제 (HGPRT) 또는 단순 포진 바이러스 (HSV-TK)의 티미딘 키나제를 포함한다. 일 실시 형태에서, 제1, 제2, 또는 제3 gRNA는 (i) 제1, 제2 또는 제3 gRNA 표적 부위에 혼성화되는 뉴클레오티드 서열 및 (ii) 트랜스-활성화 CRISPR RNA (tracrRNA)를 포함한다. 일 실시 형태에서, 제1, 제2, 또는 제3 표적 좌는 제1, 제2 또는 제3 gRNA 표적 부위에 아주 가깝게 위치해서 gRNA 표적 부위에서의 Nick 또는 이중-가닥 브레이크가 표적 좌에서 표적화 벡터의 상동성 재조합을 촉진하도록 한다. 일 실시 형태에서, Cas 단백질은 Cas9이다. 일 실시 형태에서, 제1, 제2, 또는 제3 gRNA 표적 부위는 프로토스페이스어 인접 모티프 (PAM) 서열에 의해 바로 플랭킹된다.

[0030] 일 실시 형태에서, 세포는 원핵 세포이다. 다른 실시 형태에서, 세포는 진핵 세포이다. 일 실시 형태에서, 진핵 세포는 포유류 세포이다. 일 실시 형태에서, 포유류 세포는 섬유아세포이다. 일 실시 형태에서, 포유류 세포는 인간 섬유아세포이다. 일 실시 형태에서, 포유류 세포는 비-인간 포유류 세포이다. 일 실시 형태에서, 포유류 세포는 설치류 유래이다. 일 실시 형태에서, 설치류는 래트, 마우스, 또는 햄스터이다.

[0031] 일 실시 형태에서, 진핵 세포는 다능성 세포이다. 일 실시 형태에서, 다능성 세포는 조혈 줄기 세포 또는 신경 줄기 세포이다. 일 실시 형태에서, 다능성 세포는 인간 유도 다능성 줄기 (iPS) 세포이다. 일 실시 형태에서, 다능성 세포는 마우스 배아 줄기 (ES) 세포 또는 래트 배아 줄기 (ES) 세포이다.

[0032] 일 실시 형태에서, 제1, 제2, 또는 제3 gRNA 표적 부위는 제1, 제2, 또는 제3 선택 마커를 인코딩하는 제1, 제2, 또는 제3 핵산에서의 인트론, 엑손, 프로모터, 또는 프로모터 조절 영역에 위치한다. 일 실시 형태에서, 제1, 제2, 또는 제3 표적화 벡터는 약 10 kb 이상이다. 일 실시 형태에서, 제1, 제2, 또는 제3 삽입 핵산은 약 5 kb 내지 약 300 kb의 범위이다.

[0033] 일 실시 형태에서, 제1, 제2, 또는 제3 삽입 핵산은 인간 T 세포 수용체 알파 좌의 게놈 영역을 포함한다. 일 실시 형태에서, 게놈 영역은 인간 T 세포 수용체 알파 좌의 적어도 하나의 가변 영역 유전자 분절 및/또는 결합 영역 유전자 분절을 포함한다.

[0034] 일 실시 형태에서, 제1 및 제3 선택 마커는 동일하다. 일 실시 형태에서, 제1 및 제3 선택 마커는 동일하고 제2 및 제4 선택 마커는 동일하다. 일 실시 형태에서, 제1 및 제3 gRNA는 동일하다.

[0035] 세포에서 표적 좌를 변형시키는 방법 및 조성물이 추가로 제공된다. 그러한 방법은 세포를 제공하는 단계를 포함하고, 세포는 세포에서 활성인 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 제1 폴리뉴클레오티드를 포함하는 표적 좌를 포함하고, 제1 폴리뉴클레오티드는 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 추가로 포함한다. 제1 뉴클레아제 작용제가 세포 내로 도입되고, 제1 뉴클레아제 작용제는 제1 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도한다. 제1 인식 부위에 충분히 가깝게 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹된 제1 삽입 폴리뉴클레오티드를 포함하는 제1 표적화 벡터가 추가로 세포 내로 도입된다. 이어서, 표적 좌에 통합된 제1 삽입 폴리뉴클레오티드를 포함하는 적어도 하나의 세포가 확인된다.

[0036] 세포에서 표적 좌를 변형시키는 방법이 또한 제공되고, 본 방법은 (a) 세포를 제공하는 단계로서, 세포는 세포에서 활성인 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 제1 폴리뉴클레오티드를 포함하는 표적 좌를 포함하고, 제1 폴리뉴클레오티드는 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 추가로 포함하는, 상기 세포를 제공하는 단계, (b) (i) 제1 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도하는 제1 뉴클레아제 작용제; 및 (ii) 제1 인식 부위에 충분히 가깝게 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹된 제1 삽입 폴리뉴클레오티드를 포함하는 제1 표적화 벡터를 세포 내로 도입하는 단계; 및 (c) 표적 좌에 통합된 제1 삽입 폴리뉴클레오티드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함한다. 일 실시 형태에서, 표적 좌는 세포의 게놈 내에 있다. 다른 실시 형태에서, 표적 좌는 세포 내의 벡터에 위치한다. 일 실시 형태에서, 제1 인식 부위에서의 Nick 또는 이중-가닥 브레이크는 제1 선택 마커의 활성

성을 방해한다. 또 다른 실시 형태에서, 확인 단계 (c)는 제1 선택 마커의 활성을 갖지 않는 세포의 확인을 가능하게 하는 조건 하에서 세포를 배양하는 단계를 포함한다. 일 실시 형태에서, 제1 선택 마커를 포함하는 제1 폴리뉴클레오티드는 제1 표적 부위 및 제2 표적 부위에 의해 플랭킹된다. 일 실시 형태에서, 확인 단계 (c)는 제1 및 제2 표적 부위에 통합된 제1 삽입 폴리뉴클레오티드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함한다. 일 실시 형태에서, 제1 삽입 폴리뉴클레오티드는 (a) 제1 관심 폴리뉴클레오티드; 및 (b) 세포에서 활성인 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 제2 폴리뉴클레오티드로서, 제2 폴리뉴클레오티드는 제2 뉴클레아제 작용제에 대한 제2 인식 부위를 포함하는, 상기 제2 폴리뉴클레오티드를 포함한다.

[0037] 일 실시 형태에서, 본 방법은 (a) 표적 좌에 통합된 제1 삽입 폴리뉴클레오티드를 포함하는 세포 내로 (i) 제2 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도하는 제2 뉴클레아제 작용제; 및 (ii) 제2 인식 부위에 충분히 가깝게 위치한 제3 및 제4 표적 부위에 상응하는 제3 및 제4 상동성 아암에 의해 플랭킹된 제2 삽입 폴리뉴클레오티드를 포함하는 제2 표적화 벡터를 도입하는 단계; 및 (b) 표적 좌에 통합된 제2 삽입 폴리뉴클레오티드를 포함하는 적어도 하나의 세포를 확인하는 단계를 추가로 포함한다. 일 실시 형태에서, 제2 인식 부위에서의 Nick 또는 이중 가닥 브레이크는 제2 선택 마커의 활성을 방해한다. 일 실시 형태에서, 확인 단계 (b)는 제2 선택 마커의 활성을 갖지 않는 세포의 확인을 가능하게 하는 조건 하에서 세포를 배양하는 단계를 포함한다. 일 실시 형태에서, 제2 선택 마커를 포함하는 제2 폴리뉴클레오티드는 제3 표적 부위 및 제4 표적 부위에 의해 플랭킹된다. 일 실시 형태에서, 확인 단계 (b)는 제3 및 제4 표적 부위에 통합된 제2 삽입 폴리뉴클레오티드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함한다. 일 실시 형태에서, 제2 삽입 폴리뉴클레오티드는 (a) 제2 관심 폴리뉴클레오티드; 및 (b) 세포에서 활성인 제3 프로모터에 작동가능하게 연결된 제3 선택 마커를 인코딩하는 제3 폴리뉴클레오티드로서, 제3 폴리뉴클레오티드는 제3 뉴클레아제 작용제에 대한 제3 인식 부위를 포함하는, 상기 제3 폴리뉴클레오티드를 포함한다. 일 실시 형태에서, 제1 뉴클레아제 작용제는 제2 뉴클레아제 작용제와 상이하다. 일 실시 형태에서, 제1 선택 마커는 제2 선택 마커와 상이하다. 일 실시 형태에서, 제1 및 제3 뉴클레아제 인식 부위는 서로 동일하고 제2 뉴클레아제 인식 부위와 상이하고; 제1 및 제3 뉴클레아제 작용제는 서로 동일하고 제2 뉴클레아제 작용제와 상이하다. 일 실시 형태에서, 제1 및 제3 선택 마커는 동일하다. 일 실시 형태에서, 제1, 제2 또는 제3 선택 마커 중 하나는 항생제에 대한 저항성을 부여한다. 일 실시 형태에서, 항생제는 G418, 하이그로마이신, 블라스토시딘, 네오마이신, 또는 퓨로마이신을 포함한다. 일 실시 형태에서, 제1, 제2 또는 제3 선택 마커 중 하나는 유도성 프로모터에 작동가능하게 연결되고, 선택 마커의 발현은 세포에 독성이다. 일 실시 형태에서, 제1, 제2 또는 제3 선택 마커는 하이포잔틴-구아닌 포스포리보실트랜스퍼라제 (HGPRT) 또는 단순 포진 바이러스 (HSV-TK)의 티미딘 키나제를 포함한다. 일 실시 형태에서, 상기 세포는 원핵 세포이다. 일 실시 형태에서, 세포는 진핵 세포이다. 일 실시 형태에서, 진핵 세포는 포유류 세포이다. 일 실시 형태에서, 포유류 세포는 비-인간 포유류 세포이다. 일 실시 형태에서, 포유류 세포는 설치류 유래이다. 일 실시 형태에서, 설치류는 래트 또는 마우스이다. 일 실시 형태에서, 포유류 세포는 인간 섬유아세포이다.

[0038] 일 실시 형태에서, 세포는 다능성 세포이다. 일 실시 형태에서, 포유류 세포는 인간 유도 다능성 줄기 (iPS) 세포이다. 일 실시 형태에서, 다능성 세포는 비-인간 배아 줄기 (ES) 세포이다. 일 실시 형태에서, 다능성 세포는 마우스 배아 줄기 (ES) 세포 또는 래트 배아 줄기 (ES) 세포이다. 일 실시 형태에서, 다능성 세포는 조혈 줄기 세포이다. 일 실시 형태에서, 다능성 세포는 신경 줄기 세포이다.

[0039] 일 실시 형태에서, 제1 표적화 벡터와 제1 뉴클레아제 작용제의 조합 사용은 제1 표적화 벡터의 단독 사용과 비교하여 증가된 표적화 효율을 가져온다. 일 실시 형태에서, 제1 표적화 벡터의 표적화 효율은 제1 표적화 벡터의 단독 사용과 비교하여 2배 이상 증가된다.

[0040] 일 실시 형태에서, 제1 또는 제2 뉴클레아제 작용제는 뉴클레아제 작용제를 인코딩하는 핵산 서열을 포함하는 발현 작제물을 포함하고, 핵산은 세포에서 활성인 제4 프로모터에 작동가능하게 연결된다. 일 실시 형태에서, 제1 또는 제2 뉴클레아제 작용제는 뉴클레아제를 인코딩하는 mRNA이다. 일 실시 형태에서, 제1 또는 제2 뉴클레아제 작용제는 징크 핑거 뉴클레아제 (ZFN)이다. 일 실시 형태에서, 제1 또는 제2 뉴클레아제 작용제는 전사 활성화인자-유사 이펙터 뉴클레아제 (TALEN)이다. 일 실시 형태에서, 제1 또는 제2 뉴클레아제 작용제는 메가 뉴클레아제이다.

[0041] 일 실시 형태에서, 제1 또는 제2 뉴클레아제 작용제는 클러스터화된 규칙적으로 사이 공간을 둔 짧은 회문성 반복 (CRISPR)-관련 (Cas) 단백질 및 가이드 RNA (gRNA)를 포함한다. 일 실시 형태에서, 가이드 RNA (gRNA)는 (a) 제1, 제2, 또는 제3 인식 부위를 표적으로 하는 클러스터화된 규칙적으로 사이 공간을 둔 짧은 회문성 반복

(CRISPR) RNA (crRNA); 및 (b) 트랜스-활성화 CRISPR RNA (tracrRNA)를 포함한다. 일 실시 형태에서, 제1 또는 제2 인식 부위는 프로토스페이서 인접 모티프 (PAM) 서열에 의해 바로 플랭킹된다. 일 실시 형태에서, 관심 게놈 좌는 서열 번호 1의 뉴클레오타이드 서열을 포함한다. 일 실시 형태에서, Cas 단백질은 Cas9이다. 일 실시 형태에서, gRNA는 (a) 서열 번호 2의 핵산 서열의 키메라 RNA 또는 (b) 서열 번호 3의 핵산 서열의 키메라 RNA를 포함한다. 일 실시 형태에서, crRNA는 서열 번호 4; 서열 번호 5; 또는 서열 번호 6을 포함한다. 일 실시 형태에서, tracrRNA는 서열 번호 7 또는 서열 번호 8을 포함한다. 일 실시 형태에서, 제1, 제2, 및/또는 제3 인식 부위는 제1, 제2, 또는 제3 선택 마커의 인트론, 엑손, 프로모터, 프로모터 조절 영역, 또는 인핸서 영역에 위치한다. 일 실시 형태에서, 제1 표적 부위 및 제2 표적 부위는 제1 인식 부위에 바로 인접한다. 일 실시 형태에서, 제1 표적 부위 및 제2 표적 부위는 제1 인식 부위로부터 약 10개 뉴클레오타이드 내지 약 14 kb에 위치한다. 일 실시 형태에서, 제3 표적 부위 및 제4 표적 부위는 제2 인식 부위에 바로 인접한다. 일 실시 형태에서, 제3 표적 부위 및 제4 표적 부위는 제2 인식 부위로부터 약 10개 뉴클레오타이드 내지 약 14 kb에 위치한다. 일 실시 형태에서, 제1 상동성 아암 및 제2 상동성 아암의 총 합계는 약 10 kb 이상이다. 일 실시 형태에서, 제3 상동성 아암 및 제4 상동성 아암의 총 합계는 약 10 kb 이상이다. 일 실시 형태에서, 제1 삽입 폴리뉴클레오타이드의 길이는 약 5 kb 내지 약 300 kb 범위이다. 일 실시 형태에서, 제2 삽입 폴리뉴클레오타이드의 길이는 약 5 kb 내지 약 300 kb 범위이다. 일 실시 형태에서, 제1 삽입 폴리뉴클레오타이드의 표적 좌 내로의 통합은 녹아웃, 녹-인, 점 돌연변이, 도메인 교체, 엑손 교체, 인트론 교체, 조절 서열 교체, 유전자 교체, 또는 이들의 조합을 야기한다. 일 실시 형태에서, 제2 삽입 폴리뉴클레오타이드의 표적 좌 내로의 통합은 녹아웃, 녹-인, 점 돌연변이, 도메인 교체, 엑손 교체, 인트론 교체, 조절 서열 교체, 유전자 교체, 또는 이들의 조합을 야기한다. 일 실시 형태에서, 제1 삽입 폴리뉴클레오타이드는 인간 폴리뉴클레오타이드를 포함하는 관심 폴리뉴클레오타이드를 포함한다. 일 실시 형태에서, 제2 삽입 폴리뉴클레오타이드는 인간 폴리뉴클레오타이드를 포함하는 관심 폴리뉴클레오타이드를 포함한다. 일 실시 형태에서, 제1 삽입 폴리뉴클레오타이드는 T 세포 수용체 알파 좌의 영역을 포함하는 관심 폴리뉴클레오타이드를 포함한다. 일 실시 형태에서, 제2 삽입 폴리뉴클레오타이드는 T 세포 수용체 알파 좌의 영역을 포함하는 관심 폴리뉴클레오타이드를 포함한다. 일 실시 형태에서, 제1 또는 제2 삽입 폴리뉴클레오타이드는 T 세포 수용체 알파 좌의 적어도 하나의 가변 영역 유전자 분절 및/또는 결합 영역 유전자 분절을 포함하는 관심 폴리뉴클레오타이드를 포함한다. 일 실시 형태에서, T 세포 수용체 알파 좌의 영역은 인간 유래의 것이다. 일 실시 형태에서, 제1 삽입 폴리뉴클레오타이드는 비-인간 면역글로불린 중쇄 불변 영역 핵산 서열에 작동가능하게 연결된 비재배열된 인간 면역글로불린 중쇄 가변 영역 핵산 서열을 포함하는 관심 폴리뉴클레오타이드를 포함한다. 일 실시 형태에서, 확인 단계는 대립유전자의 변형 (MOA) 검정을 통해 수행된다. 일 실시 형태에서, 제1 삽입 폴리뉴클레오타이드는 세포의 게놈 내의 핵산 서열에 상동성 또는 이중상동성인 핵산 서열을 포함하는 관심 폴리뉴클레오타이드를 포함한다. 일 실시 형태에서, 제2 삽입 폴리뉴클레오타이드는 세포의 게놈 내의 핵산 서열에 상동성 또는 이중상동성인 핵산 서열을 포함한다. 일 실시 형태에서, 제1 삽입 폴리뉴클레오타이드는 외인성 핵산 서열을 포함하는 관심 폴리뉴클레오타이드를 포함한다. 일 실시 형태에서, 제2 삽입 폴리뉴클레오타이드는 외인성 핵산 서열을 포함하는 관심 폴리뉴클레오타이드를 포함한다.

## 도면의 간단한 설명

[0042]

도 1은 게놈 표적화 이벤트에 대한 개략도를 제공하는데, 여기서는, 8개의 인간 가변 (V) 유전자 분절들 및 61개의 인간 결합 (J) 유전자 분절들의 상류에 위치한 네오마이신 선택 카세트를 포함하는, 하나의 대립유전자가 인간화된 TCR 알파 A-neo 대립유전자인, 마우스 염색체 14 상에 TCR 알파 좌의 이형접합성 (heterozygous) 변형을 갖는 세포가 하이그로마이신 선택 카세트 및 11개의 추가적인 인간 가변 유전자 분절들을 포함하는 100 kb 초과 단편을 포함하는, 인간화된 TCR 알파 대립유전자 B-hyg 표적화 벡터로 표적화된다. TCR 알파 A-neo 대립유전자에서의 네오마이신 카세트를 표적으로 하는, 대립유전자 B-hyg 표적화 벡터 및 징크 핑거 뉴클레아제 (ZFN) 쌍의 2개의 반쪽들을 발현하는 플라스미드의 전기천공이, 5'에서 3'으로, 하이그로마이신 카세트, 내인성 불변 영역 뉴클레오타이드 서열의 상류에 위치하는 19개의 인간 V 유전자 분절들, 및 61개의 인간 J 유전자 분절들을 포함하는, 변형된 TCR 알파 좌 (대립유전자 B-hyg)를 생성하였다. 표적화 이벤트는 마우스 TCR 알파 좌 내로 100 kb 초과 인간 TCR 알파 유전자 서열을 정확하게 삽입하였다.

도 2는 게놈 표적화 이벤트에 대한 개략도를 제공하는데, 여기서는, 19개의 인간 V 유전자 분절들 및 61개의 인간 J 유전자 분절들의 상류에 위치한 하이그로마이신 선택 카세트를 포함하는, 하나의 대립유전자가 인간화된 TCR 알파 B-hyg 대립유전자인, 마우스 염색체 14 상에 TCR 알파 좌의 이형접합성 변형을 갖는 세포가 네오마이신 선택 카세트 및 11개의 추가적인 인간 가변 유전자 분절들을 포함하는 100 kb 초과 단편을 포함하는, 인간화된 TCR 알파 대립유전자 C-neo 표적화 벡터로 표적화된다. TCR 알파 B-hyg 대립유전자에서의 하이그로마이신



카세트를 표적으로 하는, 대립유전자 C-neo 표적화 벡터 및 징크 핑거 뉴클레아제 (ZFN) 쌍의 2개의 반쪽들을 발현하는 플라스미드의 전기천공이, 5'에서 3'으로, 네오마이신 카세트, 내인성 불변 영역 뉴클레오티드 서열의 상류에 위치하는 30개의 인간 V 유전자 분절들, 및 61개의 인간 J 유전자 분절들을 포함하는, 변형된 TCR 알파 좌 (대립유전자 C-neo)를 생성하였다. 표적화 이벤트는 마우스 TCR 알파 좌 내로 100 kb 초과인 인간 TCR 알파 유전자 서열을 정확하게 삽입하였다.

도 3은 네오마이신 포스포트랜스퍼라제를 인코딩하는 *neo<sup>r</sup>* 약물 선택 카세트, 및 하이그로마이신 B 포스포트랜스퍼라제를 인코딩하는 *hyg<sup>r</sup>* 약물 선택 카세트의 개략도를 제공한다. *neo<sup>r</sup>*을 표적으로 하는 Neo-ZFN(1,2) 및 Neo-ZFN(3,4) 징크 핑거 뉴클레아제 (ZFN, 도 3a) 및 *hyg<sup>r</sup>*을 표적으로 하는 Hyg-ZFN(1,2) 및 Hyg-ZFN(3,4) ZFN (도 3b)에 대한 인식 부위 (아래 주어진 서열)의 위치는 각각의 포스포트랜스퍼라제 코딩 서열을 나타낸 두꺼운 화살표 위 또는 아래에 빗금친 상자로 표시되어 있다.

### 발명을 실시하기 위한 구체적인 내용

- [0043] 이제, 이하에서는 본 발명의 모든 실시 형태는 아닌 일부 실시 형태를 보여주는 첨부된 도면을 참조하여 본 발명을 더욱 상세히 설명할 것이다. 사실, 이들 발명은 다수의 상이한 형태로 구현될 수 있고 본 명세서에 제시된 실시 형태들로 제한되는 것으로 이해되어서는 안 되며; 오히려, 이들 실시 형태들은 본 개시가 적용가능한 법적 요건들을 만족하도록 제공된다. 전체에 걸쳐 유사한 도면부호는 유사한 요소들을 말한다.
- [0044] 본 명세서에 제시된 본 발명의 다수의 변형들 및 다른 실시 형태들이 전술한 설명 및 관련된 도면에 제시된 교시의 이익을 갖는, 본 발명이 속하는 기술분야의 당업자에게 착상될 것이다. 그러므로, 본 발명은 개시된 특정 실시 형태에 제한되지 않고 변형 및 다른 실시 형태들은 첨부된 청구범위의 범주 내에 포함되는 것으로 이해되어야 한다. 본 명세서에서 특정 용어가 사용되지만, 이는 제한의 목적이 아니라 포괄적이고 기술적인 의미로만 사용된다.
- [0045] I. 개요
- [0046] 세포에서 표적 좌, 예를 들어, 게놈 좌를 변형시키는 방법 및 조성물이 제공된다. 본 방법 및 조성물은 표적 좌 내로의 삽입 폴리뉴클레오티드의 상동성 재조합 이벤트를 향상시키기 위해 뉴클레아제 작용제 및 뉴클레아제 작용제 인식 부위를 사용한다. 본 명세서에 제공되는 다양한 방법 및 조성물은 뉴클레아제 작용제 인식 부위를 선택 마커, 리포터, 또는 외인성 단백질 (예를 들어, 마우스 세포 내의 eGFP 또는 인간 서열)을 인코딩하는 폴리뉴클레오티드 내에 전략적으로 위치시킨다.
- [0047] 표적 좌 (즉, 게놈 좌)에서 관심 폴리뉴클레오티드의 연속적인 변형 (즉, 타일링 (tiling))을 가능하게 하는 방법이 추가로 제공된다. 아래에서 더욱 상세하게 설명된 바와 같이, 표적 좌 (즉, 게놈 좌)에서 관심 폴리뉴클레오티드들을 순차적으로 타일링하는 방법이 제공되고, 이 방법에 사용된 표적 좌 (즉, 게놈 좌) 및 다양한 표적화 벡터들은 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 포함하는 제1 선택 마커 및 제2 뉴클레아제 작용제에 대한 제2 인식 부위를 포함하는 제2 선택 마커의 사용을 대체한다. 그렇게 함으로써, 본 방법은 새로운 인식 부위를 인식하도록 조작된 뉴클레아제들의 지속적인 공급을 필요로 하지 않는다. 대신, 특정 실시 형태에서, 표적화된 순차적인 타일링은 2개의 뉴클레아제 작용제들 및 2개의 뉴클레아제 작용제들의 상응하는 인식 부위만을 필요로 한다. 더욱이, 뉴클레아제 작용제는 외인성 서열 (즉, 선택 마커를 인코딩하는 폴리뉴클레오티드 내의 인식 부위)을 표적으로 하기 때문에 그리고 임의의 주어진 인식 부위의 효능 및 탈표적 효과 (off-target effect)는 이전에 확인되었을 것이기 때문에, 타일링 과정의 시간 및 비용 효율을 증가시키면서 내인성 게놈 서열의 비-특이적 절단이 최소화될 수 있다.
- [0048] II. 표적화된 통합 시스템
- [0049] 세포에서 표적 좌를 변형시키는 방법 및 조성물이 제공된다. 이 시스템은 뉴클레아제 작용제, 뉴클레아제 작용제에 대한 인식 부위, 표적 좌, 선택 마커, 표적화 벡터, 및 삽입 폴리뉴클레오티드를 사용한다. 이들 성분들 각각은 아래에서 더욱 상세하게 설명된다.
- [0050] A. 뉴클레아제 작용제 및 뉴클레아제 작용제에 대한 인식 부위
- [0051] 용어 "뉴클레아제 작용제에 대한 인식 부위"는 뉴클레아제 작용제에 의해 Nick 또는 이중-가닥 브레이크가 유도되는 DNA 서열을 포함한다. 뉴클레아제 작용제에 대한 인식 부위는 세포에 대해 내인성 (또는 고유의 것)일 수 있거나 인식 부위는 세포에 대해 외인성일 수 있다. 특정 실시 형태에서, 인식 부위는 세포에 대해 외인성이므로

로 세포의 게놈에서 자연적으로 발생하지 않는다. 또 다른 실시 형태에서, 인식 부위는 세포에 대해 그리고 표적 좌에 위치하기를 원하는 관심 폴리뉴클레오티드에 대해 외인성이다. 다른 실시 형태에서, 외인성 또는 내인성 인식 부위는 숙주 세포의 게놈에 한 번만 존재한다. 특정 실시 형태에서, 게놈 내에서 한 번만 발생하는 내인성 또는 고유의 부위가 확인된다. 이어서, 그러한 부위는 내인성 인식 부위에 Nick 또는 이중-가닥 브레이크를 생성할 뉴클레아제 작용제를 설계하는 데 사용될 수 있다.

[0052] 인식 부위의 길이는 다양할 수 있고, 예를 들어, 징크 핑거 뉴클레아제 (ZFN) 쌍에 대해서는 약 30 내지 36 bp (즉, 각각의 ZFN에 대해 약 15 내지 18 bp), 전사 활성화인자-유사 이펙터 뉴클레아제 (TALEN)에 대해서는 약 36 bp, 또는 CRISPR/Cas9 가이드 RNA에 대해서는 약 20 bp인 인식 부위를 포함한다.

[0053] 원하는 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도하는 임의의 뉴클레아제 작용제가 본 명세서에 개시된 방법 및 조성물에 사용될 수 있다. 자연적으로 발생하거나 고유의 뉴클레아제 작용제가 원하는 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도하는 한, 그러한 뉴클레아제 작용제가 사용될 수 있다. 대안적으로, 변형되거나 조작된 뉴클레아제 작용제가 사용될 수 있다. "조작된 뉴클레아제 작용제"는 원하는 인식 부위를 특이적으로 인식하고 Nick 또는 이중-가닥 브레이크를 유도하기 위해 그의 고유형으로부터 조작된 (변형되거나 유래한) 뉴클레아제를 포함한다. 따라서, 조작된 뉴클레아제 작용제는 고유의 자연적으로 발생한 뉴클레아제 작용제로부터 유래할 수 있거나 이는 인공적으로 만들어지거나 합성될 수 있다. 뉴클레아제 작용제의 변형은, 작게는, 단백질 절단 작용제에서의 하나의 아미노산 또는 핵산 절단 작용제에서의 하나의 뉴클레오티드일 수 있다. 일부 실시 형태에서, 조작된 뉴클레아제는 인식 부위에서 Nick 또는 이중-가닥 브레이크를 유도하는데, 여기서 인식 부위는 고유의 (비-조작된 또는 비-변형된) 뉴클레아제 작용제에 의해 인식되었을 서열이 아니었다. 본 명세서에서 인식 부위 또는 다른 DNA에서 Nick 또는 이중-가닥 브레이크를 생성하는 것은 인식 부위 또는 다른 DNA를 "자르기 (cutting)" 또는 "절단 (cleaving)"하는 것을 의미할 수 있다.

[0054] 예시된 인식 부위의 활성 변이체 및 단편이 또한 제공된다. 그러한 활성 변이체는 주어진 인식 부위에 대해 65% 이상, 70% 이상, 75% 이상, 80% 이상, 85% 이상, 90% 이상, 91% 이상, 92% 이상, 93% 이상, 94% 이상, 95% 이상, 96% 이상, 97% 이상, 98% 이상, 99% 이상 또는 그 이상의 서열 동일성을 포함할 수 있고, 여기서 활성 변이체는 생물학적 활성을 보유하고 있어서 서열-특이적 방식으로 뉴클레아제 작용제에 의해 인식되고 절단될 수 있다. 뉴클레아제 작용제에 의한 인식 부위의 이중-가닥 브레이크를 측정하는 검정법들이 당업계에 알려져 있다 (예를 들어, 태크맨® (TaqMan®) qPCR 검정법, 전체적으로 본 명세서에 참고로 포함된 문헌[Frendewey D. *et al.*, *Methods in Enzymology*, 2010, 476:295-307]).

[0055] 특정 실시 형태에서, 인식 부위는 선택 마커를 인코딩하는 폴리뉴클레오티드 내에 위치한다. 그러한 위치는, 선택 마커의 코딩 영역 내에 또는 선택 마커의 발현에 영향을 미치는 조절 영역 내에 위치할 수 있다. 따라서, 뉴클레아제 작용제의 인식 부위는 선택 마커의 인트론, 프로모터, 인핸서, 조절 영역, 또는 선택 마커를 인코딩하는 폴리뉴클레오티드의 임의의 비-단백질-코딩 영역에 위치할 수 있다. 특정 실시 형태에서, 인식 부위에서의 Nick 또는 이중-가닥 브레이크는 선택 마커의 활성을 방해한다. 기능적 선택 마커의 존재 또는 부재를 검정하기 위한 방법이 공지되어 있다.

[0056] 일 실시 형태에서, 뉴클레아제 작용제는 전사 활성화인자-유사 이펙터 뉴클레아제 (TALEN)이다. TAL 이펙터 뉴클레아제는 원핵 또는 진핵 유기체의 게놈에서 특정 표적 서열에 이중-가닥 브레이크를 만드는 데 사용될 수 있는 서열-특이적 뉴클레아제의 한 부류이다. TAL 이펙터 뉴클레아제는 고유의 또는 조작된 전사 활성화인자-유사 (TAL) 이펙터, 또는 이의 기능적 부분을 엔도뉴클레아제, 예컨대, *FokI*의 촉매 도메인에 융합시킴으로써 만들어진다. 고유의, 모듈성 TAL 이펙터 DNA 결합 도메인은, 잠재적으로 임의의 주어진 DNA 인식 특이성을 가진 단백질의 설계를 가능하게 한다. 따라서, TAL 이펙터 뉴클레아제의 DNA 결합 도메인은 특이적 DNA 표적 부위를 인식하도록 조작될 수 있고, 따라서 원하는 표적 서열에 이중-가닥 브레이크를 만드는 데 사용될 수 있다. 국제특허 공개 WO 2010/079430호; 문헌[Moritz *et al.* (2010) *PNAS* 10.1073/pnas.1013133107]; 문헌[Scholze & Boch (2010) *Virulence* 1:428-432]; 문헌[Christian *et al.* *Genetics* (2010) 186:757-761]; 문헌[Li *et al.* (2010) *Nuc. Acids Res.* (2010) doi:10.1093/nar/gkq704]; 및 문헌[Miller *et al.* (2011) *Nature Biotechnology* 29:143-148]을 참조하며; 이들 모두는 본 명세서에 참고로 포함된다.

[0057] 적합한 TAL 뉴클레아제의 예, 및 적합한 TAL 뉴클레아제를 제조하는 방법은, 예를 들어, 미국 특허 출원 공개 제2011/0239315 A1호, 제2011/0269234 A1호, 제2011/0145940 A1호, 제2003/0232410 A1호, 제2005/0208489 A1호, 제2005/0026157 A1호, 제2005/0064474 A1호, 제2006/0188987 A1호, 및 제2006/0063231 A1호에 개시되어 있다 (각각은 본 명세서에 참고로 포함된다). 다양한 실시 형태에서, TAL 이펙터 뉴클레아제는, 예를 들어, 관심

좌 또는 관심 게놈 좌에서 표적 핵산 서열 내 또는 그 부근을 자르도록 조작되고, 여기서 표적 핵산 서열은 표적화 벡터에 의해 변형될 서열에 또는 그 부근에 있다. 본 명세서에 제공되는 다양한 방법들 및 조성물들과 사용되기에 적합한 TAL 뉴클레아제는 본 명세서에 기재된 바와 같은 표적화 벡터에 의해 변형될 표적 핵산 서열에 또는 그 부근에 결합하도록 특이적으로 설계된 것을 포함한다.

[0058] 일 실시 형태에서, TALEN의 각각의 단량체는 2개의 초가변 잔기들을 통하여 단일 염기쌍을 인식하는 33개 내지 35개의 TAL 반복 (repeat)을 포함한다. 일 실시 형태에서, 뉴클레아제 작용제는 독립적인 뉴클레아제에 작동가능하게 연결된 TAL 반복-기반 DNA 결합 도메인을 포함하는 키메라 단백질이다. 일 실시 형태에서, 독립적인 뉴클레아제는 FokI 엔도뉴클레아제이다. 일 실시 형태에서, 뉴클레아제 작용제는 제1 TAL-반복-기반 DNA 결합 도메인 및 제2 TAL-반복-기반 DNA 결합 도메인을 포함하고, 여기서 제1 및 제2 TAL-반복-기반 DNA 결합 도메인 각각은 FokI 뉴클레아제 서브유닛에 작동가능하게 연결되고, 제1 및 제2 TAL-반복-기반 DNA 결합 도메인은 표적 DNA 서열의 각 가닥에서 다양한 길이 (12 내지 20 bp)의 스페이서 서열에 의해 분리된 2개의 연속 표적 DNA 서열을 인식하고, FokI 뉴클레아제 서브유닛은 이량체화되어, 표적 서열에 이중 가닥 브레이크를 만드는 활성 뉴클레아제를 만든다.

[0059] 본 명세서에 개시된 다양한 방법들 및 조성물들에 사용된 뉴클레아제 작용제는 징크-핑거 뉴클레아제 (ZFN)를 추가로 포함할 수 있다. 일 실시 형태에서, ZFN의 각각의 단량체는 3개 이상의 징크 핑거-기반 DNA 결합 도메인들을 포함하고, 여기서 각각의 징크 핑거-기반 DNA 결합 도메인은 3 bp 하위부위에 결합한다. 다른 실시 형태에서, ZFN은 독립적인 뉴클레아제에 작동가능하게 연결된 징크 핑거-기반 DNA 결합 도메인을 포함하는 키메라 단백질이다. 일 실시 형태에서, 독립적인 엔도뉴클레아제는 FokI 엔도뉴클레아제이다. 일 실시 형태에서, 뉴클레아제 작용제는 제1 ZFN 및 제2 ZFN을 포함하고, 여기서 제1 ZFN 및 제2 ZFN 각각은 FokI 뉴클레아제 서브유닛에 작동가능하게 연결되고, 제1 및 제2 ZFN은 표적 DNA 서열의 각 가닥에서 약 5 내지 7 bp 스페이서에 의해 분리된 2개의 연속 표적 DNA 서열을 인식하고, FokI 뉴클레아제 서브유닛은 이량체화되어, 이중 가닥 브레이크를 만드는 활성 뉴클레아제를 만든다. 예를 들어, 미국 특허 출원 공개 제20060246567호; 미국 특허 출원 공개 제20080182332호; 미국 특허 출원 공개 제20020081614호; 미국 특허 출원 공개 제20030021776호; 국제특허 공개 WO/2002/057308A2호; 미국 특허 출원 공개 제20130123484호; 미국 특허 출원 공개 제20100291048호; 국제특허 공개 WO/2011/017293A2호; 및 문헌[Gaj *et al.* (2013) *Trends in Biotechnology*, 31(7):397-405]을 참조하고, 이들 각각은 본 명세서에 참고로 포함된다.

[0060] 또 다른 실시 형태에서, 뉴클레아제 작용제는 메가뉴클레아제이다. 메가뉴클레아제는 보존된 서열 모티프에 기초하여 4개의 패밀리로 분류되어 왔는데, 이들 패밀리는 LAGLIDADG, GIY-YIG, H-N-H, 및 His-Cys 박스 패밀리아이다. 이들 모티프는 금속 이온들의 배위(coordination)와 포스포다이에스테르 결합의 가수분해에 참여한다. 메가뉴클레아제는 이의 긴 인식 부위, 및 이의 DNA 기질에서의 일부 서열 다형성을 용인한다는 점이 주목할 만하다. 메가뉴클레아제 도메인, 구조 및 기능은 공지되어 있고, 예를 들어, 문헌[Guhan and Muniyappa (2003) *Crit Rev Biochem Mol Biol* 38:199-248]; 문헌[Lucas *et al.*, (2001) *Nucleic Acids Res* 29:960-9]; 문헌[Jurica and Stoddard, (1999) *Cell Mol Life Sci* 55:1304-26]; 문헌[Stoddard, (2006) *Q Rev Biophys* 38:49-95]; 및 문헌[Moure *et al.*, (2002) *Nat Struct Biol* 9:764]을 참조한다. 일부 예에서는, 자연적으로 발생하는 변이체, 및/또는 조작된 유도 메가뉴클레아제가 사용된다. 동태(kinetics), 보조인자 상호작용, 발현, 최적 조건, 및/또는 인식 부위 특이성을 변형시키는 방법, 및 활성을 스크리닝하는 방법이 공지되어 있으며, 예를 들어, 문헌[Epinat *et al.*, (2003) *Nucleic Acids Res* 31:2952-62]; 문헌[Chevalier *et al.*, (2002) *Mol Cell* 10:895-905]; 문헌[Gimble *et al.*, (2003) *Mol Biol* 334:993-1008]; 문헌[Seligman *et al.*, (2002) *Nucleic Acids Res* 30:3870-9]; 문헌[Sussman *et al.*, (2004) *J Mol Biol* 342:31-41]; 문헌[Rosen *et al.*, (2006) *Nucleic Acids Res* 34:4791-800]; 문헌[Chames *et al.*, (2005) *Nucleic Acids Res* 33:e178]; 문헌[Smith *et al.*, (2006) *Nucleic Acids Res* 34:e149]; 문헌[Gruen *et al.*, (2002) *Nucleic Acids Res* 30:e29]; 문헌[Chen and Zhao, (2005) *Nucleic Acids Res* 33:e154]; 국제특허 공개 WO2005105989호; WO2003078619호; WO2006097854호; WO2006097853호; WO2006097784호; 및 WO2004031346호를 참조한다.

[0061] 임의의 메가뉴클레아제가 본 명세서에 사용될 수 있으며, 이는 I-SceI, I-SceII, I-SceIII, I-SceIV, I-SceV, I-SceVI, I-SceVII, I-CeuI, I-CeuAIIP, I-CreI, I-CrepsbIP, I-CrepsbIIP, I-CrepsbIIIP, I-CrepsbIIP, I-TliI, I-PpoI, PI-PspI, F-SceI, F-SceII, F-SuvI, F-TevI, F-TevII, I-AmaI, I-AniI, I-ChuI, I-CmoeI, I-CpaI, I-CpaII, I-CsmI, I-CvuI, I-CvuAIP, I-DdiI, I-DdiII, I-DirI, I-DmoI, I-HmuI, I-HmuII, I-HsNIP, I-LlaI, I-MsoI, I-NaaI, I-NanI, I-NcIIP, I-NgrIP, I-NitI, I-NjaI, I-Nsp236IP, I-PakI, I-PboIP, I-PcuIP, I-PcuAI, I-PcuVI, I-PgrIP, I-PobIP, I-PorI, I-PorIIP, I-PbpIP, I-SpBetaIP, I-ScaI, I-SexIP, I-SneIP,

I-SpomI, I-SpomCP, I-SpomIP, I-SpomIIP, I-SquIP, I-Ssp6803I, I-SthPhiJP, I-SthPhiST3P, I-SthPhiSTe3bP, I-TdeIP, I-TevI, I-TevII, I-TevIII, I-UarAP, I-UarHGPAIP, I-UarHGPA13P, I-VinIP, I-ZbiIP, PI-MtuI, PI-MtuHIP, PI-MtuHIIP, PI-PfuI, PI-PfuII, PI-PkoI, PI-PkoII, PI-Rma43812IP, PI-SpBetaIP, PI-SceI, PI-TfuI, PI-TfuII, PI-ThyI, PI-TliI, PI-TliII, 또는 이들의 임의의 활성 변이체 또는 단편을 포함하지만, 이로 제한되지는 않는다.

[0062] 일 실시 형태에서, 메가뉴클레아제는 12 내지 40개 염기쌍의 이중-가닥 DNA 서열을 인식한다. 일 실시 형태에서, 메가뉴클레아제는 게놈에서 하나의 완벽하게 매칭되는 표적 서열을 인식한다. 일 실시 형태에서, 메가뉴클레아제는 호밍(homing) 뉴클레아제이다. 일 실시 형태에서, 호밍 뉴클레아제는 호밍 뉴클레아제의 LAGLIDADG 패밀리아이다. 일 실시 형태에서, 호밍 뉴클레아제의 LAGLIDADG 패밀리는 I-SceI, I-CreI, 및 I-DmI로부터 선택된다.

[0063] 뉴클레아제 작용제는 제한 엔도뉴클레아제를 추가로 포함할 수 있고, 이는 타입 I, 타입 II, 타입 III, 및 타입 IV 엔도뉴클레아제를 포함한다. 타입 I 및 타입 III 제한 엔도뉴클레아제는 특이적 인식 부위를 인식하지만, 전형적으로는 뉴클레아제 결합 부위로부터 가변적인 위치에서 절단하며, 이는 절단 부위(인식 부위)로부터 수백 개의 염기쌍이 떨어진 곳일 수 있다. 타입 II 시스템에서, 제한 활성은 임의의 메틸라제 활성과는 관계가 없고, 절단은 전형적으로 결합 부위 내 또는 그 부근의 특정 부위에서 일어난다. 대부분의 타입 II 효소는 회문성 서열을 자르지만, 타입 IIa 효소는 비-회문성 인식 부위를 인식하고 그 인식 부위의 외부를 절단하고, 타입 IIb 효소는 인식 부위 외부의 양쪽 부위에서 서열을 2회 자르고, 타입 IIc 효소는 비대칭 인식 부위를 인식하고, 인식 부위로부터 약 1 내지 20개 뉴클레오타이드의 한정된 거리에서 그리고 한쪽만 절단한다. 타입 IV 제한 효소는 메틸화된 DNA를 표적으로 한다. 제한 효소는, 예를 들어 리베이스(REBASE) 데이터베이스(웹페이지: rebase.neb.com; 문헌[Roberts *et al.*, (2003) *Nucleic Acids Res* 31:418-20]), 문헌[Roberts *et al.*, (2003) *Nucleic Acids Res* 31:1805-12], 및 문헌[Belfort *et al.*, (2002) in *Mobile DNA II*, pp. 761-783], 문헌[Eds. Craigie *et al.*, (미국 워싱턴 DC 소재의 에이에스엠 프레스(ASM Press))에 추가로 설명되고 분류되어 있다.

[0064] 다양한 방법들 및 조성물들에서 사용된 뉴클레아제 작용제는 또한 CRISPR/Cas 시스템을 포함할 수 있다. 그러한 시스템은 Cas9 뉴클레아제를 사용할 수 있고, 이것은 일부 경우에, 이것이 발현되기를 원하는 세포 타입에 대해 코돈-최적화된다. 시스템은 코돈-최적화된 Cas9와 함께 기능하는 융합된 crRNA-tracrRNA 작제물을 추가로 사용한다. 이 단일 RNA는 종종 가이드 RNA 또는 gRNA라고 지칭된다. gRNA 내에, crRNA 부분은 주어진 인식 부위에 대한 '표적 서열'로서 확인되고 tracrRNA는 종종 '스캐폴드(scaffold)'로 지칭된다. 이 시스템은 다양한 진핵 및 원핵 세포에서 기능하는 것으로 밝혀져 있다. 간략히 말하면, 표적 서열을 함유하는 짧은 DNA 단편이 가이드 RNA 발현 플라스미드 내로 삽입된다. gRNA 발현 플라스미드는 tracrRNA 서열(스캐폴드) 형태인 표적 서열(일부 실시 형태에서는, 약 20개의 뉴클레오타이드)뿐만 아니라, 세포에서 활성인 적합한 프로모터 및 진핵 세포에서의 적절한 프로세싱에 필요한 요소들을 포함한다. 시스템들 중 다수는 이중 가닥 DNA를 형성하기 위해 어닐링된 후 gRNA 발현 플라스미드 내로 클로닝되는 주문 제작한 상보적인 올리고(oligo)를 이용한다. 이어서, gRNA 발현 카세트 및 Cas9 발현 카세트는 세포 내로 도입된다. 예를 들어, 문헌[Mali P *et al.* (2013) *Science* 2013 Feb 15; 339 (6121):823-6]; 문헌[Jinek M *et al.* *Science* 2012 Aug 17;337(6096):816-21]; 문헌[Hwang WY *et al.* *Nat Biotechnol* 2013 Mar;31(3):227-9]; 문헌[Jiang W *et al.* *Nat Biotechnol* 2013 Mar;31(3):233-9]; 및 문헌[Cong L *et al.* *Science* 2013 Feb 15;339(6121):819-23]을 참고하고, 이들 각각은 본 명세서에 참고로 포함된다.

[0065] 본 명세서에 개시된 방법 및 조성물은 클러스터화된 규칙적으로 사이 공간을 둔 짧은 회문성 반복(CRISPR)/CRISPR-관련(Cas) 시스템 또는 그러한 시스템의 성분들을 사용하여 세포 내의 게놈을 변형시킬 수 있다. CRISPR/Cas 시스템은 전사체 및 Cas 유전자의 발현에 관여하거나, 이의 활성을 유도하는 다른 요소들을 포함한다. CRISPR/Cas 시스템은 타입 I, 타입 II, 또는 타입 III 시스템일 수 있다. 본 명세서에 개시된 방법 및 조성물은 핵산의 부위-지정 절단에 대한 CRISPR 복합체(Cas 단백질과 복합체화된 가이드 RNA(gRNA)를 포함)를 사용함으로써 CRISPR/Cas 시스템을 사용한다.

[0066] 본 명세서에 개시된 방법에서 사용된 일부 CRISPR/Cas 시스템은 비-자연적으로 발생한다. "비-자연적으로 발생한" 시스템은, 시스템의 하나 이상의 성분들이 그들의 자연적으로 발생한 상태로부터 변경 또는 돌연변이되거나, 그들이 자연에서 자연적으로 관련된 적어도 하나의 다른 성분이 적어도 실질적으로 없거나, 그들이 자연적으로 관련되지 않은 적어도 하나의 다른 성분과 관련되는 것과 같이 사람의 손의 개입을 나타내는 임의의 것을 포함한다. 예를 들어, 일부 CRISPR/Cas 시스템은 자연적으로는 함께 발생하지 않는 gRNA 및 Cas



단백질을 포함하는 비-자연 발생 CRISPR 복합체를 사용한다.

[0067]

i. *Cas RNA-가이드된 엔도뉴클레아제*

[0068]

Cas 단백질은 일반적으로 적어도 하나의 RNA 인식 또는 결합 도메인을 포함한다. 그러한 도메인은 가이드 RNA (gRNA, 아래에 더욱 자세하게 설명됨)와 상호작용할 수 있다. Cas 단백질은 또한 뉴클레아제 도메인 (예를 들어, DNase 또는 RNase 도메인), DNA 결합 도메인, 헬리카제 도메인, 단백질-단백질 상호작용 도메인, 이량체화 도메인, 및 다른 도메인을 포함할 수 있다. 뉴클레아제 도메인은 핵산 절단을 위한 촉매 활성을 갖는다. 절단은 핵산 분자의 공유 결합의 파괴를 포함한다. 절단은 평활 말단 (blunt end) 또는 엇갈린 말단 (staggered end)을 생성할 수 있고, 이는 단일-가닥 또는 이중-가닥이 되게 할 수 있다.

[0069]

Cas 단백질의 예는 Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas5e (CasD), Cas6, Cas6e, Cas6f, Cas7, Cas8a1, Cas8a2, Cas8b, Cas8c, Cas9 (Csn1 또는 Csx12), Cas10, Cas10d, CasF, CasG, CasH, Csy1, Csy2, Csy3, Cse1 (CasA), Cse2 (CasB), Cse3 (CasE), Cse4 (CasC), Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, 및 Cu1966, 및 이들의 상동체 또는 변형된 버전을 포함한다.

[0070]

Cas 단백질은 타입 II CRISPR/Cas 시스템 유래일 수 있다. 예를 들어, Cas 단백질은 Cas9 단백질일 수 있거나 Cas9 단백질로부터 유래될 수 있다. 전형적으로, Cas9 단백질은 보존된 구조와 함께 4개의 주요 모티프를 공유한다. 모티프 1, 2, 및 4는 RuvC-유사 모티프이고, 모티프 3은 HNH 모티프이다. Cas9 단백질은, 예를 들어, 스트렙토코쿠스 피오게네스 (*Streptococcus pyogenes*), 스트렙토코쿠스 테르모필루스 (*Streptococcus thermophilus*), 스트렙토코쿠스 종(*Streptococcus sp.*), 스태필로코쿠스 아우레우스 (*Staphylococcus aureus*), 노카르디오피스 다손빌레이 (*Nocardia dasonvillei*), 스트렙토미세스 프리스티나이스피랄리스 (*Streptomyces pristinaespiralis*), 스트렙토미세스 비리도크로모게네스 (*Streptomyces viridochromogenes*), 스트렙토미세스 비리도크로모게네스, 스트렙토스포랑기움 로세움 (*Streptosporangium roseum*), 스트렙토스포랑기움 로세움, 알리시클로바실루스 아시도칼다리우스 (*Alicyclobacillus acidocaldarius*), 바실루스 슈도미코이데스 (*Bacillus pseudomycooides*), 바실루스 셀레니티레두센스 (*Bacillus selenitireducens*), 엑시구오박테리움 시비리쿰 (*Exiguobacterium sibiricum*), 락토바실루스 델브루엑키 (*Lactobacillus delbrueckii*), 락토바실루스 살리바리우스 (*Lactobacillus salivarius*), 미크로스킬라 마리나 (*Microscilla marina*), 부르크홀데리아레스 박테리움 (*Burkholderiales bacterium*), 폴라로모나스 나프탈레니보란스 (*Polaromonas naphthalenivorans*), 폴라로모나스 종 (*Polaromonas sp.*), 크로코스파이라 왓소니 (*Crocospaera watsonii*), 시아노테세 종 (*Cyanotheca sp.*), 미크로시스티스 아이루기노사 (*Microcystis aeruginosa*), 시네크로코쿠스 종 (*Synechococcus sp.*), 아세토할로비움 아라바티쿰 (*Acetohalobium arabaticum*), 암모니펙스 데젠시 (*Ammonifex degensii*), 칼디세룰로시룸토르 벅스키 (*Caldicelulosiruptor becsii*), 칸디다투스 데솔포르디스 (*Candidatus Desulforudis*), 클로스트리디움 보툴리눔 (*Clostridium botulinum*) 클로스트리디움 디피실레 (*Clostridium difficile*), 피넨골디아 마그나 (*Finegoldia magna*), 나트라나어로비우스 테르모필루스 (*Natraneroobius thermophilus*), 펠로토마쿨룸 테르모프로피오니쿰 (*Pelotomaculum thermopropionicum*), 아시디티오바실루스 칼두스 (*Acidithiobacillus caldus*), 아시디티오바실루스 페로옥시단스 (*Acidithiobacillus ferrooxidans*), 알로크로마티움 비노숨 (*Allochromatium vinosum*), 마리노박테르 종 (*Marinobacter sp.*), 니트로소코쿠스 할로필루스 (*Nitrosococcus halophilus*), 니트로소코쿠스 왓소니 (*Nitrosococcus watsoni*), 슈도알테로모나스 할로플란크티스 (*Pseudoalteromonas haloplanktis*), 크테노도박테르 라세미페르 (*Ktedonobacter racemifer*), 메타노할로비움 에베스티가툼 (*Methanohalobium vestigatum*), 아나바이나 바리아빌리스 (*Anabaena variabilis*), 노둘라리아 스푼미게나 (*Nodularia spumigena*), 노스톡 종 (*Nostoc sp.*), 아르트로스피라 막시마 (*Arthrospira maxima*), 아르트로스피라 플라텐시스 (*Arthrospira platensis*), 아르트로스피라 종 (*Arthrospira sp.*), 링비아 종 (*Lyngbya sp.*), 미크로콜레우스 크토노플라스테스 (*Microcoleus chthonoplastes*), 오스킬라토리아 종 (*Oscillatoria sp.*), 페트로토가 모빌리스 (*Petrotoğa mobilis*), 테르모시포 아프리카누스 (*Thermosipho africanus*) 또는 아카리오클로리스 마리나 (*Acaryochloris marina*)로부터 유래될 수 있다. Cas9 패밀리 멤버의 추가적인 예는 전체적으로 본 명세서에 참고로 포함되는 국제특허 공개 WO 2014/131833호에 기재되어 있다. S. 피오게네스 유래의 Cas9 단백질 또는 그로부터 유도된 Cas9 단백질이 바람직한 효소이다. S. 피오게네스 유래의 Cas9 단백질은 스위스프로트 (SwissProt) 기탁 번호 Q99ZW2로 할당된다.

[0071]

Cas 단백질은 야생형 단백질 (즉, 자연에서 발생한 것), 변형된 Cas 단백질 (즉, Cas 단백질 변이체), 또는 야생형 또는 변형된 Cas 단백질의 단편일 수 있다. Cas 단백질은 또한 야생형 또는 변형된 Cas 단백질의 활성 변

이체 또는 단편일 수 있다. 활성 변이체 또는 단편은 야생형 또는 변형된 Cas 단백질 또는 이들의 일부에 대해 80% 이상, 85% 이상, 90% 이상, 91% 이상, 92% 이상, 93% 이상, 94% 이상, 95% 이상, 96% 이상, 97% 이상, 98% 이상, 99% 이상 또는 그 이상의 서열 동일성을 포함할 수 있고, 여기서 활성 변이체는 원하는 절단 부위를 자르는 능력을 보유하고 있어서 Nick-유도 또는 이중-가닥-브레이크-유도 활성을 보유한다. Nick-유도 또는 이중-가닥-브레이크-유도 활성에 대한 검정법은 공지되어 있고, 일반적으로 절단 부위를 함유하는 DNA 기질 상에서 Cas 단백질의 전체 활성 및 특이성을 측정한다.

[0072] Cas 단백질은 핵산 결합 친화성, 핵산 결합 특이성, 및/또는 효소 활성을 증가 또는 감소시키기 위해 변형될 수 있다. Cas 단백질은 또한 단백질의 임의의 다른 활성 또는 성질, 예컨대 안정성을 변화시키기 위해 변형될 수 있다. 예를 들어, Cas 단백질의 하나 이상의 뉴클레아제 도메인은 변형되거나, 결실되거나, 비활성화될 수 있거나, 또는 Cas 단백질은 단백질의 기능에 본질적이지 않은 도메인을 제거하거나 Cas 단백질의 활성을 최적화(예를 들어, 향상 또는 감소)시키기 위해 절단될 수 있다.

[0073] 일부 Cas 단백질은 적어도 2개의 뉴클레아제 도메인, 예컨대 DNase 도메인을 포함한다. 예를 들어, Cas9 단백질은 RuvC-유사 뉴클레아제 도메인 및 HNH-유사 뉴클레아제 도메인을 포함할 수 있다. RuvC 및 HNH 도메인 각각은 이중-가닥 DNA의 상이한 가닥을 잘라서 DNA에 이중-가닥 브레이크를 만들 수 있다. 예를 들어, 전체적으로 본 명세서에 참고로 포함되는 문헌[Jinek *et al.* (2012) *Science* 337:816-821]을 참고한다.

[0074] 뉴클레아제 도메인들 중 하나 또는 둘 다는 더 이상 기능적이지 않거나 뉴클레아제 활성이 감소되도록 결실되거나 돌연변이될 수 있다. 뉴클레아제 도메인들 중 하나가 결실되거나 돌연변이된다면, 생성되는 Cas 단백질(예를 들어, Cas9)은 Nickase(nickase)로서 지칭될 수 있고 이중-가닥 DNA 내 CRISPR RNA 인식 서열에 이중-가닥 브레이크가 아니라 단일-가닥 브레이크를 생성할 수 있다(즉, 이것은 상보적 가닥과 비-상보적 가닥 둘 다가 아니라 상보적 가닥 또는 비-상보적 가닥을 절단할 수 있다). 뉴클레아제 도메인 둘 다가 결실되거나 돌연변이된다면, 생성되는 Cas 단백질(예를 들어, Cas9)은 이중-가닥 DNA의 가닥 둘 다를 절단하는 능력이 감소될 것이다. Cas9을 Nickase로 변환시키는 돌연변이의 예는 *S. 피오게네스* 유래의 Cas9의 RuvC 도메인에서의 D10A(Cas9의 위치 10에서 아스파르트산을 알라닌으로) 돌연변이이다. 마찬가지로, *S. 피오게네스* 유래의 Cas9의 HNH 도메인에서의 H939A(아미노산 위치 839에서 히스티딘을 알라닌으로) 또는 H840A(아미노산 위치 840에서 히스티딘을 알라닌으로)가 Cas9을 Nickase로 변환시킬 수 있다. Cas9을 Nickase로 변환시키는 돌연변이의 다른 예는 *S. 테르모필루스* 유래의 Cas9에 상응하는 돌연변이를 포함한다. 예를 들어, 각각이 전체적으로 본 명세서에 참고로 포함되는 문헌[Sapranaukas *et al.* (2011) *Nucleic Acids Research* 39:9275-9282] 및 국제특허 공개 WO 2013/141680호를 참조한다. 그러한 돌연변이는 부위-지정 돌연변이생성, PCR-매개 돌연변이생성, 또는 전체 유전자 합성과 같은 방법을 사용하여 생성될 수 있다. Nickase를 만드는 다른 돌연변이의 예는, 예를 들어, 국제특허 공개 WO/2013/176772A1호 및 WO/2013/142578A1호에서 찾을 수 있고, 이들 각각은 본 명세서에 참고로 포함된다.

[0075] Cas 단백질은 또한 융합 단백질일 수 있다. 예를 들어, Cas 단백질은 절단 도메인, 후생적 변형 도메인, 전사 활성화 도메인, 또는 전사 억제인자 도메인에 융합될 수 있다. 전체적으로 본 명세서에 참고로 포함되는 국제특허 공개 WO 2014/089290호를 참조한다. Cas 단백질은 또한 증가된 또는 감소된 안정성을 제공하는 이중성 폴리펩티드에 융합될 수 있다. 융합된 도메인 또는 이중성 폴리펩티드는 N-말단에, C-말단에, 또는 내부적으로 Cas 단백질 내에 위치할 수 있다.

[0076] Cas 단백질은 하위세포 국소화(localization)를 위해 제공되는 이중성 폴리펩티드에 융합될 수 있다. 그러한 이중성 펩티드는, 예를 들어, 핵 국소화 신호(nuclear localization signal)(NLS), 예컨대 핵으로의 표적화를 위한 SV40 NLS, 미토콘드리아로의 표적화를 위한 미토콘드리아 국소화 신호(mitochondrial localization signal), ER 잔류 신호 등을 포함한다. 예를 들어, 문헌[Lange *et al.* (2007) *J. Biol. Chem.* 282:5101-5105]을 참조한다. 그러한 하위세포 국소화 신호는 N-말단에, C-말단에, 또는 Cas 단백질 내의 어디든지에 위치할 수 있다. NLS는 염기성 아미노산들의 스트레치(stretch)를 포함할 수 있고, 단립형(monopartite) 서열 또는 양립형(bipartite) 서열일 수 있다.

[0077] Cas 단백질은 또한 세포-투과 도메인에 연결될 수 있다. 예를 들어, 세포-투과 도메인은 HIV-1 TAT 단백질, 인간 B형 간염 바이러스로부터의 TLM 세포-투과 모티프, MPG, Pep-1, VP22, 단순 포진 바이러스로부터의 세포 투과 펩티드, 또는 폴리아르기닌 펩티드 서열로부터 유래될 수 있다. 예를 들어, 전체적으로 본 명세서에 참고로 포함되는 국제특허 공개 WO 2014/089290호를 참조한다. 세포-투과 도메인은 N-말단에, C-말단에, 또는 Cas 단백질 내의 어디든지에 위치할 수 있다.

- [0078] Cas 단백질은 또한 추적 또는 정제를 용이하게 하기 위한 이중성 폴리펩티드, 예컨대 형광 단백질, 정제 태그, 또는 에피토프 태그를 포함할 수 있다. 형광 단백질의 예는 녹색 형광 단백질 (예를 들어, GFP, GFP-2, 태그 GFP (tagGFP), 터보GFP (turboGFP), eGFP, 에메랄드 (Emerald), 아자미 그린 (Azami Green), 단량체성 (Monomeric) 아자미 그린, 콕GFP (CopGFP), 에이스GFP (AceGFP), 지에스그린1 (ZsGreen1)), 황색 형광 단백질 (예를 들어, YFP, eYFP, 시트린 (Citrine), 비너스 (Venus), 와이펫 (YPet), 파이YFP (PhiYFP), 지에스옐로우1 (ZsYellow1)), 청색 형광 단백질 (예를 들어, eBFP, eBFP2, 아주라이트 (Azurite), 엠칼라말 (mKalamal), GFPuv, 사파이어 (Sapphire), T-사파이어)), 시안 형광 단백질 (예를 들어 eCFP, 세룰리안 (Cerulean), 사이펫 (CyPet), 에이엠시안1 (AmCyan1), 미도리시-시안 (Midoriishi-Cyan)), 적색 형광 단백질 (엠케이트 (mKate), 엠케이트2, 엠플럼 (mPlum), 디에스레드 (DsRed) 단량체, 엠체리 (mCherry), mRFP1, 디에스레드-익스프레스 (DsRed-Express), 디에스레드2, 디에스레드-단량체, 에이치시레드-탠덤 (HcRed-Tandem), 에이치시레드1 (HcRed1), 에이에스레드2 (AsRed2), eqFP611, 엠라즈베리 (mRaspberry), 엠스트로베리 (mStrawberry), 제이레드 (Jred)), 오렌지색 형광 단백질 (엠오렌지 (mOrange), mKO, 쿠사비라-오렌지 (Kusabira-Orange), 단량체성 쿠사비라-오렌지, 엠타렌지 (mTangerine), 티디토마토 (tdTomato)), 및 임의의 다른 적합한 형광 단백질을 포함한다. 태그의 예는 글루타티온-S-트랜스퍼라제 (GST), 키틴 결합 단백질 (CBP), 말토스 결합 단백질, 티오레독신 (TRX), 폴리(NANP), 탠덤 친화성 정제 (TAP) 태그, myc, AcV5, AU1, AU5, E, ECS, E2, FLAG, 혈구응집소 (HA), nus, 소프트태그 (Softag) 1, 소프트태그 3, 스트랩 (Strep), SBP, Glu-Glu, HSV, KT3, S, S1, T7, V5, VSV-G, 히스티딘 (His), 비오틴 카르복실 캐리어 단백질 (BCCP), 및 칼모둘린을 포함한다.
- [0079] Cas 단백질은 임의의 형태로 제공될 수 있다. 예를 들어, Cas 단백질은 단백질의 형태, 예컨대 gRNA와 복합체화된 Cas 단백질의 형태로 제공될 수 있다. 대안적으로, Cas 단백질은 Cas 단백질을 인코딩하는 핵산의 형태, 예컨대 RNA (예를 들어, 메신저 RNA (mRNA)) 또는 DNA로 제공될 수 있다. 선택적으로, Cas 단백질을 인코딩하는 핵산은 특정 세포 또는 유기체에서 단백질로의 효율적인 번역에 최적화된 코돈일 수 있다.
- [0080] Cas 단백질을 인코딩하는 핵산은 세포의 게놈에 안정적으로 통합될 수 있고 세포에서 활성인 프로모터에 작동가능하게 연결될 수 있다. 대안적으로, Cas 단백질을 인코딩하는 핵산은 발현 작제물에서 프로모터에 작동가능하게 연결될 수 있다. 발현 작제물은, 관심 유전자 또는 다른 핵산 서열 (예를 들어, Cas 유전자)의 발현을 유도할 수 있고 그런 관심 핵산 서열을 표적 세포로 전달할 수 있는 임의의 핵산 작제물을 포함한다. 발현 작제물에 사용될 수 있는 프로모터는, 예를 들어, 다능성 래트 세포, 진핵 세포, 포유류 세포, 비-인간 포유류 세포, 인간 세포, 설치류 세포, 마우스 세포, 또는 햄스터 세포에서 활성인 프로모터를 포함한다. 다른 프로모터의 예가 본 명세서의 어딘가 다른 곳에 설명되어 있다.
- [0081] *ii. 가이드 RNA (gRNA)*
- [0082] "가이드 RNA" 또는 "gRNA"는, Cas 단백질에 결합하고 표적 DNA 내의 특이적 위치로 Cas 단백질을 표적화하는 RNA 분자를 포함한다. 가이드 RNA는 다음 2개의 분절을 포함할 수 있다: "DNA-표적화 분절" 및 "단백질-결합 분절". "분절"은 분자의 분절, 부분 (section), 또는 영역, 예컨대 RNA에서 뉴클레오타이드들의 인접 스트레치를 포함한다. 일부 gRNA는 다음 2개의 별개의 RNA 분자들을 포함한다: "활성화인자-RNA" 및 "표적자 (targeter)-RNA". 다른 gRNA는 단일 RNA 분자 (단일 RNA 폴리뉴클레오타이드)이고, 이는 "단일-분자 gRNA", "단일-가이드 RNA", 또는 "sgRNA"로도 지칭된다. 예를 들어, 국제특허 공개 WO/2013/176772A1호, WO/2014/065596A1호, WO/2014/089290A1호, WO/2014/093622A2호, WO/2014/099750A2호, WO/2013142578A1호, 및 WO 2014/131833A1호를 참조하고, 이들 각각은 본 명세서에 참고로 포함된다. 용어 "가이드 RNA" 및 "gRNA"는 이중-분자 gRNA 및 단일-분자 gRNA 둘 다를 포함한다.
- [0083] 예시적인 2-분자 gRNA는 crRNA-유사 ("CRISPR RNA" 또는 "표적자-RNA" 또는 "crRNA" 또는 "crRNA 반복") 분자 및 상응하는 tracrRNA-유사 ("트랜스-활성화 CRISPR RNA" 또는 "활성화인자-RNA" 또는 "tracrRNA" 또는 "스캐폴드") 분자를 포함한다. crRNA는 gRNA의 DNA-표적화 분절 (단일-가닥으로 됨) 및 gRNA의 단백질-결합 분절의 dsRNA 듀플렉스 (duplex)의 반쪽을 형성하는 뉴클레오타이드들의 스트레치 둘 다를 포함한다.
- [0084] 상응하는 tracrRNA (활성화인자-RNA)는 gRNA의 단백질-결합 분절의 dsRNA 듀플렉스의 반쪽을 형성하는 뉴클레오타이드들의 스트레치를 포함한다. crRNA의 뉴클레오타이드들의 스트레치는 tracrRNA의 뉴클레오타이드의 스트레치에 상보적이고 이와 혼성화되어 gRNA의 단백질-결합 도메인의 dsRNA 듀플렉스를 형성한다. 그렇기 때문에, 각각의 crRNA는 상응하는 tracrRNA를 갖고 있다고 할 수 있다.
- [0085] crRNA와 상응하는 tracrRNA는 혼성화되어 gRNA를 형성한다. crRNA는 추가적으로 CRISPR RNA 인식 서열에 혼성화되는 단일-가닥 DNA-표적화 분절을 제공한다. 세포 내에서의 변형에 사용된다면, 주어진 crRNA 또는

tracrRNA 분자의 정확한 서열은 RNA 분자가 사용될 중에 특이적이 되도록 설계될 수 있다. 예를 들어, 문헌[Mali *et al.* (2013) *Science* 339:823-826]; 문헌[Jinek *et al.* (2012) *Science* 337:816-821]; 문헌[Hwang *et al.* (2013) *Nat. Biotechnol.* 31:227-229]; 문헌[Jiang *et al.* (2013) *Nat. Biotechnol.* 31:233-239]; 및 문헌[Cong *et al.* (2013) *Science* 339:819-823]을 참조하고, 이들 각각은 본 명세서에 참고로 포함된다.

[0086] 주어진 gRNA의 DNA-표적화 분절 (crRNA)은 표적 DNA에서의 서열에 상보적인 뉴클레오티드 서열을 포함한다. gRNA의 DNA-표적화 분절은 혼성화 (즉, 염기쌍 형성)을 통해 서열-특이적 방식으로 표적 DNA와 상호작용한다. 그렇기 때문에, DNA-표적화 분절의 뉴클레오티드 서열은 다양할 수 있고, gRNA 및 표적 DNA와 상호작용할 표적 DNA 내의 위치를 결정한다. 대상 gRNA의 DNA-표적화 분절은 표적 DNA 내의 임의의 원하는 서열에 혼성화되도록 변형될 수 있다. 자연적으로 발생한 crRNA는 Cas9 시스템 및 유기체에 따라 다르지만, 21 내지 46개 뉴클레오티드 길이의 2개의 직접 반복 (DR)에 의해 플랭킹된 21 내지 72개 뉴클레오티드 길이의 표적화 분절을 종종 함유한다 (예를 들어, 국제특허 공개 WO2014/131833호를 참조한다). S. 피로게네스의 경우, DR은 36개 뉴클레오티드 길이이고 표적화 분절은 30개 뉴클레오티드 길이이다. 3'에 위치한 DR은 상응하는 tracrRNA에 상보적이고 이와 혼성화되며, 이는 다시, Cas9 단백질에 결합한다.

[0087] DNA-표적화 분절의 길이는 약 12개 뉴클레오티드 내지 약 100개 뉴클레오티드일 수 있다. 예를 들어, DNA-표적화 분절의 길이는 약 12개 뉴클레오티드 (nt) 내지 약 80 nt, 약 12 nt 내지 약 50 nt, 약 12 nt 내지 약 40 nt, 약 12 nt 내지 약 30 nt, 약 12 nt 내지 약 25 nt, 약 12 nt 내지 약 20 nt, 또는 약 12 nt 내지 약 19 nt 일 수 있다. 대안적으로, DNA-표적화 분절의 길이는 약 19 nt 내지 약 20 nt, 약 19 nt 내지 약 25 nt, 약 19 nt 내지 약 30 nt, 약 19 nt 내지 약 35 nt, 약 19 nt 내지 약 40 nt, 약 19 nt 내지 약 45 nt, 약 19 nt 내지 약 50 nt, 약 19 nt 내지 약 60 nt, 약 19 nt 내지 약 70 nt, 약 19 nt 내지 약 80 nt, 약 19 nt 내지 약 90 nt, 약 19 nt 내지 약 100 nt, 약 20 nt 내지 약 25 nt, 약 20 nt 내지 약 30 nt, 약 20 nt 내지 약 35 nt, 약 20 nt 내지 약 40 nt, 약 20 nt 내지 약 45 nt, 약 20 nt 내지 약 50 nt, 약 20 nt 내지 약 60 nt, 약 20 nt 내지 약 70 nt, 약 20 nt 내지 약 80 nt, 약 20 nt 내지 약 90 nt, 또는 약 20 nt 내지 약 100 nt일 수 있다.

[0088] 표적 DNA의 뉴클레오티드 서열 (CRISPR RNA 인식 서열)에 상보적인 DNA-표적화 분절의 뉴클레오티드 서열의 길이는 약 12 nt 이상일 수 있다. 예를 들어, DNA-표적화 서열 (즉, 표적 DNA 내의 CRISPR RNA 인식 서열에 상보적인, DNA-표적화 분절 내의 서열)의 길이는 약 12 nt 이상, 약 15 nt 이상, 약 18 nt 이상, 약 19 nt 이상, 약 20 nt 이상, 약 25 nt 이상, 약 30 nt 이상, 약 35 nt 이상, 또는 약 40 nt 이상일 수 있다. 대안적으로, DNA-표적화 서열의 길이는 약 12개 뉴클레오티드 (nt) 내지 약 80 nt, 약 12 nt 내지 약 50 nt, 약 12 nt 내지 약 45 nt, 약 12 nt 내지 약 40 nt, 약 12 nt 내지 약 35 nt, 약 12 nt 내지 약 30 nt, 약 12 nt 내지 약 25 nt, 약 12 nt 내지 약 20 nt, 약 12 nt 내지 약 19 nt, 약 19 nt 내지 약 20 nt, 약 19 nt 내지 약 25 nt, 약 19 nt 내지 약 30 nt, 약 19 nt 내지 약 35 nt, 약 19 nt 내지 약 40 nt, 약 19 nt 내지 약 45 nt, 약 19 nt 내지 약 50 nt, 약 19 nt 내지 약 60 nt, 약 20 nt 내지 약 25 nt, 약 20 nt 내지 약 30 nt, 약 20 nt 내지 약 35 nt, 약 20 nt 내지 약 40 nt, 약 20 nt 내지 약 45 nt, 약 20 nt 내지 약 50 nt, 또는 약 20 nt 내지 약 60 nt일 수 있다. 일부 경우에, DNA-표적화 서열의 길이는 약 20 nt일 수 있다.

[0089] tracrRNA는 임의의 형태 (예를 들어, 전장 (full-length) tracrRNA 또는 활성의 부분 tracrRNA)일 수 있고 길이가 다양할 수 있다. 이는 일차 (primary) 전사체 또는 가공된 형태를 포함할 수 있다. 예를 들어, (단일-가이드 RNA의 일부로서의 또는 2-분자 gRNA의 일부로서의 별개 분자로서의) tracrRNA는 야생형 tracrRNA 서열의 전부 또는 일부 (예를 들어, 야생형 tracrRNA 서열의 대략 하기 개수의 또는 하기 개수를 초과하는 뉴클레오티드: 약 20, 26, 32, 45, 48, 54, 63, 67, 85개, 또는 그 이상)를 포함하거나 이로 이루어질 수 있다. S. 피로게네스 유래의 야생형 tracrRNA 서열의 예는 171-뉴클레오티드, 89-뉴클레오티드, 75-뉴클레오티드, 및 65-뉴클레오티드 버전을 포함한다. 예를 들어, 문헌[Deltcheva *et al.* (2011) *Nature* 471:602-607]; 국제특허 공개 WO 2014/093661호를 참조하고, 이들 각각은 전체적으로 본 명세서에 참고로 포함된다. 단일-가이드 RNA (sgRNA) 내의 tracrRNA의 예는 sgRNA의 +48, +54, +67, 및 +85 버전 내에서 발견되는 tracrRNA 분절을 포함하고, 여기서 "+n"은 야생형 tracrRNA의 최대 +n 뉴클레오티드가 sgRNA 내에 포함된다는 것을 나타낸다. 전체적으로 본 명세서에 참고로 포함되는 미국 특허 제8,697,359호를 참조한다.

[0090] 표적 DNA 내의 DNA-표적화 서열과 CRISPR RNA 인식 서열 사이의 %상보성은 60% 이상 (예를 들어, 65% 이상, 70% 이상, 75% 이상, 80% 이상, 85% 이상, 90% 이상, 95% 이상, 97% 이상, 98% 이상, 99% 이상, 또는 100%)일 수 있다. DNA-표적화 서열과 표적 DNA 내의 CRISPR RNA 인식 서열 사이의 %상보성은 약 20개의 인접 뉴클레오티드들에 걸쳐 60% 이상일 수 있다. 예로서, DNA-표적화 서열과 표적 DNA 내의 CRISPR RNA 인식 서열 사이의 %상보성은 표적 DNA의 상보적 가닥 내의 CRISPR RNA 인식 서열의 5' 말단에서 14개의 인접 뉴클레오티드에 걸쳐



100%이고 나머지에 걸쳐서는 0%만큼 낮다. 그러한 경우에, DNA-표적화 서열은 14개 뉴클레오티드 길이인 것으로 간주될 수 있다. 다른 예로서, DNA-표적화 서열과 표적 DNA 내의 CRISPR RNA 인식 서열 사이의 %상보성은 표적 DNA의 상보적 가닥 내의 CRISPR RNA 인식 서열의 5' 말단에서 7개의 인접 뉴클레오티드에 걸쳐 100%이고 나머지에 걸쳐서는 0%만큼 낮다. 그러한 경우에, DNA-표적화 서열은 7개 뉴클레오티드 길이인 것으로 간주될 수 있다.

[0091] gRNA의 단백질-결합 분절은 서로 상보적인 뉴클레오티드들의 2개의 스트레치를 포함할 수 있다. 단백질-결합 분절의 상보적 뉴클레오티드는 혼성화되어 이중-가닥 RNA 듀플렉스 (dsRNA)를 형성한다. 대상 gRNA의 단백질-결합 분절은 Cas 단백질과 상호작용하고, gRNA는 결합된 Cas 단백질을 DNA-표적화 분절을 통해 표적 DNA 내의 특정 뉴클레오티드 서열로 유도한다.

[0092] 가이드 RNA는 추가적인 바람직한 특성 (예를 들어, 변형된 또는 조절된 안정성; 하위세포 표적화; 형광 표지를 사용한 추적; 단백질 또는 단백질 복합체에 대한 결합 부위 등)을 제공하는 변형 또는 서열을 포함할 수 있다. 그러한 변형의 예는, 예를 들어, 5' 캡 (예를 들어, 7-메틸구아닐레이트 캡 (m7G)); 3' 폴리아데닐화된 꼬리 (즉, 3' 폴리(A) 꼬리); (예를 들어, 단백질 및/또는 단백질 복합체에 의한 조절된 안정성 및/또는 조절된 접근성을 허용하기 위한) 리보스위치 (riboswitch) 서열; 안정성 제어 서열; dsRNA 듀플렉스 (즉, 헤어핀 (hairpin))를 형성하는 서열; RNA를 하위세포 위치 (예를 들어, 핵, 미토콘드리아, 엽록체 등)로 표적화시키는 변형 또는 서열; 추적 (예를 들어, 형광 분자에 대한 직접 접합 (conjugation), 형광 검출을 촉진하는 모이어티 (moiety)에 대한 접합, 형광 검출을 허용하는 서열 등)을 제공하는 변형 또는 서열; 단백질 (예를 들어, 전사 활성화인자, 전사 억제인자, DNA 메틸트랜스퍼라제, DNA 데메틸라제, 히스톤 아세틸트랜스퍼라제, 히스톤 데아세틸라제 등을 포함하는, DNA 상에서 작용하는 단백질)에 대한 결합 부위를 제공하는 변형 또는 서열; 및 이들의 조합을 포함한다.

[0093] 가이드 RNA는 임의의 형태로 제공될 수 있다. 예를 들어, gRNA는 2개의 분자 (별개의 crRNA 및 tracrRNA)로서 또는 1개의 분자 (sgRNA)로서 RNA의 형태로, 그리고 선택적으로는 Cas 단백질과의 복합체의 형태로 제공될 수 있다. gRNA는 또한 RNA를 인코딩하는 DNA의 형태로 제공될 수 있다. gRNA를 인코딩하는 DNA는 단일 RNA 분자 (sgRNA) 또는 별개의 RNA 분자들 (예를 들어, 별개의 crRNA 및 tracrRNA)을 인코딩할 수 있다. 후자의 경우, gRNA를 인코딩하는 DNA는 crRNA 및 tracrRNA를 각각 인코딩하는 별개의 DNA 분자로서 제공될 수 있다.

[0094] gRNA를 인코딩하는 DNA는 세포의 게놈에 안정적으로 통합될 수 있고 세포에서 활성인 프로모터에 작동가능하게 연결될 수 있다. 대안적으로, gRNA를 인코딩하는 DNA는 발현 작제물에서 프로모터에 작동가능하게 연결될 수 있다. 그러한 프로모터는, 예를 들어, 다능성 래트 세포, 진행 세포, 포유류 세포, 비-인간 포유류 세포, 인간 세포, 설치류 세포, 마우스 세포, 또는 햄스터 세포에서 활성일 수 있다. 일부 경우에, 프로모터는 RNA 폴리머라제 III 프로모터, 예컨대 인간 U6 프로모터, 래트 U6 폴리머라제 III 프로모터, 또는 마우스 U6 폴리머라제 III 프로모터이다. 다른 프로모터의 예가 본 명세서의 어딘가 다른 곳에 설명되어 있다.

[0095] 대안적으로, gRNA는 다양한 다른 방법으로 제조될 수 있다. 예를 들어, gRNA는, 예를 들어, T7 RNA 폴리머라제를 사용하는 시험관내 (*in vitro*) 전사에 의해 제조될 수 있다 (예를 들어, 국제특허 공개 WO 2014/089290호 및 WO 2014/065596호 참조). 가이드 RNA는 또한 화학적 합성에 의해 제조된, 합성적으로 생산된 분자일 수 있다.

[0096] *iii. CRISPR RNA 인식 서열*

[0097] 용어 "CRISPR RNA 인식 서열"은, 결합에 대한 충분한 조건이 존재하기만 한다면, gRNA의 DNA-표적화 분절이 결합할, 표적 DNA에 존재하는 핵산 서열을 포함한다. 예를 들어, CRISPR RNA 인식 서열은 가이드 RNA가 상보성을 가지도록 설계된 서열을 포함하고, 여기서 CRISPR RNA 인식 서열과 DNA 표적화 서열 사이의 혼성화는 CRISPR 복합체의 형성을 촉진한다. 완전한 상보성이 반드시 필요한 것은 아니고, 혼성화를 야기하고 CRISPR 복합체의 형성을 촉진하기에 충분한 상보성이 있기만 하면 된다. CRISPR RNA 인식 서열은 또한 Cas 단백질에 대한 절단 부위를 포함하며, 이는 아래에 더욱 상세하게 설명되어 있다. CRISPR RNA 인식 서열은, 예를 들어, 세포의 핵 또는 세포질에 또는 세포의 세포소기관, 예컨대 미토콘드리아 또는 엽록체 내에 위치할 수 있는 임의의 폴리뉴클레오티드를 포함할 수 있다.

[0098] 표적 DNA 내의 CRISPR RNA 인식 서열은 Cas 단백질 또는 gRNA에 의해 표적화될 수 있다 (즉, 이에 의해 결합되거나, 이와 혼성화되거나, 이에 상보적일 수 있다). 적합한 DNA/RNA 결합 조건은 세포에 정상적으로 존재하는 생리적 조건을 포함한다. 다른 적합한 DNA/RNA 결합 조건 (예를 들어, 무세포 시스템에서의 조건)은 당업계에 공지되어 있다 (예를 들어, 문헌[Molecular Cloning: A Laboratory Manual, 3rd Ed. (Sambrook *et al.*,

Harbor Laboratory Press 2001)] 참조). Cas 단백질 또는 gRNA에 상보적이고 이와 혼성화되는 표적 DNA의 가닥은 "상보적 가닥"으로 불릴 수 있고, "상보적 가닥"에 상보적인 (따라서 Cas 단백질 또는 gRNA에 상보적이지 않은) 표적 DNA의 가닥은 "비상보적 가닥" 또는 "주형 가닥"으로 불릴 수 있다.

[0099] Cas 단백질은 gRNA의 DNA-표적화 분절이 결합할 표적 DNA에 존재하는 핵산 서열 내의 또는 그 외부의 부위에서 핵산을 절단할 수 있다. "절단 부위"는 Cas 단백질이 단일-가닥 브레이크 또는 이중-가닥 브레이크를 생성하는 핵산의 위치를 포함한다. 예를 들어, (CRISPR RNA 인식 서열에 혼성화되고 Cas 단백질과 복합체화된 gRNA를 포함하는) CRISPR 복합체의 형성은 gRNA의 DNA-표적화 분절이 결합할 표적 DNA에 존재하는 핵산 서열 내에서 또는 그 부근에서 (예를 들어, 이로부터 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 50개, 또는 그 이상의 염기쌍 이내에) 한 가닥 또는 두 가닥 모두의 절단을 야기할 수 있다. 절단 부위가 gRNA의 DNA-표적화 분절이 결합할 핵산 서열의 외부에 있다면, 절단 부위는 "CRISPR RNA 인식 서열" 내에 있는 것으로 여전히 간주된다. 절단 부위는 핵산의 단지 한 가닥 상에 있거나 두 가닥 모두의 상에 있을 수 있다. 절단 부위는 핵산의 두 가닥 모두 상의 동일한 위치에 있을 수 있거나 (평할 말단을 생성), 또는 각 가닥 상의 상이한 부위에 있을 수 있다 (엇갈린 말단을 생성). 엇갈린 말단은, 예를 들어, 2개의 Cas 단백질을 사용하여 생성될 수 있고, 이들 각각은 각 가닥 상의 상이한 절단 부위에서 단일-가닥 브레이크 생성하여 이중-가닥 브레이크를 생성한다. 예를 들어, 제1 Nickase는 이중-가닥 DNA (dsDNA)의 제1 가닥 상에 단일-가닥 브레이크를 만들 수 있고, 제2 Nickase는 dsDNA의 제2 가닥 상에 단일-가닥 브레이크를 만들 수 있어서 돌출된 (overhanging) 서열이 만들어지게 된다. 일부 경우에, 제1 가닥에서의 Nickase의 CRISPR RNA 인식 서열은 2개 이상, 3개 이상, 4개 이상, 5개 이상, 6개 이상, 7개 이상, 8개 이상, 9개 이상, 10개 이상, 15개 이상, 20개 이상, 25개 이상, 30개 이상, 40개 이상, 50개 이상, 75개 이상, 100개 이상, 250개 이상, 500개 이상, 또는 1,000개 이상의 염기쌍에 의해 제2 가닥 상의 Nickase의 CRISPR RNA 인식 서열로부터 분리된다.

[0100] Cas9에 의한 표적 DNA의 부위-특이적 절단은 표적 DNA에서 (i) gRNA와 표적 DNA 사이의 염기쌍 형성 상보성 및 (ii) 프로토스페이서 인접 모티프 (PAM)로 불리는, 짧은 모티프 둘 다에 의해 결정되는 위치에서 일어날 수 있다. PAM은 CRISPR RNA 인식 서열을 플랜킹할 수 있다. 선택적으로, CRISPR RNA 인식 서열은 PAM에 의해 플랜킹될 수 있다. 예를 들어, Cas9의 절단 부위는 PAM 서열의 상류 또는 하류의 약 1 내지 약 10개 또는 약 2 내지 약 5개의 염기쌍 (예를 들어, 3개의 염기쌍)일 수 있다. 일부 경우에 (예를 들어, S. 피오게네스로부터의 Cas9 또는 근연(closely related) Cas9가 사용될 때), 비-상보적 가닥의 PAM 서열은 5'-N<sub>1</sub>GG-3'일 수 있고, 여기서 N<sub>1</sub>은 임의의 DNA 뉴클레오티드이고 표적 DNA의 비-상보적 가닥의 CRISPR RNA 인식 서열의 3'에 바로 있다. 그렇기 때문에, 상보적 가닥의 PAM 서열은 5'-CC N<sub>2</sub>-3'일 것이고, 여기서 N<sub>2</sub>는 임의의 DNA 뉴클레오티드이고 표적 DNA의 상보적 가닥의 CRISPR RNA 인식 서열의 5'에 바로 있다. 일부 그러한 경우에, N<sub>1</sub>과 N<sub>2</sub>는 상보적일 수 있고 N<sub>1</sub>-N<sub>2</sub> 염기쌍은 임의의 염기쌍 (예를 들어, N<sub>1</sub>=C 및 N<sub>2</sub>=G; N<sub>1</sub>=G 및 N<sub>2</sub>=C; N<sub>1</sub>=A 및 N<sub>2</sub>=T, N<sub>1</sub>=T, 및 N<sub>2</sub>=A)일 수 있다.

[0101] CRISPR RNA 인식 서열의 예는 gRNA의 DNA-표적화 분절에 상보적인 DNA 서열, 또는 PAM 서열에 덧붙여진 그러한 DNA 서열을 포함한다. 예를 들어, 표적 모티프는 Cas 단백질에 의해 인식되는 NGG 모티프 직전의 20-뉴클레오티드 DNA 서열일 수 있다 (예를 들어, 국제특허 공개 WO 2014/165825호 참조). 5' 말단에서의 구아닌은 세포에서 RNA 폴리머라제에 의한 전사를 촉진할 수 있다. CRISPR RNA 인식 서열의 다른 예는 시험관내에서의 T7 폴리머라제에 의한 효율적인 전사를 촉진하기 위해 5' 말단에 2개의 구아닌 뉴클레오티드를 포함할 수 있다 (예를 들어, GGN<sub>20</sub>NGG; 서열 번호 21). 예를 들어, 국제특허 공개 WO 2014/065596호를 참조한다.

[0102] CRISPR RNA 인식 서열은 세포에 내인성 또는 외인성인 임의의 핵산 서열일 수 있다. CRISPR RNA 인식 서열은 유전자 산물 (예를 들어, 단백질)을 코딩하는 서열 또는 비-코딩 서열 (예를 들어, 조절 서열)일 수 있고 둘 다를 포함할 수 있다. 일 실시 형태에서, 표적 서열은 프로토스페이서 인접 모티프 (PAM) 서열에 의해 바로 플랜킹된다. 일 실시 형태에서, 관심 좌는 서열 번호 1의 뉴클레오티드 서열을 포함한다. 일 실시 형태에서, gRNA는 클러스터화된 규칙적으로 사이 공간을 둔 짧은 회문성 반복 (CRISPR) RNA (crRNA) 및 트랜스-활성화 CRISPR RNA (tracrRNA)를 인코딩하는 제3 핵산 서열을 포함한다. 다른 실시 형태에서, 다능성 래트 세포의 계놈은 표적 서열에 상보적인 표적 DNA 영역을 포함한다. 일부 그러한 방법에서, Cas 단백질은 Cas9이다. 일부 실시 형태에서, gRNA는 (a) 서열 번호 2의 핵산 서열의 키메라 RNA; 또는 (b) 서열 번호 3의 핵산 서열의 키메라 RNA를 포함한다. 일부 그러한 방법에서, crRNA는 서열 번호 4, 서열 번호 5, 또는 서열 번호 6에 제시된 서열을 포함한다. 일부 그러한 방법에서, tracrRNA는 서열 번호 7 또는 서열 번호 8에 제시된 서열을 포함한다.

- [0103] 뉴클레아제 작용제의 활성 변이체 및 단편 (즉 조작된 뉴클레아제 작용제)이 또한 제공된다. 그러한 활성 변이체는 고유의 뉴클레아제 작용제에 대해 65% 이상, 70% 이상, 75% 이상, 80% 이상, 85% 이상, 90% 이상, 91% 이상, 92% 이상, 93% 이상, 94% 이상, 95% 이상, 96% 이상, 97% 이상, 98% 이상, 99% 이상 또는 그 이상의 서열 동일성을 포함할 수 있고, 여기서 활성 변이체는 원하는 인식 부위를 자르는 능력을 보유하고 있어서 닉 또는 이중-가닥-브레이크-유도 활성을 보유한다. 예를 들어, 본 명세서에 기재된 임의의 뉴클레아제 작용제는 고유의 엔도뉴클레아제 서열로부터 변형될 수 있고, 고유의 뉴클레아제 작용제에 의해 인식되지 않았던 인식 부위에서 닉 또는 이중-가닥 브레이크를 인식하고 유도하도록 설계될 수 있다. 따라서, 일부 실시 형태에서, 조작된 뉴클레아제는 상응하는 고유의 뉴클레아제 작용제 인식 부위와 상이한 인식 부위에서 닉 또는 이중-가닥 브레이크를 유도하기 위한 특이성을 갖는다. 닉 또는 이중-가닥-브레이크-유도 활성에 대한 검정법은 공지되어 있고 일반적으로 인식 부위를 함유하는 DNA 기질 상에서 엔도뉴클레아제의 전체 활성 및 특이성을 측정한다.
- [0104] 예를 들어, 도 3은 선택 카세트 상의 ZFN 결합 부위 및 절단 부위의 위치를 나타낸다. 부위는 다음과 같다: Neo-ZFN(1,2): 뉴클레아제 결합 부위/절단 부위 gggcgcccggttctttt/gtcaag/accgacctgtccggtG (서열 번호 9); Neo-ZFN(3,4): 뉴클레아제 결합 부위/절단 부위 ccggttcttttGTC/aagacc/gacctgtccggtgcc (서열 번호 10); Hyg-ZFN(1,2): 뉴클레아제 결합 부위/절단 부위 TGCGATCGTGCGGCCGA/tcttag/CCAGACGAGCGGGTTCGG (서열 번호 11); 및 Hyg-ZFN(3,4): 뉴클레아제 결합 부위/절단 부위 CGCTGCGGCCGATCT/tagcca/GACGAGCGGGTTCGG (서열 번호 12).
- [0105] 뉴클레아제 작용제는 당업계에 공지된 임의의 수단에 의해 세포 내로 도입될 수 있다. 뉴클레아제 작용제를 인코딩하는 폴리펩티드는 세포 내로 직접 도입될 수 있다. 대안적으로, 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드가 세포 내로 도입될 수 있다. 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드가 세포 내로 도입될 때, 뉴클레아제 작용제는 세포 내에서 일시적으로, 조건부로 또는 구성적으로 발현될 수 있다. 따라서, 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드는 발현 카세트에 함유될 수 있고 조건적 프로모터, 유도성 프로모터, 구성적 프로모터, 또는 조직-특이적 프로모터에 작동가능하게 연결될 수 있다. 그러한 관심 프로모터는 본 명세서의 어딘가 다른 곳에서 더욱 상세하게 논의된다. 대안적으로, 뉴클레아제 작용제는 뉴클레아제 작용제를 인코딩하는 mRNA로서 세포 내로 도입된다.
- [0106] 특정 실시 형태에서, 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드는 세포의 게놈에 안정적으로 통합되고 세포에서 활성인 프로모터에 작동가능하게 연결된다. 다른 실시 형태에서, 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드는 삽입 폴리뉴클레오티드를 포함하는 동일한 표적화 벡터 내에 존재하는 반면, 다른 경우에, 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드는 삽입 폴리뉴클레오티드를 포함하는 표적화 벡터와는 별개인 벡터 또는 플라스미드 내에 존재한다.
- [0107] 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드의 도입을 통해 뉴클레아제 작용제가 세포로 제공될 때, 뉴클레아제 작용제를 인코딩하는 그러한 폴리뉴클레오티드는, 뉴클레아제 작용제를 인코딩하는 자연적으로 발생한 폴리뉴클레오티드 서열과 비교하여, 관심 세포에서 더 높은 사용 빈도를 갖는 코돈을 치환하기 위해 변형될 수 있다. 예를 들어, 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드는, 자연적으로 발생한 폴리뉴클레오티드 서열과 비교하여, 관심의 박테리아 세포, 효모 세포, 인간 세포, 비-인간 세포, 포유류 세포, 설치류 세포, 마우스 세포, 래트 세포 또는 임의의 다른 숙주 세포를 포함하는, 주어진 관심 원핵 또는 진핵 세포에서 더 높은 사용 빈도를 갖는 코돈을 치환하기 위해 변형될 수 있다.
- [0108] B. 선택 마커
- [0109] 본 명세서에서 제공된 다양한 방법 및 조성물은 선택 마커와 조합하여 뉴클레아제 작용제 및 이의 상응하는 인식 부위를 사용한다. 본 명세서에서 논의된 바와 같이, 선택 마커를 인코딩하는 폴리뉴클레오티드에서 인식 부위의 위치는 표적 좌에서 통합 이벤트를 확인하기 위한 효율적인 방법을 허용한다. 더욱이, 다양한 방법이 본 명세서에 제공되며, 여기서는 뉴클레아제 인식 부위를 갖는 선택 마커를 대체하는 단계가 효율 및 효능을 개선하기 위해 사용되는데, 이를 통해 다수의 관심 폴리뉴클레오티드가 주어진 표적화된 좌 내에 통합된다.
- [0110] 다양한 선택 마커가 본 명세서에서 개시된 방법 및 조성물에 사용될 수 있다. 그러한 선택 마커는, 예를 들어, 항생제, 예컨대 G418, 하이그로마이신, 블라스토시딘, 네오마이신, 또는 퓨로마이신에 저항성을 부여할 수 있다. 그러한 선택 마커는 네오마이신 포스포트랜스퍼라제 ( $neo^r$ ), 하이그로마이신 B 포스포트랜스퍼라제 ( $hyg^r$ ), 퓨로마이신-N-아세틸트랜스퍼라제 ( $puro^r$ ), 및 블라스티시딘 S 데아미나제 ( $bsr^r$ )를 포함한다. 또 다른 실시 형태에서, 선택 마커는 유도성 프로모터에 작동가능하게 연결되고 선택 마커의 발현은 세포에 독성이다. 그러한 선택 마커의 비제한적인 예는 잔틴/구아닌 포스포리보실 트랜스퍼라제 ( $gpt$ ), 하이포잔틴-구아닌 포스포

리보실트랜스퍼라제 (HGPRT) 또는 단순 포진 바이러스 티미딘 키나제 (HSV-TK)를 포함한다.

[0111] 선택 마커를 인코딩하는 폴리뉴클레오티드는 세포에서 활성인 프로모터에 작동가능하게 연결된다. 그러한 발현 카세트 및 그의 다양한 조절 성분은 본 명세서의 어딘가 다른 곳에서 더욱 상세하게 논의된다.

[0112] C. 표적 좌

[0113] 표적 좌에서의 적어도 하나의 삽입 폴리뉴클레오티드의 통합을 가능하게 하는 다양한 방법 및 조성물이 제공된다. 용어 "표적 좌"는 삽입 폴리뉴클레오티드를 통합하기 원하는 DNA의 임의의 분절 또는 영역을 포함한다. 일 실시 형태에서, 표적 좌는 게놈 좌이다. 표적 좌는 세포에 고유할 수 있거나, 대안적으로 DNA의 이종성 또는 외인성 분절을 포함할 수 있다. 그러한 DNA의 이종성 또는 외인성 분절은 도입유전자, 발현 카세트, 선택 마커를 인코딩하는 폴리뉴클레오티드, 또는 DNA의 이종성 또는 외인성 영역 (즉, 게놈 DNA의 이종성 또는 외인성 영역)을 포함할 수 있다. 표적 좌는, 예를 들어, 인식 부위, 선택 마커, 미리 통합된 삽입 폴리뉴클레오티드, 뉴클레아제 작용제를 인코딩하는 폴리뉴클레오티드, 프로모터 등을 포함하는 임의의 표적화된 통합 시스템을 포함할 수 있다. 대안적으로, 표적 좌는 효모 인공 염색체 (yeast artificial chromosome) (YAC), 박테리아 인공 염색체 (BAC), 인간 인공 염색체, 또는 적당한 숙주 세포에 함유된 임의의 다른 조각된 게놈 영역 내에 위치할 수 있다. 따라서, 특정 실시 형태에서, 표적화된 좌는 원핵생물, 진핵생물, 효모, 박테리아, 비-인간 포유류, 비-인간 세포, 설치류, 인간, 래트, 마우스, 햄스터, 토끼, 돼지, 소, 사슴, 양, 염소, 닭, 고양이, 개, 페렛, 영장류 (예를 들어, 마모셋, 붉은털 원숭이), 사육된 포유류 또는 농업용 포유류 또는 임의의 다른 관심 유기체 또는 이들의 조합으로부터의 고유의, 이종성 또는 외인성 게놈 핵산 서열을 포함할 수 있다.

[0114] 표적 좌의 비제한적인 예는 B 세포에서 발현된 단백질을 인코딩하는 게놈 좌, 미성숙 B 세포에서 폴리펩티드를 발현하는 게놈 좌, 성숙 B 세포에서 폴리펩티드를 발현하는 게놈 좌, 면역글로불린 (Ig) 좌, 또는, 예를 들어 T 세포 수용체 알파 좌를 포함하는 T 세포 수용체 좌를 포함한다. 그러한 좌는 조류 (예를 들어, 닭), 비-인간 포유류, 설치류, 인간, 래트, 마우스, 햄스터, 토끼, 돼지, 소, 사슴, 양, 염소, 고양이, 개, 페렛, 영장류 (예를 들어, 마모셋, 붉은털 원숭이), 사육된 포유류 또는 농업용 포유류 또는 임의의 다른 관심 유기체 또는 이들의 조합으로부터 유래할 수 있다.

[0115] 다른 실시 형태에서, 표적화된 좌는, 뉴클레아제 작용제에 의해 유도된 Nick 또는 이중-가닥 브레이크의 부재 하에서, 종래의 방법을 사용하여 표적화가 가능하지 않거나 단지 부정확하게 또는 단지 유의하게 저효율로 표적화될 수 있다.

[0116] D. 표적화 벡터 및 삽입 폴리뉴클레오티드

[0117] 위에서 개략적으로 설명된 바와 같이, 본 명세서에 제공된 방법 및 조성물은 뉴클레아제 작용제, 및 상동성 재조합 이벤트와의 조합으로의 선택 카세트 내의 뉴클레아제 작용제에 대한 인식 부위의 전략적 위치화를 이용한다. 그러한 방법은 상동성 재조합과 조합하여 인식 부위에서 Nick 또는 이중-가닥 브레이크를 사용하여 표적 좌 내로 삽입 폴리뉴클레오티드의 통합을 표적화한다. "상동성 재조합"은 종래에 상동성 영역들 내의 크로스-오버 부위들에서의 2개의 DNA 분자들 사이에서 DNA 단편들의 교환을 포함하기 위해 사용된다.

[0118] i. 삽입 폴리뉴클레오티드

[0119] 용어 "삽입 폴리뉴클레오티드"는 표적 좌에 통합하기를 원하는 DNA의 분절을 포함한다. 일 실시 형태에서, 삽입 폴리뉴클레오티드는 하나 이상의 관심 폴리뉴클레오티드를 포함한다. 다른 실시 형태에서, 삽입 폴리뉴클레오티드는 하나 이상의 발현 카세트를 포함할 수 있다. 주어진 발현 카세트는 관심 폴리뉴클레오티드, 선택 마커 및/또는 리포터 유전자를 인코딩하는 폴리뉴클레오티드와, 이와 함께 발현에 영향을 미치는 다양한 조절 성분들을 포함할 수 있다. 삽입 폴리뉴클레오티드 내에 포함될 수 있는 관심 폴리뉴클레오티드, 선택 마커, 및 리포터 유전자 (예를 들어, eGFP)의 비제한적인 예는 본 명세서의 어딘가 다른 곳에서 상세하게 논의된다.

[0120] 특정 실시 형태에서, 삽입 폴리뉴클레오티드는 게놈 핵산을 포함할 수 있다. 일 실시 형태에서, 게놈 핵산은 마우스, 인간, 설치류, 비-인간, 래트, 햄스터, 토끼, 돼지, 소, 사슴, 양, 염소, 닭, 고양이, 개, 페렛, 영장류 (예를 들어, 마모셋, 붉은털 원숭이), 사육된 포유류 또는 농업용 포유류 또는 임의의 다른 관심 유기체 또는 이들의 조합으로부터 유래된다.

[0121] 다른 실시 형태에서, 삽입 폴리뉴클레오티드는 조건적 대립유전자를 포함한다. 일 실시 형태에서, 전체적으로 참고로 포함되는 미국 특허 제2011/0104799호에 기재된 바와 같이, 조건적 대립유전자는 다기능적 대립유전자이다. 특정 실시 형태에서, 조건적 대립유전자는 다음을 포함한다: (a) 표적 유전자의 전사에 관한 센스 배향



(sense orientation)의 구동 (actuating) 서열, 및 센스 또는 안티센스 배향의 약물 선택 카세트; (b) 안티센스 배향의 관심 뉴클레오타이드 서열 (nucleotide sequence of interest) (NSI) 및 조건적 역위 모듈 (conditional by inversion module) (COIN) (이는 엑손-스플리팅 (splitting) 인트론 및 가역 유전자 트랩 (genetrap)-유사 모듈을 사용함; 예를 들어, 전체적으로 참고로 포함되는 미국 특허 제2011/0104799호 참조); 및 (c) (i) 구동 서열 및 DSC가 결합되어 있고, (ii) NSI를 센스 배향으로 그리고 COIN을 안티센스 배향으로 합유하는 조건적 대립유전자를 형성하기 위해 제1 재조합효소에 대한 노출 시에 재조합하는 재조합가능 단위.

[0122] 삽입 폴리뉴클레오타이드는 약 5 kb 내지 약 200 kb, 약 5 kb 내지 약 10 kb, 약 10 kb 내지 약 20 kb, 약 20 kb 내지 약 30 kb, 약 30 kb 내지 약 40 kb, 약 40 kb 내지 약 50 kb, 약 60 kb 내지 약 70 kb, 약 80 kb 내지 약 90 kb, 약 90 kb 내지 약 100 kb, 약 100 kb 내지 약 110 kb, 약 120 kb 내지 약 130 kb, 약 130 kb 내지 약 140 kb, 약 140 kb 내지 약 150 kb, 약 150 kb 내지 약 160 kb, 약 160 kb 내지 약 170 kb, 약 170 kb 내지 약 180 kb, 약 180 kb 내지 약 190 kb, 또는 약 190 kb 내지 약 200 kb일 수 있다.

[0123] 특정 실시 형태에서, 삽입 폴리뉴클레오타이드는 부위-특이적 재조합 표적 서열로 플랭킹된 핵산을 포함한다. 삽입 폴리뉴클레오타이드 전체가 그러한 부위-특이적 재조합 표적 서열에 의해 플랭킹될 수 있지만, 삽입 폴리뉴클레오타이드 내의 임의의 관심 영역 또는 개별 관심 폴리뉴클레오타이드도 또한 그러한 부위에 의해 플랭킹될 수 있는 것으로 인식되어 있다. 용어 "재조합 부위"는, 부위-특이적 재조합효소에 의해 인식되고 재조합 이벤트에 대한 기질의 역할을 할 수 있는 뉴클레오타이드 서열을 포함한다. 용어 "부위-특이적 재조합효소"는 재조합 부위들 사이의 재조합을 촉진할 수 있는 효소들의 그룹을 포함하고, 여기서 2개의 재조합 부위는 단일 핵산 분자 내 또는 별개의 핵산 분자 상에서 물리적으로 분리된다. 부위-특이적 재조합효소의 예는 Cre, Flp, 및 Dre 재조합 효소를 포함하지만, 이로 제한되지는 않는다. 부위-특이적 재조합효소는, 재조합효소 폴리펩티드를 세포 내로 도입하거나 부위-특이적 재조합효소를 인코딩하는 폴리뉴클레오타이드를 숙주 세포 내로 도입하는 것을 포함하는 임의의 수단에 의해 세포 내로 도입될 수 있다. 부위-특이적 재조합효소를 인코딩하는 폴리뉴클레오타이드는 삽입 폴리뉴클레오타이드 내에 또는 별개의 폴리뉴클레오타이드 내에 위치할 수 있다. 부위-특이적 재조합효소는, 예를 들어, 유도성 프로모터, 세포에 대해 내인성인 프로모터, 세포에 대해 이중성인 프로모터, 세포-특이적 프로모터, 조직-특이적 프로모터, 또는 발생 단계-특이적 프로모터를 포함하는, 세포에서 활성인 프로모터에 작동가능하게 연결될 수 있다. 삽입 폴리뉴클레오타이드 또는 삽입 폴리뉴클레오타이드 내의 임의의 관심 폴리뉴클레오타이드를 플랭킹할 수 있는 부위-특이적 재조합 표적 서열은 loxP, lox511, lox2272, lox66, lox71, loxM2, lox5171, FRT, FRT11, FRT71, attP, att, FRT, rox, 및 이들의 조합을 포함할 수 있지만, 이로 제한되지는 않는다.

[0124] 다른 실시 형태에서, 부위-특이적 재조합 부위는 삽입 폴리뉴클레오타이드 내에 함유된 선택 마커 및/또는 리포터 유전자를 인코딩하는 폴리뉴클레오타이드를 플랭킹한다. 그러한 경우에는, 표적화된 좌에서의 삽입 폴리뉴클레오타이드의 통합 후에, 부위-특이적 재조합 부위들 사이의 서열은 제거될 수 있다.

[0125] 일 실시 형태에서, 삽입 폴리뉴클레오타이드는 선택 마커를 인코딩하는 폴리뉴클레오타이드를 포함한다. 그러한 선택 마커는 네오마이신 포스포트랜스퍼라제 (neo<sup>r</sup>), 하이그로마이신 B 포스포트랜스퍼라제 (hyg<sup>r</sup>), 퓨로마이신-N-아세틸트랜스퍼라제 (puro<sup>r</sup>), 블라스티시딘 S 데아미나제 (bsr<sup>r</sup>), 잔틴/구아닌 포스포리보실 트랜스퍼라제 (gpt), 또는 단순 포진 바이러스 티미딘 키나제 (HSV-k), 또는 이들의 조합을 포함하지만, 이로 제한되지는 않는다. 일 실시 형태에서, 선택 마커를 인코딩하는 폴리뉴클레오타이드는 세포에서 활성인 프로모터에 작동가능하게 연결된다. 관심 폴리뉴클레오타이드를 표적화된 좌 (즉, 게놈 좌) 내로 연속적으로 타일링할 때, 선택 마커는, 위에서 개략적으로 설명된 바와 같이, 뉴클레아제 작용제에 대한 인식 부위를 포함할 수 있다. 일 실시 형태에서, 선택 마커를 인코딩하는 폴리뉴클레오타이드는 부위-특이적 재조합 표적 서열로 플랭킹된다.

[0126] 삽입 폴리뉴클레오타이드는 프로모터에 작동가능하게 연결된 리포터 유전자를 추가로 포함할 수 있고, 여기서 리포터 유전자는 LacZ, 엠폴럼, 엠펜리, 티디토마토, 엠스트로베리, 제이-레드, 디에스레드, 엠오렌지, mKO, 엠시트린 (mCitrine), 비너스, 와이펫, 향상된 황색 형광 단백질 (enhanced yellow fluorescent protein) (EYFP), 에메랄드, 향상된 녹색 형광 단백질 (enhanced green fluorescent protein) (EGFP), 사이펫, 시안 형광 단백질 (CFP), 세룰리안, T-사파이어, 루시퍼라제, 알칼리성 포스파타제, 및 이들의 조합으로 이루어진 군으로부터 선택되는 리포터 단백질을 인코딩한다. 그러한 리포터 유전자는 세포에서 활성인 프로모터에 작동가능하게 연결될 수 있다. 그러한 프로모터는 유도성 프로모터, 리포터 유전자 또는 세포에 대해 내인성인 프로모터, 리포터 유전자 또는 세포에 대해 이중성인 프로모터, 세포-특이적 프로모터, 조직-특이적 프로모터 또는 발생 단계-특이적 프로모터일 수 있다.

[0127] ii. 표적화 벡터

[0128] 표적화 벡터는 표적화된 좌 내로 삽입 폴리뉴클레오타이드를 도입하기 위해 사용된다. 표적화 벡터는 삽입 폴리뉴클레오타이드를 포함하고, 삽입 폴리뉴클레오타이드를 플랭킹하는 상류 및 하류 상동성 아암을 추가로 포함한다. 삽입 폴리뉴클레오타이드를 플랭킹하는 상동성 아암들은 표적화된 좌 내의 영역들에 상응한다. 참고하기 쉽도록, 표적화된 좌 내의 상응하는 영역들은 본 명세서에서 "표적 부위"로 지칭된다. 따라서, 일 예에서, 표적화 벡터는 선택 마커를 인코딩하는 폴리뉴클레오타이드 내의 제1 인식 부위에 충분히 가깝게 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹된 제1 삽입 폴리뉴클레오타이드를 포함할 수 있다. 그렇기 때문에, 표적화 벡터는, 예를 들어, 세포의 게놈 내에서 상동성 아암 및 상응하는 표적 부위 사이에 발생하는 상동성 재조합 이벤트를 통해 표적화된 좌 내로의 삽입 폴리뉴클레오타이드의 통합을 돕는다.

[0129] 표적화 벡터의 상동성 아암은, 예를 들어, 50 내지 100 베이스, 100 내지 1000 베이스 또는 적어도 5 내지 10, 5 내지 15, 5 내지 20, 5 내지 25, 5 내지 30, 5 내지 35, 5 내지 40, 5 내지 45, 5 내지 50, 5 내지 55, 5 내지 60, 5 내지 65, 5 내지 70, 5 내지 75, 5 내지 80, 5 내지 85, 5 내지 90, 5 내지 95, 5 내지 100, 100 내지 200, 또는 200 내지 300 킬로베이스 길이 또는 그 이상을 포함하는, 상응하는 표적 부위와의 상동성 재조합 이벤트를 촉진하기에 충분한 임의의 길이를 가질 수 있다. 아래에서 더욱 상세하게 설명되는 바와 같이, 대형 표적화 벡터는 더 큰 길이의 표적화 아암을 사용할 수 있다.

[0130] 표적화 벡터의 상류 및 하류 상동성 아암에 상응하는 표적화된 좌 내의 표적 부위는 선택 마커를 인코딩하는 폴리뉴클레오타이드에 위치하는 "인식 부위에 충분히 가깝게" 위치한다. 표적화 벡터의 상류 및 하류 상동성 아암은 인식 부위에 "충분히 가깝게 위치하고", 여기서 거리는 인식 부위에서의 Nick 또는 이중-가닥 브레이크 상에서 표적 부위와 상동성 아암 사이에서의 상동성 재조합 이벤트의 발생을 촉진하기 위한 그러한 거리이다. 따라서, 특정 실시 형태에서, 표적화 벡터의 상류 및/또는 하류 상동성 아암에 상응하는 표적 부위는 주어진 인식 부위의 적어도 1 뉴클레오타이드 내에 있거나, 주어진 인식 부위의 적어도 10개 뉴클레오타이드 내지 약 14 kb 내에 있거나, 또는 주어진 인식 부위의 약 10개 뉴클레오타이드 내지 약 100개 뉴클레오타이드, 약 100개 뉴클레오타이드 내지 약 500개 뉴클레오타이드, 약 500개 뉴클레오타이드 내지 약 1000개 뉴클레오타이드, 약 1 kb 내지 약 5 kb, 약 5 kb 내지 약 10 kb, 또는 약 10 kb 내지 약 14 kb 내에 있다. 특정 실시 형태에서, 인식 부위는 표적 부위들 중 적어도 하나 또는 둘 다에 바로 인접해 있다.

[0131] 표적화 벡터의 상동성 아암들에 상응하는 표적 부위들과 선택 마커를 인코딩하는 폴리뉴클레오타이드 내의 인식 부위의 공간적 관계는 다양할 수 있다. 예를 들어, 표적 부위들은 인식 부위의 5'에 위치할 수 있거나, 표적 부위 둘 다는 인식 부위의 3'에 위치할 수 있거나, 표적 부위들은 인식 부위를 플랭킹할 수 있다.

[0132] 상동성 아암 및 표적 부위가 상동성 재조합 반응에 대한 기질로서 작용하기 위해 서로에 대해 충분한 수준의 서열 동일성을 공유할 때, 두 영역은 서로에게 "상응하거나" 또는 "상응하고 있다". "상동성"은 동일한 DNA 서열 또는 상응하는 서열에 대해 서열 동일성을 공유하는 DNA 서열을 의미한다. 주어진 표적 부위 및 표적화 벡터 상에서 확인되는 상응하는 상동성 아암 사이의 서열 동일성은 상동성 재조합이 발생하도록 허용하는 임의의 정도의 서열 동일성일 수 있다. 예를 들어, 표적화 벡터 (또는 이의 단편)의 상동성 아암과 표적 부위 (또는 이의 단편)에 의해 공유되는 서열 동일성의 양은 50% 이상, 55% 이상, 60% 이상, 65% 이상, 70% 이상, 75% 이상, 80% 이상, 81% 이상, 82% 이상, 83% 이상, 84% 이상, 85% 이상, 86% 이상, 87% 이상, 88% 이상, 89% 이상, 90% 이상, 91% 이상, 92% 이상, 93% 이상, 94% 이상, 95% 이상, 96% 이상, 97% 이상, 98% 이상, 99% 이상 또는 100% 서열 동일성일 수 있어서, 이들 서열은 상동성 재조합을 거치게 된다. 더욱이, 상동성 아암과 상응하는 표적 부위 사이의 상응하는 상동성 영역은 절단된 인식 부위에서 상동성 재조합을 촉진하기에 충분한 임의의 길이를 가질 수 있다. 예를 들어, 주어진 상동성 아암 및/또는 상응하는 표적 부위는 (예컨대, 본 명세서의 어딘가 다른 곳에서 기재된 LTVEC 벡터에서 기재된 바와 같이) 적어도 약 50 내지 100 베이스, 100 내지 1000 베이스, 또는 5 내지 10, 5 내지 15, 5 내지 20, 5 내지 25, 5 내지 30, 5 내지 35, 5 내지 40, 5 내지 45, 5 내지 50, 5 내지 55, 5 내지 60, 5 내지 65, 5 내지 70, 5 내지 75, 5 내지 80, 5 내지 85, 5 내지 90, 5 내지 95, 5 내지 100, 100 내지 200, 또는 200 내지 300 킬로베이스 길이 또는 그 이상의 상응하는 상동성 영역을 포함할 수 있어서, 상동성 아암이 세포의 게놈 내의 상응하는 표적 부위와의 상동성 재조합을 거치기에 충분한 상동성을 갖게 된다.

[0133] 참고하기 쉽도록, 상동성 아암은 상류 및 하류 상동성 아암을 포함한다. 이 용어는 표적화 벡터 내의 삽입 폴리뉴클레오타이드에 대한 상동성 아암의 상대적인 위치에 관한 것이다.

[0134] 그러므로, 표적화 벡터의 상동성 아암은 표적화된 좌를 갖는 표적 부위에 상응하도록 설계된다. 따라서, 상동

성 아암은 세포에 고유한 좌에 상응할 수 있거나, 대안적으로 이는 도입유전자, 발현 카세트, 또는 DNA의 이중성 또는 외인성 영역을 포함하지만, 이로 제한되지는 않는, 세포의 게놈 내로 통합된 DNA의 이중성 또는 외인성 분절의 영역에 상응할 수 있다. 대안적으로, 표적화 벡터의 상동성 아암은 효모 인공 염색체 (YAC), 박테리아 인공 염색체 (BAC), 인간 인공 염색체의 영역, 또는 적당한 숙주 세포에 함유된 임의의 다른 조작된 영역에 상응할 수 있다. 추가로 또한, 표적화 벡터의 상동성 아암은 BAC 라이브러리, 코스미드 라이브러리, 또는 P1 파지 라이브러리의 영역에 상응할 수 있거나 이로부터 유래할 수 있다. 따라서, 특정 실시 형태에서, 표적화 벡터의 상동성 아암은 원핵생물, 효모, 조류 (예를 들어, 닭), 비-인간 포유류, 설치류, 인간, 래트, 마우스, 햄스터, 토끼, 돼지, 소, 사슴, 양, 염소, 고양이, 개, 페렛, 영장류 (예를 들어, 마모셋, 붉은털 원숭이), 사육된 포유류 또는 농업용 포유류 또는 임의의 다른 관심 유기체에 대해 고유하거나, 이중성이거나 외인성인 좌에 상응한다. 다른 실시 형태에서, 상동성 아암은, 뉴클레아제 작용제에 의해 유도된 Nick 또는 이중-가닥 브레이크의 부재 하에서, 종래의 방법을 사용하여 표적화가 가능하지 않거나 단지 부정확하게 또는 단지 유의하게 저효율로 표적화될 수 있는 세포의 좌에 상응한다. 일 실시 형태에서, 상동성 아암은 합성 DNA로부터 유래된다.

[0135] 또 다른 실시 형태에서, 상류 및 하류 상동성 아암은 표적화된 게놈과 동일한 게놈에 상응한다. 일 실시 형태에서, 상동성 아암은 관련된 게놈으로부터 유래하고, 예를 들어, 표적화된 게놈은 제1 세포주 (strain)의 마우스 게놈이고, 표적화 아암은 제2 세포주의 마우스 게놈으로부터 유래하고, 여기서 제1 세포주 및 제2 세포주는 상이하다. 다른 실시 형태에서, 상동성 아암은 동일한 동물의 게놈으로부터 유래하거나 동일한 세포주의 게놈으로부터 유래하고, 예를 들어, 표적화된 게놈은 제1 세포주의 마우스 게놈이고, 표적화 아암은 동일한 마우스로부터 또는 동일한 세포주로부터의 마우스 게놈으로부터 유래한다.

[0136] 본 명세서의 어딘가 다른 곳에서 논의된 바와 같이, 표적화 벡터 (예컨대, 대형 표적화 벡터)는 또한 선택 카세트 또는 리포터 유전자를 포함할 수 있다. 선택 카세트는 선택 마커를 인코딩하는 핵산 서열을 포함할 수 있고, 여기서 핵산 서열은 프로모터에 작동가능하게 연결된다. 프로모터는 관심 원핵 세포에서 활성일 수 있고/있거나 관심 진핵 세포에서 활성일 수 있다. 그러한 프로모터는 유도성 프로모터, 리포터 유전자 또는 세포에 대해 내인성인 프로모터, 리포터 유전자 또는 세포에 대해 이중성인 프로모터, 세포-특이적 프로모터, 조직-특이적 프로모터 또는 발생 단계-특이적 프로모터일 수 있다. 일 실시 형태에서, 선택 마커는 네오마이신 포스포트랜스퍼라제 ( $neo^r$ ), 하이그로마이신 B 포스포트랜스퍼라제 ( $hyg^r$ ), 퓨로마이신-N-아세틸트랜스퍼라제 ( $puro^r$ ), 블라스티시딘 S 테아미나제 ( $bsr^r$ ), 잔틴/구아닌 포스포리보실 트랜스퍼라제 ( $gpt$ ), 및 단순 포진 바이러스 티미딘 키나제 (HSV-k), 및 이들의 조합으로부터 선택된다. 표적화 벡터의 선택 마커는 상류 및 하류 상동성 아암에 의해 플랭킹될 수 있거나 상동성 아암의 5' 또는 3'에서 발견될 수 있다.

[0137] 일 실시 형태에서, 표적화 벡터 (예컨대, 대형 표적화 벡터)는 프로모터에 작동가능하게 연결된 리포터 유전자를 포함하고, 여기서 리포터 유전자는 LacZ, 엠프렘, 엠헤리, 티디토마토, 엠스트로베리, 제이-레드, 디에스레드, 엠옐렌지, mKO, 엠시트린, 비너스, 와이펫, 향상된 황색 형광 단백질 (EYFP), 에메랄드, 향상된 녹색 형광 단백질 (EGFP), 사이펫, 시안 형광 단백질 (CFP), 세룰리안, T-사파이어, 루시퍼라제, 알칼리성 포스포타제, 및 이들의 조합으로 이루어진 군으로부터 선택되는 리포터 단백질을 인코딩한다. 그러한 리포터 유전자는 세포에서 활성인 프로모터에 작동가능하게 연결될 수 있다. 그러한 프로모터는 유도성 프로모터, 리포터 유전자 또는 세포에 대해 내인성인 프로모터, 리포터 유전자 또는 세포에 대해 이중성인 프로모터, 세포-특이적 프로모터, 조직-특이적 프로모터 또는 발생 단계-특이적 프로모터일 수 있다.

[0138] 일 실시 형태에서, 뉴클레아제 작용제와 (예를 들어, 대형 표적화 벡터를 포함하는) 표적화 벡터의 조합 사용은 표적화 벡터의 단독 사용과 비교하여 증가된 표적화 효율을 가져온다. 일 실시 형태에서, 표적화 벡터가 뉴클레아제 작용제와 함께 사용될 때, 표적화 벡터의 표적화 효율은 표적화 벡터가 단독으로 사용될 때와 비교하여 적어도 2배, 적어도 3배, 적어도 4배, 또는 적어도 10배 증가된다.

[0139] *iii. 대형 표적화 벡터*

[0140] 용어 "대형 표적화 벡터" 또는 "LTVEC"는 세포에서 상동성 재조합을 수행하도록 의도된 다른 접근에 의해 전형적으로 사용된 것들보다 큰 핵산 서열에 상응하고 이로부터 유래된 상동성 아암을 포함하고/하거나 세포에서 상동성 재조합을 수행하도록 의도된 다른 접근에 의해 전형적으로 사용되는 것들보다 큰 핵산 서열을 포함하는 삽입 뉴클레오티드를 포함하는 대형 표적화 벡터를 포함한다. 특정 실시 형태에서, LTVEC의 상동성 아암 및/또는 삽입 폴리뉴클레오티드는 진핵 세포의 게놈 서열을 포함한다. LTVEC의 크기는 너무 커서 종래의 검정법, 예를 들어, 서던 블롯팅 (southern blotting) 및 긴-범위 (예를 들면, 1 kb 내지 5 kb) PCR에 의한 표적화 이벤트의 스크리닝을 가능하게 할 수 없다. LTVEC의 예는 박테리아 인공 염색체 (BAC), 인간 인공 염색체 또는 효모 인

공 염색체 (YAC)로부터 유래된 벡터를 포함하지만, 이로 제한되지는 않는다. LTVEC 및 이를 제조하기 위한 방법의 비제한적인 예는, 예를 들어, 미국 특허 제 6,586,251호, 제6,596,541호, 제7,105,348호, 및 국제특허 공개 WO 2002/036789호 (PCT/US01/45375)에 기재되어 있고, 이들 각각은 본 명세서에 참고로 포함된다.

[0141] LTVEC는 임의의 길이를 가질 수 있고, 이는 약 20 kb 내지 약 300 kb, 약 20 kb 내지 약 30 kb, 약 30 kb 내지 약 40 kb, 약 40 kb 내지 약 50 kb, 약 50 kb 내지 약 75 kb, 약 75 kb 내지 약 100 kb, 약 100 kb to 125 kb, 약 125 kb 내지 약 150 kb, 약 150 kb 내지 약 175 kb, 약 175 kb 내지 약 200 kb, 약 200 kb 내지 약 225 kb, 약 225 kb 내지 약 250 kb, 약 250 kb 내지 약 275 kb 또는 약 275 kb 내지 약 300 kb를 포함하지만, 이로 제한되지는 않는다.

[0142] 일 실시 형태에서, LTVEC는 약 5 kb 내지 약 200 kb, 약 5 kb 내지 약 10 kb, 약 10 kb 내지 약 20 kb, 약 20 kb 내지 약 30 kb, 약 30 kb 내지 약 40 kb, 약 40 kb 내지 약 50 kb, 약 60 kb 내지 약 70 kb, 약 80 kb 내지 약 90 kb, 약 90 kb 내지 약 100 kb, 약 100 kb 내지 약 110 kb, 약 120 kb 내지 약 130 kb, 약 130 kb 내지 약 140 kb, 약 140 kb 내지 약 150 kb, 약 150 kb 내지 약 160 kb, 약 160 kb 내지 약 170 kb, 약 170 kb 내지 약 180 kb, 약 180 kb 내지 약 190 kb, 또는 약 190 kb 내지 약 200 kb 범위의 삽입 폴리뉴클레오티드를 포함한다.

[0143] 일 실시 형태에서, LTVEC의 상동성 아암은 BAC 라이브러리, 코스미드 라이브러리, 또는 P1 파지 라이브러리로부터 유래한다. 다른 실시 형태에서, 상동성 아암은 세포의 표적화된 좌 (즉, 게놈 좌)로부터 유래하고 일부 경우에 LTVEC가 표적이 되도록 설계된 표적 좌는 종래 방법을 사용하여 표적화될 수 없다. 또 다른 실시 형태에서, 상동성 아암은 합성 DNA로부터 유래된다. 일 실시 형태에서, LTVEC에서 상류 상동성 아암 및 하류 상동성 아암의 총 합계는 10 kb 이상이다. 일 실시 형태에서, 상류 상동성 아암은 약 1 kb 내지 약 100 kb 범위이다. 다른 실시 형태에서, 상류 상동성 아암은 약 5 kb 내지 약 100 kb 범위이다. 일 실시 형태에서, 하류 상동성 아암은 약 1 kb 내지 약 100 kb 범위이다. 일 실시 형태에서, 하류 상동성 아암은 약 5 kb 내지 약 100 kb 범위이다. 다른 실시 형태에서, 상류 및 하류 상동성 아암의 총 합계는 약 1 kb 내지 약 5 kb, 약 5 kb 내지 약 10 kb, 약 10 kb 내지 약 20 kb, 약 20 kb 내지 약 30 kb, 약 30 kb 내지 약 40 kb, 약 40 kb 내지 약 50 kb, 약 50 kb 내지 약 60 kb, 약 60 kb 내지 약 70 kb, 약 70 kb 내지 약 80 kb, 약 80 kb 내지 약 90 kb, 약 90 kb 내지 약 100 kb, 약 100 kb 내지 약 110 kb, 약 110 kb 내지 약 120 kb, 약 120 kb 내지 약 130 kb, 약 130 kb 내지 약 140 kb, 약 140 kb 내지 약 150 kb, 약 150 kb 내지 약 160 kb, 약 160 kb 내지 약 170 kb, 약 170 kb 내지 약 180 kb, 약 180 kb 내지 약 190 kb, 또는 약 190 kb 내지 약 200 kb이다.

[0144] 다른 실시 형태에서, LTVEC의 5' 및 3' 상동성 아암의 총 합계는 약 10 kb 내지 약 30 kb, 약 20 kb 내지 약 40 kb, 약 40 kb 내지 약 60 kb, 약 60 kb 내지 약 80 kb, 약 80 kb 내지 약 100 kb, 약 100 kb 내지 약 120 kb, 또는 약 120 kb to 150 kb이다. 다른 경우에, 5' 및 3' 상동성 아암의 총 합계는 약 16 Kb 내지 약 150 Kb이다.

[0145] 다른 실시 형태에서, LTVEC 및 삽입 폴리뉴클레오티드는 표적 좌에서 약 5 kb 내지 약 10 kb, 약 10 kb 내지 약 20 kb, 약 20 kb 내지 약 40 kb, 약 40 kb 내지 약 60 kb, 약 60 kb 내지 약 80 kb, 약 80 kb 내지 약 100 kb, 약 100 kb 내지 약 150 kb, 또는 약 150 kb 내지 약 200 kb, 약 200 kb 내지 약 300 kb, 약 300 kb 내지 약 400 kb, 약 400 kb 내지 약 500 kb, 약 500 kb 내지 약 1 Mb, 약 1 Mb 내지 약 1.5 Mb, 약 1.5 Mb 내지 약 2 Mb, 약 2 Mb 내지 약 2.5 Mb, 또는 약 2.5 Mb 내지 약 3 Mb의 결실이 가능하도록 설계된다.

[0146] 다른 경우에, LTVEC 및 삽입 폴리뉴클레오티드는 표적 좌 내로 약 5 kb 내지 약 10 kb, 약 10 kb 내지 약 20 kb, 약 20 kb 내지 약 40 kb, 약 40 kb 내지 약 60 kb, 약 60 kb 내지 약 80 kb, 약 80 kb 내지 약 100 kb, 약 100 kb 내지 약 150 kb, 약 150 kb 내지 약 200 kb, 약 200 kb 내지 약 250 kb, 약 250 kb 내지 약 300 kb, 약 300 kb 내지 약 350 kb, 또는 약 350 kb 내지 약 400 kb 범위의 외인성 핵산 서열의 삽입이 가능하도록 설계된다. 일 실시 형태에서, 삽입 폴리뉴클레오티드는 약 130 kb 또는 약 155 kb이다.

[0147] 일 실시 형태에서, LTVEC는 본 명세서의 어딘가 다른 곳에서 논의된 바와 같이 선택 카세트 또는 리포터 유전자를 포함한다.

[0148] *III. 표적 좌 내로 관심 폴리뉴클레오티드를 통합하는 방법*

[0149] *A. 상동성 재조합에 의한 인식 부위 부근에서의 삽입 폴리뉴클레오티드의 통합 방법*

[0150] 세포에서 표적 좌를 변형시키는 방법이 제공된다. 본 방법은 (a) 세포를 제공하는 단계로서, 세포는 세포에서 활성인 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 제1 폴리뉴클레오티드를 포함하고, 제



1 폴리뉴클레오티드는 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 추가로 포함하는, 상기 세포를 제공하는 단계; (b) (i) 제1 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도하는 제1 뉴클레아제 작용제, 및 (ii) 제1 인식 부위에 충분히 가깝게 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹된 제1 삽입 폴리뉴클레오티드를 포함하는 제1 표적화 벡터를 세포 내로 도입하는 단계; 및 (c) 표적 좌에 통합된 제1 삽입 폴리뉴클레오티드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함한다. 특정 실시 형태에서, 제1 선택 마커를 포함하는 제1 폴리뉴클레오티드는 제1 표적 부위 및 제2 표적 부위에 의해 플랭킹되고, 제1 표적 부위는 제1 표적화 벡터에서 제1 상동성 아암에 상응하고 제2 표적 부위는 제1 표적화 벡터에서 제2 상동성 아암에 상응한다.

[0151] 다양한 방법이 표적 좌에 통합된 삽입 폴리뉴클레오티드를 갖는 세포를 확인하는 데 사용될 수 있다. 일 실시 형태에서, 제1 인식 부위에서의 Nick 또는 이중-가닥 브레이크는 제1 선택 마커의 활성을 방해한다. 따라서, 일 실시 형태에서, 그러한 세포는, 뉴클레아제 작용제에 의해 잘린, 인식 부위를 갖는 폴리뉴클레오티드에 의해 인코딩되는 선택 마커의 활성을 갖지 않는 세포를 확인하는 조건 하에서 세포를 배양함으로써 확인된다. 그러한 선택 마커를 사용하고 이의 활성을 검정하기 위한 방법이 공지되어 있다. 표적 좌에 삽입 폴리뉴클레오티드를 갖는 세포를 확인하기 위한 추가적인 방법은 원하는 표적 부위에 통합된 삽입 폴리뉴클레오티드를 포함하여 갖는 적어도 하나의 세포를 확인하는 단계를 포함할 수 있다. 그러한 방법은 제1 및 제2 표적 부위에 통합된 제1 삽입 폴리뉴클레오티드를 게놈 내에 포함하는 적어도 하나의 세포를 확인하는 단계를 포함할 수 있다.

[0152] 추가적인 방법이 또한 표적 좌에 통합된 삽입 폴리뉴클레오티드를 갖는 세포를 확인하는 데 사용될 수 있다. 표적 좌에서의 삽입 폴리뉴클레오티드의 삽입은 "대립유전자의 변형"을 야기한다 용어 "대립유전자의 변형" 또는 "MOA"는 게놈 내의 유전자(들) 또는 염색체 좌 (좌들)의 하나의 대립유전자의 정확한 DNA 서열의 변형을 포함한다. "대립유전자의 변형 (MOA)"의 예는, 적게는 단일 뉴클레오티드의 결실, 치환, 또는 삽입, 또는 관심 유전자(들) 또는 염색체 좌 (좌들)을 스패닝 (spanning)하는 많은 킬로베이스의 결실뿐만 아니라, 이들 두 극치 사이의 임의의 그리고 모든 가능한 변형을 포함하지만, 이로 제한되지는 않는다.

[0153] 다양한 실시 형태에서, 표적화된 변형의 확인을 용이하게 하기 위해, 대량 처리 (high-throughput) 정량적 검정법, 즉, 대립유전자의 변형 (MOA) 검정법이 사용된다. 본 명세서에 기재된 MOA 검정법은 유전자 변형 후에 모 염색체 내의 변형된 대립유전자(들)의 대규모 스크리닝을 가능하게 한다. MOA 검정법은 정량적 PCR, 예를 들어, 실시간 PCR (qPCR)을 포함하지만, 이로 제한되지는 않는, 다양한 분석 기술을 통해 수행될 수 있다. 예를 들어, 실시간 PCR은 표적 좌를 인식하는 제1 프라이머 세트 및 비-표적화된 참조 좌를 인식하는 제2 프라이머 세트를 포함한다. 게다가, 프라이머 세트는 증폭된 서열을 인식하는 형광 프로브를 포함한다. 정량적 검정법은 또한 형광-매개 제자리 혼성화 (fluorescence-mediated in situ hybridization) (FISH), 비교 게놈 혼성화 (comparative genomic hybridization), 등온 DNA 증폭 (isothermal DNA amplification), 고정 프로브(들)에 대한 정량적 혼성화, 인베더 프로브즈 (Invader Probes)®, MMP 어세이즈 (MMP assays)®, 태크맨® 분자 비콘 (Molecular Beacon), 및 이클립스 (Eclipse)™ 프로브 기술을 포함하지만, 이로 제한되지 않는, 다양한 분석 기술을 통해 수행될 수 있다. (예를 들어, 전체적으로 본 명세서에 참고로 포함되는 미국 특허 출원 공개 제 2005/0144655호 참조).

[0154] 다양한 실시 형태에서, 선택 마커 내의 인식 부위에서의 Nick 또는 이중-가닥 브레이크의 존재는 표적화 벡터 (예컨대, LTVEC)와 표적화된 좌 사이의 재조합의 효율 및/또는 빈도를 증가시킨다. 일 실시 형태에서, 재조합은 상동성 재조합이다. 다양한 실시 형태에서, Nick 또는 이중 가닥 브레이크의 존재 하에서, 표적 좌에서의 표적화 벡터 (예컨대, LTVEC)의 표적화 효율은 Nick 또는 이중-가닥 브레이크의 부재 하에서보다 (예를 들어, 관심 좌에서 동일한 표적화 벡터 및 동일한 상동성 아암 및 상응하는 표적 부위를 사용하지만, Nick 또는 이중 가닥 브레이크를 만드는 첨가된 뉴클레아제 작용제의 부재 하에서보다) 적어도 약 2배 더 높고, 적어도 약 3배 더 높고, 적어도 약 4배, 적어도 약 10배 더 높다.

[0155] B. 표적화된 좌에 다수의 관심 폴리뉴클레오티드를 통합하는 방법

[0156] 본 명세서에 제공된 다양한 방법 및 조성물은 주어진 표적 좌 내에 다수의 관심 폴리뉴클레오티드의 표적화된 통합을 가능하게 한다. 본 방법은 본 명세서에 기재된, 선택 마커를 인코딩하는 폴리뉴클레오티드 내의 뉴클레아제 작용제 인식 부위의 전략적 위치화를 사용하는 표적화된 통합 시스템을 사용한다. 특정 실시 형태에서, 선택 마커 및 인식 부위는 각각의 삽입 폴리뉴클레오티드 내에서 대체된다. 그렇게 함으로써, 주어진 표적 좌 내의 순차적 삽입 폴리뉴클레오티드들의 타이핑이 향상된 효율 및 효능으로 일어난다.

[0157] 일 실시 형태에서, 세포에서 표적 좌를 변형시키는 방법은 (a) 세포를 제공하는 단계로서, 세포는 세포에서 활

성인 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 제1 폴리뉴클레오타이드를 포함하는 좌를 포함하고, 제1 폴리뉴클레오타이드는 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 추가로 포함하는, 상기 세포를 제공하는 단계; (b) 제1 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도하는 제1 뉴클레아제 작용제를 세포 내로 도입하고; 제1 인식 부위에 충분히 가깝게 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹된 제1 삽입 폴리뉴클레오타이드를 포함하는 제1 표적화 벡터를 세포 내로 도입하는 단계로서, 제1 삽입 폴리뉴클레오타이드는 (1) 제1 관심 폴리뉴클레오타이드; 및 (2) 세포에서 활성인 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 제2 폴리뉴클레오타이드로서, 제2 폴리뉴클레오타이드는 제2 뉴클레아제 작용제에 대한 제2 인식 부위를 포함하는, 상기 제2 폴리뉴클레오타이드를 추가로 포함하는, 상기 세포 내로 도입하는 단계; 및 (c) 표적 좌에 통합된 제1 삽입 폴리뉴클레오타이드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함한다.

[0158]

다른 실시 형태에서, 추가적인 관심 폴리뉴클레오타이드가 표적 좌에 통합될 수 있다. 세포에서 표적 좌를 변형시키는 그러한 방법은 (a) 세포를 제공하는 단계로서, 세포는 세포에서 활성인 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 제1 폴리뉴클레오타이드를 포함하는 좌를 포함하고, 제1 폴리뉴클레오타이드는 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 추가로 포함하는, 상기 세포를 제공하는 단계; (b) 제1 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도하는 제1 뉴클레아제 작용제를 세포 내로 도입하고; 제1 인식 부위에 충분히 가깝게 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹된 제1 삽입 폴리뉴클레오타이드를 포함하는 제1 표적화 벡터를 세포 내로 도입하는 단계로서, 제1 삽입 폴리뉴클레오타이드는 (1) 제1 관심 폴리뉴클레오타이드; 및 (2) 세포에서 활성인 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 제2 폴리뉴클레오타이드로서, 제2 폴리뉴클레오타이드는 제2 뉴클레아제 작용제에 대한 제2 인식 부위를 포함하는, 상기 제2 폴리뉴클레오타이드를 추가로 포함하는, 상기 세포 내로 도입하는 단계; (c) 표적 좌에 통합된 제1 삽입 폴리뉴클레오타이드를 포함하는 적어도 하나의 세포를 확인하는 단계; (d) 표적 좌에 통합된 제1 삽입 폴리뉴클레오타이드를 게놈 내에 포함하는 세포 내로 (i) 제2 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도하는 제2 뉴클레아제 작용제; 및 (ii) 제3 및 제4 상동성 아암에 의해 플랭킹된 제2 삽입 폴리뉴클레오타이드를 포함하는 제2 표적화 벡터를 도입하는 단계; 및 (b) 표적 좌에 통합된 제2 삽입 폴리뉴클레오타이드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함한다. 특정 실시 형태에서, 제2 인식 마커에서의 Nick 또는 이중 가닥 브레이크는 제2 선택 마커의 활성을 방해한다. 다른 실시 형태에서, 표적 좌에 통합된 제2 삽입 폴리뉴클레오타이드를 포함하는 적어도 하나의 세포를 확인하는 단계는 제2 선택 마커의 활성을 갖지 않는 세포를 확인하는 조건 하에서 세포를 배양하는 단계를 포함한다. 또 다른 실시 형태에서, 제2 선택 마커를 포함하는 제2 폴리뉴클레오타이드는 제3 표적 부위 및 제4 표적 부위에 의해 플랭킹되고, 제3 표적 부위는 제2 표적화 벡터에서 제3 상동성 아암에 상응하고 제4 표적 부위는 제2 표적화 벡터에서 제4 상동성 아암에 상응한다. 또 다른 실시 형태에서, 표적 좌에 통합된 제2 삽입 폴리뉴클레오타이드를 포함하는 적어도 하나의 세포를 확인하는 단계는 제3 및 제4 표적 부위에 통합된 제2 삽입 폴리뉴클레오타이드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함한다.

[0159]

세포에서 표적 좌를 변형시키기 위한 추가적인 방법은 (a) 세포를 제공하는 단계로서, 세포는 세포에서 활성인 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 제1 폴리뉴클레오타이드를 포함하는 표적 좌를 포함하고, 제1 폴리뉴클레오타이드는 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 추가로 포함하는, 상기 세포를 제공하는 단계; (b) (i) 제1 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도하는 제1 뉴클레아제 작용제; 및 (ii) 제1 인식 부위에 충분히 가깝게 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹된 제1 삽입 폴리뉴클레오타이드를 포함하는 제1 표적화 벡터를 세포 내로 도입하는 단계로서, 제1 삽입 폴리뉴클레오타이드는 (1) 제1 관심 폴리뉴클레오타이드; 및 (2) 세포에서 활성인 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 제2 폴리뉴클레오타이드로서, 제2 폴리뉴클레오타이드는 제2 뉴클레아제 작용제에 대한 제2 인식 부위를 포함하고, 제2 선택 마커를 포함하는 제2 폴리뉴클레오타이드는 제3 표적 부위 및 제4 표적 부위에 의해 플랭킹되고, 제3 표적 부위는 제2 표적화 벡터에서 제3 상동성 아암에 상응하고 제4 표적 부위는 제2 표적화 벡터에서 제4 상동성 아암에 상응하는, 상기 제2 폴리뉴클레오타이드를 추가로 포함하는, 상기 세포 내로 도입하는 단계; (c) 표적 좌에 통합된 제1 삽입 폴리뉴클레오타이드를 포함하는 적어도 하나의 세포를 확인하는 단계; (d) 표적 좌에 통합된 제1 삽입 폴리뉴클레오타이드를 포함하는 세포 내로 (i) 제2 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도하는 제2 뉴클레아제 작용제; 및 (ii) 제3 및 제4 상동성 아암에 의해 플랭킹된 제2 삽입 폴리뉴클레오타이드를 포함하는 제2 표적화 벡터를 도입하는 단계로서, 제2 삽입 폴리뉴클레오타이드는 (1) 제2 관심 폴리뉴클레오타이드; 및 (2) 세포에서 활성인 제3 프로모터에 작동가능하게 연결된 제3 선택 마커를 인코딩하는 제3 폴리뉴클레오타이드로서, 제3 폴리뉴클레오타이드는 제3 뉴클레아제 작용제에 대한 제3 인식 부위를

포함하는, 상기 제3 폴리뉴클레오타이드를 포함하는, 상기 제2 표적화 벡터를 도입하는 단계; 및 (b) 표적 좌에 통합된 제2 삽입 폴리뉴클레오타이드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함한다. 특정 실시 형태에서, 제2 인식 마커에서의 Nick 또는 이중 가닥 브레이크는 제2 선택 마커의 활성을 방해한다. 다른 실시 형태에서, 표적 좌에 통합된 제2 삽입 폴리뉴클레오타이드를 게놈 내에 포함하는 적어도 하나의 세포를 확인하는 단계는 제2 선택 마커의 활성을 갖지 않는 세포를 확인하는 조건 하에서 세포를 배양하는 단계를 포함한다. 다른 실시 형태에서, 표적 좌에 통합된 제2 삽입 폴리뉴클레오타이드를 게놈 내에 포함하는 적어도 하나의 세포를 확인하는 단계는 제3 및 제4 표적 부위에 통합된 제2 삽입 폴리뉴클레오타이드를 게놈 내에 포함하는 적어도 하나의 세포를 확인하는 단계를 포함한다.

[0160] 위에서 제시된 다양한 방법은 순차적으로 반복되어 주어진 표적화된 좌 내로의 임의의 수의 삽입 폴리뉴클레오타이드의 표적화된 통합을 가능하게 할 수 있다. 따라서, 다양한 방법은 표적 좌 내로 1개 이상, 2개 이상, 3개 이상, 4개 이상, 5개 이상, 6개 이상, 7개 이상, 8개 이상, 9개 이상, 10개 이상, 11개 이상, 12개 이상, 13개 이상, 14개 이상, 15개 이상, 16개 이상, 17개 이상, 18개 이상, 19개 이상, 20개 이상, 또는 그 이상의 삽입 폴리뉴클레오타이드의 삽입을 제공한다. 특정 실시 형태에서, 그러한 순차적인 타일링 방법은 표적화된 좌 (즉, 게놈 좌) 내로의 포유류 세포 (즉, 인간, 비-인간, 설치류, 마우스, 원숭이, 래트, 햄스터, 사육된 포유류 또는 농업용 동물)로부터의 큰 게놈 영역의 재구성을 가능하게 한다. 그러한 경우에, 코딩 및 비-코딩 영역 둘 다를 포함하는 게놈 영역의 전달 및 재구성은 주어진 영역의 복잡성이 코딩 영역, 비-코딩 영역 및 고유의 게놈 영역 내에서 발견되는 카피 수 변화의 적어도 일부를 유지함으로써 보존될 수 있게 한다. 따라서, 다양한 방법은, 예를 들어, 임의의 관심 포유류 세포 또는 동물 내에 "이종성" 또는 "외인성" 게놈 영역을 생성하는 방법을 제공한다. 하나의 비제한적인 예에서, 비-인간 동물 내에 "인간화된" 게놈 영역이 생성된다.

[0161] 주어진 표적 좌 내의 다수의 삽입 폴리뉴클레오타이드의 통합을 수행할 때, 선택 마커를 인코딩하고 뉴클레아제 작용제 인식 부위를 포함하는 폴리뉴클레오타이드는 통합 라운드들 사이에서 대체될 수 있다. 예를 들어, 특정 방법에서, 제1 뉴클레아제 작용제는 제2 뉴클레아제 작용제와 상이하고/상이하거나 제1 선택 마커는 제2 선택 마커와 상이하다. 다른 예에서, 표적화된 좌 내로 3개의 삽입 폴리뉴클레오타이드를 삽입할 때, 제1 및 제3 선택 마커는 서로 동일할 수 있고, 특정 실시 형태에서는 동일한 인식 부위를 추가로 포함할 수 있고, 제2 선택 마커는 제1 및 제3 선택 마커와 상이할 수 있고 상이한 인식 부위를 함유할 수 있다. 그러한 방식으로 선택 마커 및 인식 부위의 선택은 생성되어야 하는 뉴클레아제 작용제의 수를 최소화하여, 통합 이벤트의 효율 및 효능을 개선한다.

#### [0162] C. CRISPR/Cas 시스템을 사용하여 하나 이상의 표적 좌를 변형시키기 위한 방법

[0163] 본 명세서의 어딘가 다른 곳에서 기재된 바와 같은 CRISPR/Cas 시스템을 사용하여 세포에서 하나 이상의 관심 표적 좌를 변형시키기 위한 방법 및 조성물이 제공된다. CRISPR/Cas 시스템에 대해, 용어 "표적 부위" 또는 "표적 서열"은 상호교환적으로 사용될 수 있고, 결합에 충분한 조건이 존재하지만 한다면, 가이드 RNA (gRNA)의 DNA-표적화 분절이 결합할 표적 DNA에 존재하는 핵산 서열을 포함할 수 있다. 예를 들어, 표적 DNA 내의 표적 부위 (또는 표적 서열)은 Cas 뉴클레아제 또는 gRNA에 의해 표적화된다 (즉, 이에 의해 결합되거나, 이와 혼성화되거나, 이에 상보적이다). 적합한 DNA/RNA 결합 조건은 세포에 정상적으로 존재하는 생리적 조건을 포함한다. 다른 적합한 DNA/RNA 결합 조건 (예를 들어, 무세포 시스템에서의 조건)은 당업계에 공지되어 있다 (예를 들어, 문헌[Molecular Cloning: A Laboratory Manual, 3rd Ed. (Sambrook *et al.*, Harbor Laboratory Press 2001)] 참조). Cas 단백질 또는 gRNA에 상보적이고 이와 혼성화되는 표적 DNA의 가닥은 "상보적 가닥"으로 지칭되고 "상보적 가닥"에 상보적인 (따라서 Cas 단백질 또는 gRNA에 상보적이지 않은) 표적 DNA의 가닥은 "비상보적 가닥" 또는 "주형 가닥"으로 지칭된다.

[0164] Cas 단백질은 표적 서열 내의 또는 표적 서열 외부의 부위에서 핵산을 절단할 수 있다. "절단 부위"는 Cas 단백질이 단일-가닥 브레이크 또는 이중-가닥 브레이크를 생성하는 핵산의 위치를 포함한다. 각 가닥 상의 절단 부위에서 단일-가닥 브레이크를 생성하는 2개의 Cas9 단백질을 사용하여 점착성 말단이 또한 생성될 수 있다. Cas9에 의한 표적 DNA의 부위-특이적 절단은 표적 DNA에서 (i) 가이드 RNA와 표적 DNA 사이의 염기쌍 형성 상보성 및 (ii) 프로토스페이서 인접 모티프 (PAM)로 지칭되는, 짧은 모티프 둘 다에 의해 결정되는 위치에서 일어날 수 있다. 예를 들어, Cas9의 절단 부위는 PAM 서열의 상류의 약 1 내지 약 10개 또는 약 2 내지 약 5개의 염기쌍 (예를 들어, 3개의 염기쌍)일 수 있다. 일부 실시 형태에서 (예를 들어, S. 피오케네스로부터의 Cas9 또는 근연 Cas9가 사용될 때), 비-상보적 가닥의 PAM 서열은 5'-XGG-3'일 수 있고, 여기서 X는 임의의 DNA 뉴클레오타이드이고 X는 표적 DNA의 비-상보적 가닥의 표적 서열의 3'에 바로 있다. 그렇기 때문에, 상보적 가닥의 PAM 서열은 5'-CCY-3'일 것이고, 여기서 Y는 임의의 DNA 뉴클레오타이드이고 Y는 표적 DNA의 상보적 가닥의 표적



서열의 5'에 바로 있다. 일부 그러한 실시 형태에서, X 및 Y는 상보적일 수 있고 X-Y 염기쌍은 임의의 염기쌍 (예를 들어, X=C 및 Y=G; X=G 및 Y=C; X=A 및 Y=T, X=T 및 Y=A)일 수 있다.

[0165] 따라서, 일부 실시 형태에서, 세포에서 관심 표적 좌를 변형시키는 방법은 (a) 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 핵산을 포함하는 제1 표적 좌를 포함하는 세포를 제공하는 단계; (b) (i) Cas 단백질 및 제1 가이드 RNA (gRNA)를 인코딩하는 하나 이상의 발현 작제물로서, 이들 각각은 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제1 핵산에서의 제1 gRNA 표적 부위에 닉 또는 이중-가닥 브레이크를 유도하여 제1 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 발현 작제물, 및 (ii) 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 제2 핵산을 포함하는 제1 삽입 핵산을 포함하는 제1 표적화 벡터로서, 제1 삽입 핵산은 제1 표적 좌에 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹되는, 상기 제1 표적화 벡터를 세포 내로 도입하는 단계; 및 (c) 제1 표적 좌에 제1 삽입 핵산을 포함하는 변형된 세포를 확인하는 단계로서, 변형된 세포는 제2 선택 마커의 활성은 갖지만 제1 선택 마커의 활성은 갖지 않는, 상기 변형된 세포를 확인하는 단계를 포함하며, 제1 및 제2 선택 마커는 상이하다. 일 실시 형태에서, 제1 gRNA는 제1 삽입 핵산에 혼성화되지 않는다. 일 실시 형태에서, 관심 표적 좌는 세포의 게놈에 위치한다. 다른 실시 형태에서, 관심 표적 좌는 세포 내의 벡터에 위치한다. 일 실시 형태에서, 확인 단계 (c)는 제2 선택 마커의 활성은 갖지만 제1 선택 마커의 활성은 갖지 않는 변형된 세포의 확인을 가능하게 하는 조건 하에서 세포를 배양하는 단계를 포함한다.

[0166] 일 실시 형태에서, 본 방법은 (d) 제1 표적 좌에 제1 삽입 핵산을 포함하는 변형된 세포 내로 (i) Cas 단백질 및 제2 gRNA를 인코딩하는 하나 이상의 핵산으로서, 이들 각각은 변형된 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제2 핵산을 포함하는 제1 삽입 핵산에서 제2 gRNA 표적 부위에 닉 또는 이중-가닥 브레이크를 유도하여 제2 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 핵산, 및 (ii) 제3 프로모터에 작동가능하게 연결된 제3 선택 마커를 인코딩하는 제3 핵산을 포함하는 제2 삽입 핵산을 포함하는 제2 표적화 벡터로서, 제2 삽입 핵산은 제2 표적 좌에 위치한 제3 및 제4 표적 부위에 상응하는 제3 및 제4 상동성 아암에 의해 플랭킹되는, 상기 제2 표적화 벡터를 도입하는 단계; 및 (e) 제2 표적 좌에 제2 삽입 핵산을 포함하는 제2 변형된 세포를 확인하는 단계로서, 제2 변형된 세포는 제3 선택 마커의 활성은 갖지만 제2 선택 마커의 활성은 갖지 않는, 상기 제2 변형된 세포를 확인하는 단계를 추가로 포함하며, 제2 및 제3 선택 마커는 상이하다. 일 실시 형태에서, 제1 및 제2 표적 좌는 서로 바로 인접하게 위치한다. 다른 실시 형태에서, 제1 또는 제2 표적 좌는 제1 또는 제2 gRNA 표적 부위로부터 약 10개 뉴클레오티드 내지 약 14 kb에 위치한다. 일 실시 형태에서, 제2 gRNA는 제2 삽입 핵산에 혼성화되지 않는다. 일 실시 형태에서, 확인 단계 (e)는 제3 선택 마커의 활성은 갖지만 제2 선택 마커의 활성은 갖지 않는 제2 변형된 세포의 확인을 가능하게 하는 조건 하에서 변형된 세포를 배양하는 단계를 포함한다.

[0167] 일 실시 형태에서, 본 방법은 (f) 제2 표적 좌에 제2 삽입 핵산을 포함하는 제2 변형된 세포 내로 (i) Cas 단백질 및 제3 gRNA를 인코딩하는 하나 이상의 발현 작제물로서, 이들 각각은 제2 변형된 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제3 핵산을 포함하는 제2 삽입 핵산에서 제3 gRNA 표적 부위에 닉 또는 이중-가닥 브레이크를 유도하여 제3 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 발현 작제물, 및 (ii) 제4 프로모터에 작동가능하게 연결된 제4 선택 마커를 인코딩하는 제4 핵산을 포함하는 제3 삽입 핵산을 포함하는 제3 표적화 벡터로서, 제3 삽입 핵산은 제3 표적 좌에 위치한 제5 및 제6 표적 부위에 상응하는 제5 및 제6 상동성 아암에 의해 플랭킹되는, 상기 제3 표적화 벡터를 도입하는 단계; 및 (g) 제3 표적 좌에 제3 삽입 핵산을 포함하는 제3 변형된 세포를 확인하는 단계로서, 제3 변형된 세포는 제4 선택 마커의 활성은 갖지만 제3 선택 마커의 활성은 갖지 않는, 상기 제3 변형된 세포를 확인하는 단계를 추가로 포함하며, 제3 및 제4 선택 마커는 상이하다. 일 실시 형태에서, 제2 및 제3 표적 좌는 서로 바로 인접하게 위치한다. 다른 실시 형태에서, 제2 또는 제3 표적 좌는 제1 또는 제2 gRNA 표적 부위로부터 약 10개 뉴클레오티드 내지 약 14 kb에 위치한다.

[0168] 일 실시 형태에서, 제1, 제2, 제3, 또는 제4 마커는 항생제에 대한 저항성을 부여한다. 일 실시 형태에서, 항생제는 G418, 하이그로마이신, 블라스토시딘, 네오마이신, 또는 퓨로마이신을 포함한다. 일 실시 형태에서, 제1, 제2, 제3, 또는 제4 선택 마커는 하이포잔틴-구아닌 포스포리보실트랜스퍼라제 (HGPRT) 또는 단순 포진 바이러스 (HSV-TK)의 티미딘 키나제를 포함한다. 일 실시 형태에서, 제1, 제2, 또는 제3 gRNA는 (i) 제1, 제2 또는 제3 gRNA 표적 부위에 혼성화되는 뉴클레오티드 서열 및 (ii) 트랜스-활성화 CRISPR RNA (tracrRNA)를 포함한다. 일 실시 형태에서, 제1, 제2, 또는 제3 표적 좌는 제1, 제2 또는 제3 gRNA 표적 부위에 아주 가깝게 위치해서 gRNA 표적 부위에서의 닉 또는 이중-가닥 브레이크가 표적 좌에서 표적화 벡터의 상동성 제조합을 측

진하도록 한다. 일 실시 형태에서, Cas 단백질은 Cas9이다. 일 실시 형태에서, 제1, 제2, 또는 제3 gRNA 표적 부위는 프로토스페이서 인접 모티프 (PAM) 서열에 의해 바로 플랭킹된다.

[0169] 특정 실시 형태에서, gRNA는 제2 선택 마커 (예를 들어, Neo<sup>r</sup>)를 인코딩하는 제1 삽입 핵산의 삽입을 위한 제1 항생제 선택 마커 (예를 들어, Hyg<sup>r</sup>)를 표적으로 하도록 설계되어, 제1 삽입 핵산의 삽입은 제1 항생제 선택 마커의 활성을 방해한다. 제2 gRNA 발현 플라스미드는 제1 선택 마커를 인코딩하는 제2 삽입 핵산의 삽입을 위한 제2 선택 마커를 표적으로 하는 gRNA를 발현하도록 설계될 수 있어서, 제2 삽입 핵산의 삽입은 제2 항생제 선택 마커의 활성을 방해한다. 이러한 방식으로, gRNA는 핵산 삽입물을 대체하는 데 사용될 수 있는 2개의 항생제 선택 마커 각각을 표적으로 하도록 설계되지만 하면 된다. Neo 저항성 선택 마커에 특이적인 gRNA를 인코딩하는 예시적인 핵산은 서열 번호 13, 14, 15, 및 16에서 확인될 수 있다. Hyg 저항성 선택 마커에 특이적인 gRNA를 인코딩하는 예시적인 핵산은 서열 번호 17, 18, 19, 및 20에서 확인될 수 있다.

[0170] 일 실시 형태에서, 세포는 원핵 세포이다. 다른 실시 형태에서, 세포는 진핵 세포이다. 일 실시 형태에서, 진핵 세포는 포유류 세포 또는 비-인간 포유류 세포이다. 일 실시 형태에서, 포유류 세포는 섬유아세포이다. 일 실시 형태에서, 포유류 세포는 인간 섬유아세포이다. 일 실시 형태에서, 포유류 세포는 인간 성체 줄기 세포이다. 일 실시 형태에서, 포유류 세포는 발생적으로 제한된 선조 세포이다. 일 실시 형태에서, 포유류 세포는 발생적으로 제한된 인간 선조 세포이다.

[0171] 일 실시 형태에서, 포유류 세포는 비-인간 포유류 세포이다. 일 실시 형태에서, 포유류 세포는 설치류 유래이다. 일 실시 형태에서, 설치류는 래트, 마우스, 또는 햄스터이다. 일 실시 형태에서, 진핵 세포는 다능성 세포이다. 일 실시 형태에서, 다능성 세포는 조혈 줄기 세포 또는 신경 줄기 세포이다. 일 실시 형태에서, 다능성 세포는 인간 유도 다능성 줄기 (iPS) 세포이다. 일 실시 형태에서, 다능성 세포는 비-인간 ES 세포, 인간 ES 세포, 설치류 배아 줄기 (ES) 세포, 마우스 배아 줄기 (ES) 세포 또는 래트 배아 줄기 (ES) 세포이다.

[0172] 일 실시 형태에서, 제1, 제2, 또는 제3 gRNA 표적 부위는 제1, 제2, 또는 제3 선택 마커를 인코딩하는 제1, 제2, 또는 제3 핵산에서의 인트론, 엑손, 프로모터, 또는 프로모터 조절 영역에 위치한다. 일 실시 형태에서, 제1, 제2, 또는 제3 표적화 벡터는 약 10 kb 이상이다. 일 실시 형태에서, 제1, 제2, 또는 제3 삽입 핵산은 약 5 kb 내지 약 300 kb의 범위이다.

[0173] 일 실시 형태에서, 제1, 제2, 또는 제3 삽입 핵산은 인간 T 세포 수용체 알파 좌의 게놈 영역을 포함한다. 일 실시 형태에서, 게놈 영역은 인간 T 세포 수용체 알파 좌의 적어도 하나의 가변 영역 유전자 분절 및/또는 결합 영역 유전자 분절을 포함한다.

[0174] 일 실시 형태에서, 제1 및 제3 선택 마커는 동일하다. 일 실시 형태에서, 제1 및 제3 선택 마커는 동일하고 제2 및 제4 선택 마커는 동일하다. 일 실시 형태에서, 제1 및 제3 gRNA는 동일하다.

[0175] 일부 실시 형태에서, 세포에서 관심 표적 좌를 변형시키는 방법은 (a) 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 핵산을 포함하는 제1 표적 좌를 포함하는 세포를 제공하는 단계; (b) (i) Cas 단백질 및 제1 gRNA를 인코딩하는 하나 이상의 발현 작제물로서, 이들 각각은 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제1 핵산에서의 제1 gRNA 표적 부위에 Nick 또는 이중-가닥 브레이크를 유도하여 제1 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 발현 작제물, 및 (ii) 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 제2 핵산을 포함하는 제1 삽입 핵산을 포함하는 제1 표적화 벡터로서, 제1 삽입 핵산은 제1 표적 좌에 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹되는, 상기 제1 표적화 벡터를 세포 내로 도입하는 단계; (c) 제1 표적 좌에 제1 삽입 핵산을 포함하는 변형된 세포를 확인하는 단계로서, 변형된 세포는 제2 선택 마커의 활성은 갖지만 제1 선택 마커의 활성은 갖지 않는, 상기 변형된 세포를 확인하는 단계 - 여기서, 제1 및 제2 선택 마커는 상이함 -; (d) 제1 표적 좌에 제1 삽입 핵산을 포함하는 변형된 세포 내로 (i) Cas 단백질 및 제2 gRNA를 인코딩하는 하나 이상의 핵산으로서, 이들 각각은 변형된 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제2 핵산을 포함하는 제1 삽입 핵산에서 제2 gRNA 표적 부위에 Nick 또는 이중-가닥 브레이크를 유도하여 제2 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 핵산; 및 (ii) 제3 프로모터에 작동가능하게 연결된 제3 선택 마커를 인코딩하는 제3 핵산을 포함하는 제2 삽입 핵산을 포함하는 제2 표적화 벡터로서, 제2 삽입 핵산은 제2 표적 좌에 위치한 제3 및 제4 표적 부위에 상응하는 제3 및 제4 상동성 아암에 의해 플랭킹되는, 상기 제2 표적화 벡터를 도입하는 단계; 및 (e) 제2 표적 좌에 제2 삽입 핵산을 포함하는 제2 변형된 세포를 확인하는 단계로서, 제2 변형된 세포는 제3 선택 마커의 활성은 갖지만 제2 선택 마커의 활성은 갖지 않는, 상기 제2 변형된 세포를 확인하는 단계를

포함하며, 제1 및 제3 선택 마커는 동일하고, 제2 및 제3 선택 마커는 상이하다.

[0176] 다른 실시 형태에서, 세포에서 관심 표적 좌를 변형시키는 방법은 (a) 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 핵산을 포함하는 제1 표적 좌를 포함하는 세포를 제공하는 단계; (b) (i) Cas 단백질 및 제1 gRNA를 인코딩하는 하나 이상의 발현 작제물로서, 이들 각각은 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제1 핵산에서의 제1 gRNA 표적 부위에 Nick 또는 이중-가닥 브레이크를 유도하여 제1 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 발현 작제물, 및 (ii) 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 제2 핵산을 포함하는 제1 삽입 핵산을 포함하는 제1 표적화 벡터로서, 제1 삽입 핵산은 제1 표적 좌에 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹되는, 상기 제1 표적화 벡터를 세포 내로 도입하는 단계; (c) 제1 표적 좌에 제1 삽입 핵산을 포함하는 변형된 세포를 확인하는 단계로서, 변형된 세포는 제2 선택 마커의 활성은 갖지만 제1 선택 마커의 활성은 갖지 않는, 상기 변형된 세포를 확인하는 단계 - 여기서, 제1 및 제2 선택 마커는 상이함 -; (d) 제1 표적 좌에 제1 삽입 핵산을 포함하는 변형된 세포 내로 (i) Cas 단백질 및 제2 gRNA를 인코딩하는 하나 이상의 핵산으로서, 이들은 각각은 변형된 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제2 핵산을 포함하는 제1 삽입 핵산에서 제2 gRNA 표적 부위에 Nick 또는 이중-가닥 브레이크를 유도하여 제2 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 핵산; 및 (ii) 제3 프로모터에 작동가능하게 연결된 제3 선택 마커를 인코딩하는 제3 핵산을 포함하는 제2 삽입 핵산을 포함하는 제2 표적화 벡터로서, 제2 삽입 핵산은 제2 표적 좌에 위치한 제3 및 제4 표적 부위에 상응하는 제3 및 제4 상동성 아암에 의해 플랭킹되는, 상기 제2 표적화 벡터를 도입하는 단계; (e) 제2 표적 좌에 제2 삽입 핵산을 포함하는 제2 변형된 세포를 확인하는 단계로서, 제2 변형된 세포는 제3 선택 마커의 활성은 갖지만 제2 선택 마커의 활성은 갖지 않는, 상기 제2 변형된 세포를 확인하는 단계 - 여기서, 제2 및 제3 선택 마커는 상이함 -; (f) 제2 표적 좌에 제2 삽입 핵산을 포함하는 제2 변형된 세포 내로 (i) Cas 단백질 및 제3 gRNA를 인코딩하는 하나 이상의 발현 작제물로서, 이들 각각은 제2 변형된 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제3 핵산을 포함하는 제2 삽입 핵산에서 제3 gRNA 표적 부위에 Nick 또는 이중-가닥 브레이크를 유도하여 제3 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 발현 작제물, 및 (ii) 제4 프로모터에 작동가능하게 연결된 제4 선택 마커를 인코딩하는 제4 핵산을 포함하는 제3 삽입 핵산을 포함하는 제3 표적화 벡터로서, 제3 삽입 핵산은 제3 표적 좌에 위치한 제5 및 제6 표적 부위에 상응하는 제5 및 제6 상동성 아암에 의해 플랭킹되는, 상기 제3 표적화 벡터를 도입하는 단계; 및 (g) 제3 표적 좌에 제3 삽입 핵산을 포함하는 제3 변형된 세포를 확인하는 단계로서, 제3 변형된 세포는 제4 선택 마커의 활성은 갖지만 제3 선택 마커의 활성은 갖지 않는, 상기 제3 변형된 세포를 확인하는 단계를 포함하며, 제3 및 제4 선택 마커는 상이하다. 일부 실시 형태에서, 제1 및 제3 선택 마커는 동일하고 제2 및 제4 선택 마커는 동일하다. 일 실시 형태에서, 제1 및 제3 선택 마커는 동일하고, 제2 및 제4 선택 마커는 동일하고, 제1 및 제3 gRNA는 동일하다.

[0177] IV. 관심 폴리뉴클레오티드

[0178] 임의의 관심 폴리뉴클레오티드가 다양한 삽입 폴리뉴클레오티드에 삽입되어 표적 좌에 통합될 수 있다. 본 명세서에 개시된 방법은 표적화된 좌 내로 통합될 1개 이상, 2개 이상, 3개 이상, 4개 이상, 5개 이상, 6 이상 또는 그 이상의 관심 폴리뉴클레오티드를 제공한다.

[0179] 표적 좌에 통합될 때 삽입 폴리뉴클레오티드 내의 관심 폴리뉴클레오티드는 세포 내로 하나 이상의 유전자 변형을 도입할 수 있다. 게놈 변형은 내인성 핵산 서열의 결실 및/또는 표적 좌 내로의 외인성 또는 이중성 또는 이중상동성 폴리뉴클레오티드의 첨가를 포함할 수 있다. 일 실시 형태에서, 유전자 변형은 표적 좌에서 내인성 핵산 서열의 외인성 관심 폴리뉴클레오티드로의 대체를 포함한다. 따라서, 본 명세서에 제공된 방법은 녹아웃, 결실, 삽입, 대체 ("녹-인"), 점 돌연변이, 도메인 교체, 엑손 교체, 인트론 교체, 조절 서열 교체, 유전자 교체, 또는 이들의 조합을 포함하는 유전자 변형의 생성을 가능하게 한다. 그러한 변형은 표적 좌 내로의 제1, 제2, 제3, 제4, 제5, 제6, 제7, 또는 임의의 후속적인 삽입 폴리뉴클레오티드의 통합 시에 발생할 수 있다.

[0180] 삽입 폴리뉴클레오티드 내에 있고/있거나 표적 좌에 통합된 관심 폴리뉴클레오티드는 그것이 도입되는 세포에 고유하거나 그에 상동성인 서열을 포함할 수 있거나; 관심 폴리뉴클레오티드는 그것이 도입되는 세포에 이종성일 수 있거나; 관심 폴리뉴클레오티드는 그것이 도입되는 세포에 외인성일 수 있거나; 관심 폴리뉴클레오티드는 그것이 도입되는 세포에 이중상동성일 수 있거나; 관심 폴리뉴클레오티드는 그것이 도입되는 세포와 상이한 종으로부터 유래할 수 있다. 서열에 대하여, 용어 "상동성"은 세포에 고유한 서열을 포함한다. 서열에 대하여, 용어 "이종성"은 외래 종으로부터 기원하는 서열을 포함하거나, 동일한 종으로부터 기원하는 경우라면, 의도적인 인간 중재에 의해 조성 및/또는 좌에서 이의 고유 형태로부터 실질적으로 변형된 서열을 포함한다. 서열에



대하여, 용어 "외인성"은 외래 종으로부터 기원하는 서열을 포함한다. 용어 "이종상동성"은 다른 종 (즉, 종 변이체)에서 공지된 참조 서열과 기능적으로 동등한 하나의 종으로부터의 폴리뉴클레오타이드를 포함한다. 관심 폴리뉴클레오타이드는 비-인간, 설치류, 햄스터, 마우스, 래트, 인간, 원숭이, 농업용 포유류 또는 비-농업용 포유류를 포함하지만, 이로 제한되지는 않는 임의의 관심 유기체로부터 유래할 수 있다. 관심 폴리뉴클레오타이드는 코딩 영역, 비-코딩 영역, 조절 영역, 또는 게놈 DNA를 추가로 포함할 수 있다. 따라서, 제1, 제2, 제3, 제4, 제5, 제6, 제7, 및/또는 임의의 후속적인 삽입 폴리뉴클레오타이드는 그러한 서열을 포함할 수 있다.

[0181] 일 실시 형태에서, 삽입 폴리뉴클레오타이드 내의 및/또는 표적 좌에 통합된 관심 폴리뉴클레오타이드는 마우스 핵산 서열, 인간 핵산, 비-인간 핵산, 설치류 핵산, 래트 핵산, 햄스터 핵산, 원숭이 핵산, 농업용 포유류 핵산, 또는 비-농업용 포유류 핵산에 상동성이다. 또 다른 실시 형태에서, 표적 좌에 통합된 관심 폴리뉴클레오타이드는 게놈 핵산의 단편이다. 일 실시 형태에서, 게놈 핵산은 마우스 게놈 핵산, 인간 게놈 핵산, 비-인간 핵산, 설치류 핵산, 래트 핵산, 햄스터 핵산, 원숭이 핵산, 농업용 포유류 핵산 또는 비-농업용 포유류 핵산 또는 이들의 조합이다.

[0182] 일 실시 형태에서, 위에서 설명된 바와 같이, 관심 폴리뉴클레오타이드는 약 500개 뉴클레오타이드 내지 약 200 kb의 범위일 수 있다. 관심 폴리뉴클레오타이드는 약 500개 뉴클레오타이드 내지 약 5 kb, 약 5 kb 내지 약 200 kb, 약 5 kb 내지 약 10 kb, 약 10 kb 내지 약 20 kb, 약 20 kb 내지 약 30 kb, 약 30 kb 내지 약 40 kb, 약 40 kb 내지 약 50 kb, 약 60 kb 내지 약 70 kb, 약 80 kb 내지 약 90 kb, 약 90 kb 내지 약 100 kb, 약 100 kb 내지 약 110 kb, 약 120 kb 내지 약 130 kb, 약 130 kb 내지 약 140 kb, 약 140 kb 내지 약 150 kb, 약 150 kb 내지 약 160 kb, 약 160 kb 내지 약 170 kb, 약 170 kb 내지 약 180 kb, 약 180 kb 내지 약 190 kb, 또는 약 190 kb 내지 약 200 kb일 수 있다.

[0183] 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 삽입된 관심 폴리뉴클레오타이드는 폴리펩티드를 인코딩할 수 있거나, miRNA를 인코딩할 수 있거나, 이것은, 예를 들어, 조절 서열, 프로모터 서열, 인핸서 서열, 전사 억제 인자-결합 서열, 또는 비-단백질-코딩 서열의 결실을 포함하는 임의의 관심 조절 영역 또는 비-코딩 영역을 포함할 수 있다. 게다가, 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 삽입된 관심 폴리뉴클레오타이드는 신경계, 골격계, 소화계, 순환계, 근육계, 호흡계, 심혈관계, 림프계, 내분비계, 비뇨계, 생식계 또는 이들의 조합에서 발현되는 단백질을 인코딩할 수 있다. 일 실시 형태에서, 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 삽입된 관심 폴리뉴클레오타이드는 골수 또는 골수-유래 세포에서 발현되는 단백질을 인코딩한다. 일 실시 형태에서, 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 통합된 관심 폴리뉴클레오타이드는 비장 세포에서 발현되는 단백질을 인코딩한다. 또 다른 실시 형태에서, 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 삽입된 관심 폴리뉴클레오타이드는 B 세포에서 발현되는 단백질을 인코딩하거나, 미성숙 B 세포에서 발현되는 단백질을 인코딩하거나, 성숙 B 세포에서 발현되는 단백질을 인코딩한다.

[0184] 일 실시 형태에서, 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 삽입된 관심 폴리뉴클레오타이드는 면역글로불린 중쇄 가변 영역 아미노산 서열을 인코딩하는 게놈 핵산 서열을 포함한다. 문구 "중쇄," 또는 "면역글로불린 중쇄"는 임의의 유기체로부터의, 면역글로불린 중쇄 불변 영역 서열을 포함하는 면역글로불린 중쇄 서열을 포함한다. 달리 특정되지 않는다면, 중쇄 가변 도메인은 3개의 중쇄 CDR 및 4개의 프레임워크 (framework) (FR) 영역을 포함한다. 중쇄의 단편은 CDR, CDR 및 FR, 및 이들의 조합을 포함한다. 전형적인 중쇄는 (N-말단부터 C-말단으로의) 가변 도메인 이후에, CH1 도메인, 힌지 (hinge), CH2 도메인, 및 CH3 도메인을 갖는다. 중쇄의 기능적 단편은, 에피토프를 특이적으로 인식할 (예를 들어, 마이크로몰, 나노몰, 또는 피코몰 범위의 KD로 에피토프를 인식할) 수 있고, 세포로부터 발현하고 분비할 수 있고, 적어도 하나의 CDR을 포함하는 단편을 포함한다. 중쇄 가변 도메인은, 일반적으로 생식세포계열(germline)에 존재하는 VH, DH, 및 JH 분절의 레퍼토리오로부터 유래된 VH, DH, 및 JH 분절을 포함하는 가변 영역 뉴클레오타이드 서열에 의해 인코딩된다. 다양한 유기체에 대한 V, D, 및 J 중쇄 분절의 서열, 위치 및 명명법은 URL "imgt.org"에서 월드 와이드 웹 (www) 상의 인터넷을 통해 접근할 수 있는 IMGT 데이터베이스에서 확인될 수 있다.

[0185] 일 실시 형태에서, 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 통합된 관심 폴리뉴클레오타이드는 인간 면역글로불린 중쇄 가변 영역 아미노산 서열을 인코딩하는 게놈 핵산 서열을 포함한다. 일 실시 형태에서, 게놈 핵산 서열은 면역글로불린 중쇄 불변 영역 핵산 서열에 작동가능하게 연결된 비재배열된 인간 면역글로불린 중쇄 가변 영역 핵산 서열을 포함한다. 일 실시 형태에서, 면역글로불린 중쇄 불변 영역 핵산 서열은 마우스 면역글로불린 중쇄 불변 영역 핵산 서열 또는 인간 면역글로불린 중쇄 불변 영역 핵산 서열, 또는 이들의 조합이다. 일 실시 형태에서, 면역글로불린 중쇄 불변 영역 핵산 서열은 CH1, 힌지, CH2, CH3, 및 이들의 조합으로부터 선택된다. 일 실시 형태에서, 중쇄 불변 영역 핵산 서열은 CH1-힌지-CH2-CH3을 포함한다. 일 실시 형태

에서, 게놈 핵산 서열은 면역글로불린 중쇄 불변 영역 핵산 서열에 작동가능하게 연결된 재배열된 인간 면역글로불린 중쇄 가변 영역 핵산 서열을 포함한다. 일 실시 형태에서, 면역글로불린 중쇄 불변 영역 핵산 서열은 마우스 면역글로불린 중쇄 불변 영역 핵산 서열 또는 인간 면역글로불린 중쇄 불변 영역 핵산 서열, 또는 이들의 조합이다. 일 실시 형태에서, 면역글로불린 중쇄 불변 영역 핵산 서열은 CH1, 힌지, CH2, CH3, 및 이들의 조합으로부터 선택된다. 일 실시 형태에서, 중쇄 불변 영역 핵산 서열은 CH1-힌지-CH2-CH3을 포함한다.

[0186] 일 실시 형태에서, 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 통합된 관심 폴리뉴클레오타이드는 면역글로불린 경쇄 가변 영역 아미노산 서열을 인코딩하는 게놈 핵산 서열을 포함한다. 문구 "경쇄"는 임의의 유기체로부터의 면역글로불린 경쇄 서열을 포함하고, 달리 특정되지 않는다면, 인간 카파 ( $\kappa$ ) 및 람다 ( $\lambda$ ) 경쇄 및 VpreB, 및 대리 (surrogate) 경쇄를 포함한다. 달리 특정되지 않는다면, 경쇄 가변 도메인은 전형적으로 3개의 경쇄 CDR 및 4개의 FR을 포함한다. 일반적으로, 전장 경쇄는 아미노 말단에서 카르복실 말단으로, FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4, 및 경쇄 불변 영역 아미노산 서열을 포함하는 가변 도메인을 포함한다. 경쇄 가변 도메인은, 일반적으로 생식세포계열에 존재하는 경쇄 V 및 J 유전자 분절의 레퍼토리로부터 유래된 경쇄 VL 및 경쇄 JL 유전자 분절을 포함하는 경쇄 가변 영역 뉴클레오타이드 서열에 의해 인코딩된다. 다양한 유기체에 대한 경쇄 V 및 J 유전자 분절의 서열, 위치 및 명명법은 URL "imgt.org"에서 월드 와이드 웹 (www) 상의 인터넷을 통해 접근할 수 있는 IMGT 데이터베이스에서 확인될 수 있다. 경쇄는, 예를 들어, 그것이 나타나는 에피토프-결합 단백질에 의해 선택적으로 결합된 제1 또는 제2 에피토프 중 어느 하나와도 선택적으로 결합하지 않는 것을 포함한다. 경쇄는 또한, 그것이 나타나는 에피토프 결합 단백질에 의해 선택적으로 결합된 하나 이상의 에피토프에 결합하고 이를 인식하는 것과 함께, 중쇄에 결합하고 인식하거나 보조하는 것을 포함한다.

[0187] 일 실시 형태에서, 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 통합된 관심 폴리뉴클레오타이드는 인간 면역글로불린 경쇄 가변 영역 아미노산 서열을 인코딩하는 게놈 핵산 서열을 포함한다. 일 실시 형태에서, 게놈 핵산 서열은 비재배열된 인간  $\lambda$  및/또는  $\kappa$  경쇄 가변 영역 핵산 서열을 포함한다. 일 실시 형태에서, 게놈 핵산 서열은 재배열된 인간  $\lambda$  및/또는  $\kappa$  경쇄 가변 영역 핵산 서열을 포함한다. 일 실시 형태에서, 비재배열된 또는 재배열된  $\lambda$  및/또는  $\kappa$  경쇄 가변 영역 핵산 서열은  $\lambda$  경쇄 불변 영역 핵산 서열 및  $\kappa$  경쇄 불변 영역 핵산 서열로부터 선택되는 마우스, 래트, 또는 인간 면역글로불린 경쇄 불변 영역 핵산 서열에 작동가능하게 연결된다.

[0188] 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 통합된 관심 폴리뉴클레오타이드는 세포외 단백질 또는 수용체에 대한 리간드를 인코딩할 수 있다. 특정 실시 형태에서, 인코딩된 리간드는 사이토카인이다. 관심 사이토카인은 CCL, CXCL, CX3CL, 및 XCL로부터 선택되는 케모카인을 포함한다. 사이토카인은 또한 종양 괴사 인자 (tumor necrosis factor) (TNF)를 포함할 수 있다. 또 다른 실시 형태에서, 사이토카인은 인터류킨 (IL)이다. 일 실시 형태에서, 인터류킨은 IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IL-19, IL-20, IL-21, IL-22, IL-23, IL-24, IL-25, IL-26, IL-27, IL-28, IL-29, IL-30, IL-31, IL-32, IL-33, IL-34, IL-35, 및 IL-36으로부터 선택된다. 일 실시 형태에서, 인터류킨은 IL-2이다. 특정 실시 형태에서, 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 통합된 그러한 관심 폴리뉴클레오타이드는 인간으로부터 유래하고, 더욱 특정한 실시 형태에서는, 인간 서열을 포함할 수 있다.

[0189] 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 통합된 관심 폴리뉴클레오타이드는 세포질 단백질 또는 막 단백질을 인코딩할 수 있다. 일 실시 형태에서, 막 단백질은 수용체, 예컨대 사이토카인 수용체, 인터류킨 수용체, 인터류킨 2 수용체 알파, 인터류킨 2 수용체 베타, 또는 인터류킨 2 수용체 감마이다.

[0190] 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 통합된 관심 폴리뉴클레오타이드는, T 세포 수용체 알파를 포함하는 T 세포 수용체의 영역을 적어도 인코딩하는 폴리뉴클레오타이드를 포함할 수 있다. 특정 방법에서, 삽입 폴리뉴클레오타이드 각각은 T 세포 수용체 좌 (즉, T 세포 수용체 알파 좌)의 영역을 포함하여, 연속적인 통합의 완료 시, T 세포 수용체 좌의 일부 또는 전체가 표적 좌에 통합되게 되었다. 그러한 삽입 폴리뉴클레오타이드는 T 세포 수용체 좌의 (즉, T 세포 수용체 알파 좌의) 적어도 하나 또는 그 이상의 가변 분절 또는 결합 분절을 포함할 수 있다. 또 다른 실시 형태에서, T 세포 수용체의 영역을 인코딩하는 관심 폴리뉴클레오타이드는, 예를 들어, 돌연변이 단백질을 인코딩하는 포유류, 비-인간 포유류, 설치류, 마우스, 래트, 인간, 원숭이, 농업용 포유류 또는 사육용 포유류 폴리뉴클레오타이드로부터 유래할 수 있다.

[0191] 다른 실시 형태에서, 표적 좌에 통합된 관심 폴리뉴클레오타이드는 핵 단백질을 인코딩한다. 일 실시 형태에서, 핵 단백질은 핵 수용체이다. 특정 실시 형태에서, 삽입 폴리뉴클레오타이드 내에 있고/있거나 표적 좌에 통합된



그러한 관심 폴리뉴클레오티드는 인간으로부터 유래하고, 더욱 특정한 실시 형태에서는, 인간 게놈 서열을 포함할 수 있다.

[0192] 삽입 폴리뉴클레오티드 내에 있고/있거나 표적 좌에 통합된 관심 폴리뉴클레오티드는 코딩 서열에서의 유전자 변형을 포함할 수 있다. 그러한 유전자 변형은 코딩 서열의 결실 돌연변이 또는 2개의 코딩 서열의 융합을 포함하지만, 이로 제한되지는 않는다.

[0193] 삽입 폴리뉴클레오티드 내에 있고/있거나 표적 좌에 통합된 관심 폴리뉴클레오티드는 돌연변이 단백질을 인코딩하는 폴리뉴클레오티드를 포함할 수 있다. 일 실시 형태에서, 돌연변이 단백질은 변경된 결합 특성, 변경된 국소화, 변경된 발현, 및/또는 변경된 발현 패턴을 특징으로 한다. 일 실시 형태에서, 삽입 폴리뉴클레오티드 내에 있고/있거나 표적 좌에 통합된 관심 폴리뉴클레오티드는, 예를 들어, 신경학적 질병의 대립유전자, 심혈관 질병의 대립유전자, 신장 질병의 대립유전자, 근육 질병의 대립유전자, 혈액 질병의 대립유전자, 발암 유전자의 대립유전자, 또는 면역계 질병의 대립유전자를 포함하는 적어도 하나의 질병 대립유전자를 포함한다. 그러한 경우에, 질병 대립유전자는 우성 대립유전자일 수 있거나 질병 대립유전자는 열성 대립유전자이다. 더욱이, 질병 대립유전자는 단일 뉴클레오티드 다형성 (single nucleotide polymorphism) (SNP) 대립유전자를 포함할 수 있다. 돌연변이 단백질을 인코딩하는 관심 폴리뉴클레오티드는 돌연변이 단백질을 인코딩하는 포유류, 비-인간 포유류, 설치류, 마우스, 래트, 인간, 원숭이, 농업용 포유류 또는 사육용 포유류 폴리뉴클레오티드를 포함하지만, 이로 제한되지는 않는 임의의 유기체로부터 유래할 수 있다.

[0194] 삽입 폴리뉴클레오티드 내에 있고/있거나 표적 좌에 통합된 관심 폴리뉴클레오티드는 또한, 예를 들어, 프로모터 서열, 인핸서 서열, 또는 전사 억제인자-결합 서열을 포함하는 조절 서열을 포함할 수 있다. 특정 실시 형태에서, 삽입 폴리뉴클레오티드 내에 있고/있거나 표적 좌에 통합된 관심 폴리뉴클레오티드는, 비-단백질-코딩 서열의 결실을 갖지만, 단백질-코딩 서열의 결실은 포함하지 않는 폴리뉴클레오티드를 포함한다. 일 실시 형태에서, 비-단백질-코딩 서열의 결실은 조절 서열의 결실을 포함한다. 다른 실시 형태에서, 조절 요소의 결실은 프로모터 서열의 결실을 포함한다. 일 실시 형태에서, 조절 요소의 결실은 인핸서 서열의 결실을 포함한다. 그러한 관심 폴리뉴클레오티드는 돌연변이 단백질을 인코딩하는 포유류, 비-인간 포유류, 설치류, 마우스, 래트, 인간, 원숭이, 농업용 포유류 또는 사육용 포유류 폴리뉴클레오티드를 포함하지만, 이로 제한되지는 않는 임의의 유기체로부터 유래할 수 있다.

[0195] V. 서열 도입 방법 및 유전자도입 (Transgenic) 동물의 생성 방법

[0196] 위에서 개략적으로 설명된 바와 같이, 본 명세서에서는 하나 이상의 관심 폴리뉴클레오티드의 표적화된 통합을 가능하게 하는 방법 및 조성물이 제공된다. 그러한 시스템은 다양한 성분들을 사용하고 참고하기 쉽도록, 본 명세서에서 용어 "표적화된 통합 시스템"은 일반적으로 통합 이벤트에 필요한 모든 성분 (즉, 다양한 뉴클레아제 작용제, 인식 부위, 삽입 DNA 폴리뉴클레오티드, 표적화 벡터, 표적 좌, 및 관심 폴리뉴클레오티드)을 말한다.

[0197] 본 명세서에 제공되는 방법은 표적화된 통합 시스템의 다양한 성분들을 포함하는 하나 이상의 폴리뉴클레오티드 또는 폴리펩티드 작제물을 세포 내로 도입하는 단계를 포함한다. 용어 "도입하는"은 서열이 세포의 내부에 대한 접근을 얻게 하는 방식으로 서열 (폴리펩티드 또는 폴리뉴클레오티드)을 세포에 제시하는 단계를 포함한다. 본 명세서에 제공된 방법은, 단지 폴리뉴클레오티드가 적어도 하나의 세포의 내부에 대한 접근을 얻는, 표적화된 통합 시스템의 임의의 성분을 세포 내로 도입하기 위한 특정 방법에 의존하지 않는다. 다양한 세포 유형 내로 폴리뉴클레오티드를 도입하기 위한 방법은 당업계에 공지되어 있고 안정한 형질감염 방법, 일시적 형질감염 방법, 및 바이러스-매개 방법을 포함하지만 이로 제한되지는 않는다.

[0198] 일부 실시 형태에서, 본 방법 및 조성물에 사용된 세포는 그의 게놈 내로 안정하게 혼입된 DNA 작제물을 갖는다. "안정하게 혼입된" 또는 "안정하게 도입된"은, 뉴클레오티드 서열이 세포의 게놈 내로 통합되어 이의 자손에 의해 유전될 수 있도록 세포 내로의 폴리뉴클레오티드의 도입을 의미한다. 임의의 프로토콜이 DNA 작제물 또는 표적화된 통합 시스템의 다양한 성분들의 안정한 혼입을 위해 사용될 수 있다.

[0199] 형질감염 프로토콜 및 폴리펩티드 또는 폴리뉴클레오티드 서열의 세포 내로의 도입을 위한 프로토콜은 다양할 수 있다. 비제한적인 형질감염 방법은 리포좀; 나노입자; 인산칼슘 (문헌[Graham et al. (1973). Virology 52 (2): 456-67], 문헌[Bacchetti et al. (1977) Proc Natl Acad Sci USA 74 (4): 1590-4], 및 문헌[Kriegler, M (1991). Transfer and Expression: A Laboratory Manual. New York: W. H. Freeman and Company. pp. 96-97]); 덴드리머 (dendrimer); 또는 양이온성 중합체, 예컨대 DEAE-덱스트란 또는 폴리에틸렌이민의 사용을 포

합하는 화학적-기반 형질감염 방법을 포함한다. 비-화학적 방법은 전기천공; 초음파-천공 (Sono-poration); 및 광학적 형질감염을 포함한다. 입자-기반 형질감염은 유전자 총, 자성 보조 형질감염 (문헌[Bertram, J. (2006) Current Pharmaceutical Biotechnology 7, 277-28])의 사용을 포함한다. 바이러스 방법도 형질감염에 사용될 수 있다.

[0200] 일 실시 형태에서, 뉴클레아제 작용제는 표적화 벡터 또는 대형 표적화 벡터 (LTVEC)와 동시에 세포 내로 도입된다. 대안적으로, 뉴클레아제 작용제는 일정 기간에 걸쳐 표적화 벡터 또는 LTVEC와 별개로 도입된다. 일 실시 형태에서, 뉴클레아제 작용제는 표적화 벡터 또는 LTVEC의 도입 전에 도입되는 반면, 다른 실시 형태에서, 뉴클레아제 작용제는 표적화 벡터 또는 LTVEC의 도입 후에 도입된다.

[0201] 비-인간 포유류 동물은 본 명세서에 개시된 다양한 방법을 사용하여 생성될 수 있다. 그러한 방법은 다음을 포함한다: (1) 본 명세서에 개시된 방법을 사용하여 비-인간 동물의 다능성 세포의 표적 좌에 하나 이상의 관심 폴리뉴클레오티드를 통합하여 표적화된 좌에 삽입 폴리뉴클레오티드를 포함하는 유전자 변형된 다능성 세포를 생성하는 단계; (2) 표적 좌에 하나 이상의 관심 폴리뉴클레오티드를 갖는 유전자 변형된 다능성 세포를 선택하는 단계; (3) 유전자 변형된 다능성 세포를 전-상실기 (pre-morula stage)의 비-인간 동물의 숙주 배아 내로 도입하는 단계; 및 (4) 유전자 변형된 다능성 세포를 포함하는 숙주 배아를 대리모 내로 이식하여 유전자 변형된 다능성 세포로부터 유래한 F0 세대를 생성하는 단계. 비-인간 동물은 비-인간 포유류, 설치류 (예를 들어, 마우스, 래트, 햄스터), 원숭이, 농업용 포유류 또는 사육용 포유류일 수 있다. 다능성 세포는 인간 ES 세포, 인간 iPS 세포, 비-인간 ES 세포, 설치류 ES 세포 (예를 들어, 마우스 ES 세포, 래트 ES 세포, 또는 햄스터 ES 세포), 원숭이 ES 세포, 농업용 포유류 ES 세포 또는 사육용 포유류 ES 세포일 수 있다. 예를 들어, 미국 특허 출원 공개 제2014/0235933호; 미국 특허 출원 공개 제2014/0310828호; 및 문헌[Tong *et al.*(2010) *Nature*, 467(7312):211-213]을 참조하고, 이들 각각은 전체적으로 본 명세서에 참고로 포함된다.

[0202] 핵 전달 기술이 또한 비-인간 포유류 동물을 생성하기 위해 사용될 수 있다. 간략히 말해서, 핵 전달을 위한 방법은 하기 단계들을 포함한다: (1) 난모세포 (oocyte)를 제핵하는(enucleating) 단계; (2) 제핵된 난모세포와 조합될 공여자 세포 또는 핵을 분리하는 단계; (3) 세포 또는 핵을 제핵된 난모세포에 삽입하여 재구성된 세포를 형성하는 단계; (4) 재구성된 세포를 동물의 자궁 내로 이식하여 배아를 형성하는 단계; 및 (5) 배아를 발생되게 하는 단계. 그러한 방법에서 난모세포는 일반적으로 죽은 동물로부터 회수되긴 하지만, 또한 살아있는 동물의 난관 및/또는 난소로부터 분리될 수 있다. 난모세포는 제핵 전에 당업자에게 공지된 다양한 배지에서 성숙될 수 있다. 난모세포의 제핵은 당업자에게 널리 공지된 다수의 방식으로 수행될 수 있다. 공여자 세포 또는 핵을 제핵된 난모세포 내로 삽입하여 재구성된 세포를 형성하는 것은 일반적으로 융합 전에 투명대 (zona pellucida) 아래의 공여자 세포의 미세주입 (microinjection)에 의한다. 융합은 접촉/융합 평면(전기융합)을 가로지르는 DC 전기적 펄스의 적용에 의해, 폴리에틸렌 글리콜과 같은 융합-촉진 화학물질에 대한 세포의 노출에 의해, 또는 센다이 (Sendai) 바이러스와 같은 불활성화된 바이러스에 의해 유도될 수 있다. 재구성된 세포는 전형적으로 핵 공여자 및 수용자 난모세포의 융합 전, 융합 동안, 및/또는 융합 후 전기적 및/또는 비-전기적 수단에 의해 활성화된다. 활성화 방법은 전기적 펄스, 화학적으로 유도된 쇼크, 정자에 의한 침투, 난모세포에서 2가 양이온의 수준의 증가, 및 난모세포에서 (키나제 억제제에 의한) 세포 단백질의 인산화 감소를 포함한다. 활성화된 재구성된 세포, 또는 배아는 전형적으로 당업자에게 널리 공지된 배지에서 배양되고 이어서 동물의 자궁으로 전달된다. 예를 들어, 미국 특허 출원 공개 제20080092249호, 국제특허 공개 WO/1999/005266A2호, 미국 특허 출원 공개 제20040177390호, 국제특허 공개 WO/2008/017234A1호, 및 미국 특허 제 7,612,250호를 참조하고, 이들 각각은 본 명세서에 참고로 포함된다.

[0203] 본 명세서에 기재된 바와 같은 하나 이상의 유전자 변형을 생식세포계열 내에 포함하는 비-인간 동물을 만들기 위한 다른 방법이 제공되고, 본 방법은 하기 단계들을 포함한다: (a) 본 명세서에 기재된 다양한 방법을 사용하여 원핵 세포에서 비-인간 동물의 표적화된 좌를 변형시키는 단계; (b) 표적화된 좌에 유전자 변형을 포함하는 변형된 원핵 세포를 선택하는 단계; (c) 변형된 원핵 세포로부터 유전자 변형된 표적화 벡터를 분리하는 단계; (d) 유전자 변형된 표적화 벡터를 비-인간 동물의 다능성 세포 내로 도입하여 표적화된 좌에 삽입 핵산을 포함하는 유전자 변형된 다능성 세포를 생성하는 단계; (e) 유전자 변형된 다능성 세포를 선택하는 단계; (f) 유전자 변형된 다능성 세포를 전-상실기의 비-인간 동물의 숙주 배아 내로 도입하는 단계; 및 (g) 유전자 변형된 다능성 세포를 포함하는 숙주 배아를 대리모에 이식하여 유전자 변형된 다능성 세포로부터 유래한 F0 세대를 생성하는 단계. 그러한 방법에서, 표적화 벡터는 대형 표적화 벡터를 포함할 수 있다. 비-인간 동물은 비-인간 포유류, 설치류, 마우스, 래트, 햄스터, 원숭이, 농업용 포유류 또는 사육용 포유류일 수 있다. 다능성 세포는 인간 ES 세포, 인간 iPS 세포, 비-인간 ES 세포, 설치류 ES 세포 (예를 들어, 마우스 ES 세포, 래트 ES 세포,

또는 햄스터 ES 세포), 원숭이 ES 세포, 농업용 포유류 ES 세포 또는 사육용 포유류 ES 세포일 수 있다.

[0204] 추가의 방법에서, 단리하는 단계 (c)는 (c1) 유전자 변형된 표적화 벡터 (즉, 유전자 변형된 LTVEC)를 선형화하는 단계를 추가로 포함한다. 또 다른 실시 형태에서, 도입하는 단계 (d)는 (d1) 본 명세서에 기재된 바와 같은 뉴클레아제 작용제를 다능성 세포 내로 도입하는 단계를 추가로 포함한다. 다른 실시 형태에서, 도입하는 단계 (d)는 (d2)를 추가로 포함하는데, 여기서 비-인간 포유류의 다능성 세포는 세포에서 활성인 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 제1 폴리뉴클레오티드를 포함하는 표적 좌를 포함하고, 제1 폴리뉴클레오티드는 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 추가로 포함하고, 뉴클레아제 작용제를 다능성 세포 내로 도입하는 단계를 추가로 포함하고, 여기서 제1 뉴클레아제 작용제는 제1 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도한다. 다능성 세포 내로 추가로 도입하는 것은 변형된 원핵 세포의 게놈으로부터 유전자 변형된 표적화 벡터를 포함하는 제1 표적화 벡터이다. 변형된 표적화 벡터는 비-인간 포유류의 다능성 세포의 게놈 내의 제1 인식 부위에 충분히 가깝게 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암을 포함한다. 일 실시 형태에서, 선택하는 단계 (b) 및/또는 (e)는 원핵 세포 또는 다능성 세포에 본 명세서에 기재된 바와 같은 선택 작용제를 적용시킴으로써 수행된다. 일 실시 형태에서, 선택하는 단계 (b) 및/또는 (e)는 본 명세서에 기재된 바와 같은 대립유전자의 변형 (MOA) 검정법을 통해 수행된다.

[0205] 원핵 세포에서 박테리아 상동성 재조합 (BHR)을 통해 포유류 세포의 표적 좌를 변형시키기 위한 추가의 방법이 제공되고 이는 (a) 원핵 세포에서 활성인 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 제1 폴리뉴클레오티드를 포함하는 표적 좌를 포함하는 원핵 세포를 제공하는 단계로서, 제1 폴리뉴클레오티드는 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 추가로 포함하는, 상기 원핵 세포를 제공하는 단계, (b) 원핵 세포 내로 제1 상류 상동성 아암 및 제1 하류 상동성 아암으로 플랜킹된 삽입 폴리뉴클레오티드를 포함하는 표적화 벡터를 도입하는 단계로서, 삽입 폴리뉴클레오티드는 포유류 영역을 포함하는, 상기 표적화 벡터를 도입하는 단계, 및 원핵 세포 내로 제1 인식 부위에 또는 그 부근에 Nick 또는 이중-가닥 브레이크를 만드는 뉴클레아제 작용제를 도입하는 단계, 및 (c) 표적 좌에 삽입 폴리뉴클레오티드를 포함하는 표적화된 원핵 세포를 선택하는 단계로서, 원핵 세포는 BHR을 매개하는 재조합 유전성 단백질 및 효소를 발현할 수 있는, 상기 원핵 세포를 선택하는 단계를 포함한다. 단계 (a) 내지 (c)는 본 명세서에 개시된 바와 같이 연속적으로 반복되어 원핵 세포에서 표적화된 좌에 다수의 삽입 폴리뉴클레오티드의 도입을 가능하게 할 수 있다. 일단 표적화된 좌가 원핵 세포와 "조립(built)"되면, 변형된 표적 좌를 포함하는 표적화 벡터가 원핵 세포로부터 단리될 수 있고 비-인간 포유류 세포 내의 표적 좌 내로 도입될 수 있다. 이어서, 변형된 좌를 포함하는 포유류 세포는 비-인간 유전자도입 동물로 제조될 수 있다.

[0206] 일부 실시 형태에서, 본 명세서에 기재된 표적 좌의 다양한 유전자 변형은 벨로시진(VELOCIGENE)® 유전자 조작 기술을 사용하는 박테리아 인공 염색체 (BAC) DNA를 사용하는 박테리아 세포에서의 일련의 상동성 재조합 반응 (BHR)에 의해 수행될 수 있다 (예를 들어, 전체적으로 본 명세서에 참고로 포함되는 미국 특허 제6,586,251호 및 문헌[Valenzuela, D. M. et al. (2003), High-throughput engineering of the mouse genome coupled with high-resolution expression analysis, Nature Biotechnology 21(6): 652-659] 참조).

[0207] 일부 실시 형태에서, 본 명세서에 기재된 바와 같은 다양한 유전자 변형을 포함하는 표적화된 포유류 ES 세포 (즉, 비-인간 포유류, 설치류 (예를 들어, 마우스, 래트, 또는 햄스터), 농업용 포유류, 사육용 포유류, 원숭이 등 유래)는 벨로시마우스(VELOCIMOUSE)® 방법 (예를 들어, 미국 특허 제7,576,259호, 미국 특허 제7,659,442호, 미국 특허 제7,294,754호, 및 미국 특허 출원 공개 제2008-0078000 A1호를 참조하며, 이들 모두는 전체적으로 본 명세서에 참고로 포함됨)을 통해 상응하는 유기체로부터의 전-상실기 배아, 예를 들어, 8-세포기 마우스 배아 내로 도입된다. 유전자 변형된 ES 세포를 포함하는 비-인간 포유류 배아는 배반포기 (blastocyst stage) 까지 인큐베이션된 후 대리모 내로 이식되어 F0을 생성한다. 일부 다른 실시 형태에서, 본 명세서에 기재된 바와 같은 다양한 유전자 변형을 포함하는 표적화된 포유류 ES 세포는 배반포기 배아 내로 도입된다. 유전자 변형된 좌를 함유하는 비-인간 포유류는 본 명세서에 기재된 바와 같은 대립유전자의 변형 (MOA) 검정법을 통해 확인될 수 있다. 유전자 변형된 ES 세포로부터 유래된 생성된 F0 세대 비-인간 포유류는 야생형 비-인간 포유류와 교배시켜 F1 세대 자손을 수득한다. 특이적 프라이머 및/또는 프로브를 사용하여 유전자형 분석을 행한 후, 유전자 변형된 좌에 대해 이중접합성인 F1 비-인간 포유류를 서로 교배시켜, 유전자 변형된 좌에 대해 동중접합성인 비-인간 포유류를 생성한다.

[0208] VI. 세포

[0209] 본 명세서에 기재된 다양한 방법은 세포에서 좌 표적화 시스템을 사용한다. 그러한 세포는 원핵 세포, 예컨대

대장균 (*E. coli*)을 포함하는 박테리아 세포, 또는 진핵 세포, 예컨대 효모, 곤충, 양서류, 조류 (예를 들어, 닭 세포), 식물, 또는 마우스 세포, 래트 세포, 토끼 세포, 돼지 세포, 소 세포, 사슴 세포, 양 세포, 염소 세포, 고양이 세포, 개 세포, 페렛 세포, 영장류 (예를 들어, 마모셋, 붉은털 원숭이) 세포 등 및 사육된 포유류로부터의 세포 또는 농업용 포유류로부터의 세포를 포함하지만, 이로 제한되지는 않는 포유류 세포를 포함한다. 일부 세포는 비-인간, 특히 비-인간 포유류 세포이다. 일부 실시 형태에서, 적합한 유전자 변형가능한 다능성 세포가 용이하게 입수가가능하지 않은 포유류의 경우, 예를 들어 Oct3/4, Sox2, KLF4, Myc, Nanog, LIN28, 및 Glis1을 포함하지만 이로 제한되지는 않는 다능성-유도 인자들의 조합을 체세포 내로 도입시키는 것을 통해 체세포를 다능성 세포 내로 재프로그래밍하기 위해 다른 방법이 사용된다.

[0210] 일 실시 형태에서, 진핵 세포는 다능성 세포이다. 일 실시 형태에서, 다능성 세포는 배아 줄기 (ES) 세포이다. 용어 "배아 줄기 세포" 또는 "ES 세포"는 시험관내에서 미분화된 증식을 할 수 있는 배아-유래 전능성 또는 다능성 세포를 포함하고, 배아 내로 도입될 때 발달하는 배아의 어떠한 조직으로도 기여할 수 있다. 용어 "다능성 세포"는 하나 초과와 분화된 세포 유형으로 발달할 능력을 갖는 미분화된 세포를 포함한다. 폴리뉴클레오티드 서열에 대해 용어 "생식세포계열"은 자손으로 넘겨질 수 있는 핵산 서열을 포함한다.

[0211] 다능성 세포는 비-인간 ES 세포 또는 유도 다능성 줄기 (iPS) 세포일 수 있다. 일 실시 형태에서, 유도 다능성 (iPS) 세포는 섬유아세포로부터 유래된다. 특정 실시 형태에서, 유도 다능성 (iPS) 세포는 인간 섬유아세포로부터 유래된다. 일부 실시 형태에서, 다능성 세포는 조혈 줄기 세포 (HSC), 신경 줄기 세포 (NSC), 또는 상배엽 (epiblast) 줄기 세포이다. 다능성 세포는 또한 발생적으로 제한된 선조 세포일 수 있다. 다른 실시 형태에서, 다능성 세포는 설치류 다능성 세포이다. 일 실시 형태에서, 설치류 다능성 세포는 래트 다능성 세포 또는 래트 ES 세포이다. 다른 실시 형태에서, 설치류 다능성 세포는 마우스 다능성 세포 또는 마우스 ES 세포이다.

[0212] 다른 실시 형태에서, 포유류 세포는 불멸화된 마우스 세포, 래트 세포 또는 인간 세포일 수 있다. 일 실시 형태에서, 포유류 세포는 인간 섬유아세포인 반면, 다른 실시 형태에서, 포유류 세포는 인간 암 세포를 포함하는 암 세포이다.

[0213] 또 다른 실시 형태에서, 포유류는 인간이고 표적화는 생체의 (ex vivo) 인간 세포를 사용하여 수행된다.

[0214] 일 실시 형태에서, 포유류 세포는 질병을 가진 환자로부터 단리된 인간 세포이고/이거나 돌연변이 단백질을 인코딩하는 인간 폴리뉴클레오티드를 포함한다. 일 실시 형태에서, 돌연변이 인간 단백질은 변경된 결합 특성, 변경된 국소화, 변경된 발현, 및/또는 변경된 발현 패턴을 특징으로 한다. 일 실시 형태에서, 인간 핵산 서열은 적어도 하나의 인간 질병 대립유전자를 포함한다. 일 실시 형태에서, 인간 핵산 서열은 적어도 하나의 인간 질병 대립유전자를 포함한다. 일 실시 형태에서, 인간 질병 대립유전자는 신경 질병의 대립유전자이다. 일 실시 형태에서, 인간 질병 대립유전자는 심혈관 질병의 대립유전자이다. 일 실시 형태에서, 인간 질병 대립유전자는 신장 질병의 대립유전자이다. 일 실시 형태에서, 인간 질병 대립유전자는 근육 질병의 대립유전자이다. 일 실시 형태에서, 인간 질병 대립유전자는 혈액 질병의 대립유전자이다. 일 실시 형태에서, 인간 질병 대립유전자는 발암 유전자의 대립유전자이다. 일 실시 형태에서, 인간 질병 대립유전자는 면역계 질병의 대립유전자이다. 일 실시 형태에서, 인간 질병 대립유전자는 우성 대립유전자이다. 일 실시 형태에서, 인간 질병 대립유전자는 열성 대립유전자이다. 일 실시 형태에서, 인간 질병 대립유전자는 단일 뉴클레오티드 다형성 (SNP) 대립유전자를 포함한다.

[0215] 세포가 원핵 세포를 포함할 때, 특정 실시 형태에서, 원핵 세포는 대장균의 재조합-컴피턴트 균주 (recombination-competent strain)이다. 일 실시 형태에서, 원핵 세포는 재조합 유전성 단백질 및 효소를 인코딩하는 핵산을 포함한다. 일 실시 형태에서, 원핵 세포는 재조합 유전성 단백질 및 효소를 인코딩하는 핵산을 포함하지 않고, 재조합 유전성 단백질 및 효소를 인코딩하는 핵산은 원핵 세포 내로 도입된다. 일 실시 형태에서, 핵산은 재조합 유전성 단백질 및 효소를 인코딩하는 DNA 또는 mRNA를 포함한다. 일 실시 형태에서, 재조합 유전성 단백질 및 효소를 인코딩하는 핵산은 pABG이다. 일 실시 형태에서, 재조합 유전성 단백질 및 효소는 유도성 프로모터의 제어 하에서 발현된다. 일 실시 형태에서, 재조합 유전성 단백질 및 효소의 발현은 아라비노스에 의해 제어된다.

[0216] VII. 발현 카세트

[0217] 본 명세서에서는 본 명세서에서 제공되는 표적화된 통합 시스템의 다양한 성분들 (즉, 뉴클레아제 작용제, 인식 부위, 삽입 폴리뉴클레오티드, 관심 폴리뉴클레오티드, 표적화 벡터, 선택 마커 및 다른 성분들)을 포함하는 폴



리뉴클레오티드 또는 핵산 분자가 제공된다.

[0218] 용어 "폴리뉴클레오티드", "폴리뉴클레오티드 서열", "핵산 서열", 및 "핵산 단편"은 본 명세서에서 상호교환적으로 사용된다. 이들 용어는 뉴클레오티드 서열 등을 포함한다. 폴리뉴클레오티드는, 선택적으로 합성, 비-천연 또는 변경된 뉴클레오티드 염기를 함유하는 단일- 또는 이중-가닥 RNA 또는 DNA의 중합체일 수 있다. DNA의 중합체 형태의 폴리뉴클레오티드는 cDNA, 게놈 DNA, 합성 DNA, 또는 이들의 혼합물의 하나 이상의 분절로 구성될 수 있다. 폴리뉴클레오티드는 데옥시리보뉴클레오티드를 포함할 수 있고 리보뉴클레오티드는 자연 발생 분자 및 합성 유사체 둘 다, 및 이들의 임의의 조합을 포함한다. 본 명세서에 제공된 폴리뉴클레오티드는 또한 단일-가닥 형태, 이중-가닥 형태, 헤어핀, 줄기-및-루프 구조 등을 포함하지만 이로 제한되지는 않는 모든 형태의 서열을 포함한다.

[0219] 표적화된 통합 시스템의 다양한 성분들을 포함하는 재조합 폴리뉴클레오티드가 추가로 제공된다. 용어 "재조합 폴리뉴클레오티드" 및 "재조합 DNA 작제물"은 본 명세서에서 상호교환적으로 사용된다. 재조합 작제물은 핵산 서열의 인공 또는 이중성 조합, 예를 들어, 자연에서 함께 발견되지 않는 조절 및 코딩 서열을 포함한다. 다른 실시 형태에서, 재조합 작제물은 상이한 공급원으로부터 유래된 조절 서열 및 코딩 서열, 또는 동일한 공급원으로부터 유래되지만 자연에서 발견되는 것과는 상이한 방식으로 정렬된 조절 서열 및 코딩 서열을 포함할 수 있다. 그러한 작제물은 그 자체만으로 사용될 수 있거나 벡터와 연계하여 사용될 수 있다. 벡터가 사용되는 경우, 벡터의 선택은 당업자에게 널리 공지된 바와 같이 숙주 세포를 형질전환시키기 위해 사용되는 방법에 의존한다. 예를 들어, 플라스미드 벡터가 사용될 수 있다. 숙주 세포를 성공적으로 형질전환시키고, 선택하고, 증식시키기 위해 요구되고 임의의 단리된 핵산 단편을 포함하는 유전자 요소들이 본 명세서에 제공된다. 스크리닝은 무엇보다 DNA의 서던 (Southern) 분석, mRNA 발현의 노던 (Northern) 분석, 단백질 발현의 면역블롯팅 분석, 또는 표현형 분석에 의해 성취될 수 있다.

[0220] 특정 실시 형태에서, 본 명세서에 기재된 표적화된 통합 시스템의 하나 이상의 성분이 원핵 세포, 진핵 세포, 박테리아, 효모 세포, 또는 포유류 세포 또는 다른 유기체 또는 관심 세포 유형에서의 발현을 위한 발현 카세트에 제공될 수 있다. 카세트는 본 명세서에 제공된 폴리뉴클레오티드에 작동가능하게 연결된 5' 및 3' 조절 서열을 포함할 수 있다. "작동가능하게 연결된"은 2개 이상의 요소들 사이의 기능적 연결을 포함한다. 예를 들어, 관심 폴리뉴클레오티드와 조절 서열 (즉, 프로모터) 사이의 작동가능한 연결은 관심 폴리뉴클레오티드의 발현을 가능하게 하는 기능적 연결이다. 작동가능하게 연결된 요소는 연속적 또는 비-연속적일 수 있다. 2개의 단백질 코딩 영역들의 결합을 언급하기 위해 사용될 때, '작동가능하게 연결된'은 코딩 영역이 동일한 해독 프레임 (reading frame)에 있음을 의미한다. 다른 경우에, 단백질을 인코딩하는 핵산 서열이 적당한 전사 조절을 보유하도록 조절 서열 (예를 들어, 프로모터, 인핸서, 사일렌서 (silencer) 서열 등)에 작동가능하게 연결될 수 있다. 하나의 경우에, 번역글로불린 가변 영역 (또는 V(D)J 분절)의 핵산 서열이 번역글로불린 중쇄 또는 경쇄 서열 내로의 서열들 사이의 적당한 재조합을 가능하게 하도록 번역글로불린 불변 영역의 핵산 서열에 작동가능하게 연결될 수 있다.

[0221] 카세트는 유기체 내로 동시-도입될 수 있도록 적어도 하나의 추가의 관심 폴리뉴클레오티드를 추가로 함유할 수 있다. 대안적으로, 추가의 관심 폴리뉴클레오티드는 다수의 발현 카세트 상에 제공될 수 있다. 그러한 발현 카세트에는 조절 영역의 전사 조절 하에 있도록 재조합 폴리뉴클레오티드의 삽입을 위한 복수의 제한 부위 및/또는 재조합 부위가 제공된다. 발현 카세트는 추가적으로 선택 마커 유전자를 함유할 수 있다.

[0222] 발현 카세트는 5'-3'의 전사 방향으로 관심 포유류 세포 또는 숙주 세포에서 기능성인 전사 및 번역 개시 영역 (즉, 프로모터), 본 명세서에 제공된 재조합 폴리뉴클레오티드, 및 전사 및 번역 종결 영역 (즉, 종결 영역)을 포함할 수 있다. 본 명세서에 제공된 조절 영역 (즉, 프로모터, 전사 조절 영역, 및 번역 종결 영역) 및/또는 폴리뉴클레오티드는 숙주 세포에 또는 서로에 고유하고/유사할 수 있다. 대안적으로, 본 명세서에 제공된 조절 영역 및/또는 폴리뉴클레오티드는 숙주 세포에 또는 서로에 이중성일 수 있다. 예를 들어, 이중성 폴리뉴클레오티드에 작동가능하게 연결된 프로모터가, 폴리뉴클레오티드의 유래가 된 종과 상이한 종으로부터 유래하거나, 또는 만약 이것이 동일하고/유사한 종으로부터 유래한다면, 하나 또는 둘 다는 그들의 본래의 형태 및/또는 좌로부터 실질적으로 변형되거나, 프로모터는 작동가능하게 연결된 폴리뉴클레오티드에 대해 고유의 프로모터가 아니다. 대안적으로, 본 명세서에 제공된 조절 영역 및/또는 재조합 폴리뉴클레오티드는 전체적으로 합성일 수 있다.

[0223] 종결 영역은 전사 개시 영역과 함께 고유할 수 있거나, 작동가능하게 연결된 재조합 폴리뉴클레오티드와 함께 고유할 수 있거나, 숙주 세포와 함께 고유할 수 있거나, 프로모터, 재조합 폴리뉴클레오티드, 숙주 세포, 또는



이들의 임의의 조합에 대해 다른 공급원 (즉, 외래 또는 이중성)으로부터 유래할 수 있다.

[0224] 발현 카세트를 제조함에 있어서, 적당한 배향으로 DNA 서열을 제공할 수 있도록, 다양한 DNA 단편이 조작될 수 있다. 이러한 목적을 위해, 어댑터 (adapter) 또는 링커 (linker)를 사용하여 DNA 단편들을 결합할 수 있거나 다른 조각을 포함시켜 간편한 제한 부위, 피상적 DNA의 제거, 제한 부위의 제거 등을 제공할 수 있다. 이러한 목적을 위해, 시험관내 돌연변이생성, 프라이머 수선, 제한, 어닐링, 재치환, 예를 들어 전위 및 역위가 포함될 수 있다.

[0225] 다수의 프로모터가 본 명세서에 제공된 발현 카세트에 사용될 수 있다. 프로모터는 원하는 결과에 기초하여 선택될 수 있다. 상이한 적용은 관심 폴리뉴클레오티드의 타이밍, 위치 및/또는 발현 수준을 조절하기 위해 발현 카세트에서 상이한 프로모터를 사용하여 향상될 수 있는 것으로 인식된다. 그러한 발현 작제물은 또한, 원하는 경우, 프로모터 조절 영역 (예를 들어, 유도성, 구조성, 환경적으로- 또는 발생적으로- 조절된, 또는 세포- 또는 조직-특이적/선택적 발현을 부여하는 영역), 전사 개시 시작 부위, 리보솜 결합 부위, RNA 가공 신호, 전사 종결 부위, 및/또는 폴리아데닐화 신호를 함유할 수 있다.

[0226] 본 명세서에 제공된 폴리뉴클레오티드를 함유하는 발현 카세트는 또한 형질전환된 세포의 선택을 위한 선택 마커 유전자를 포함할 수 있다. 선택 마커 유전자는 형질전환된 세포 또는 조직의 선택을 위해 사용된다.

[0227] 적절한 경우, 본 방법 및 조성물에 사용된 서열 (즉, 관심 폴리뉴클레오티드, 뉴클레아제 작용제 등)은 세포에서의 증가된 발현을 위해 최적화될 수 있다. 즉, 유전자는 개선된 발현을 위해, 주어진 관심 세포에서, 예를 들어, 포유류-선택 코돈, 인간-선택 코돈, 설치류-선택 코돈, 마우스-선택 코돈, 래트-선택 코돈 등을 포함하는 선호되는 코돈을 사용하여 합성될 수 있다.

#### [0228] VIII. 서열 동일성

[0229] 본 명세서에 제공된 방법 및 조성물은 표적화된 통합 시스템의 다양한 상이한 성분들 (즉, 뉴클레아제 작용제, 인식 부위, 삽입 폴리뉴클레오티드, 관심 폴리뉴클레오티드, 표적화 벡터, 선택 마커 및 다른 성분들)을 사용한다. 표적화된 통합 시스템의 일부 성분들이 활성 변이체 및 단편을 가질 수 있다는 것이 상세한 설명 전반에 걸쳐 인식된다. 그러한 성분들은, 예를 들어, 뉴클레아제 작용제 (즉, 조작된 뉴클레아제 작용제), 뉴클레아제 작용제 인식 부위, 관심 폴리뉴클레오티드, 표적 부위 및 표적화 벡터의 상응하는 상동성 아암을 포함한다. 이들 성분 각각에 대한 생물학적 활성은 본 명세서의 어딘가 다른 곳에서 기재되어 있다.

[0230] 본 명세서에서 사용되는 바와 같이, 2개의 폴리뉴클레오티드 또는 폴리펩티드 서열과 관련하여 "서열 동일성" 또는 "동일성"은 2개의 서열에서 잔기들이 명시된 비교 창 (window)에 걸쳐 최대로 상응하도록 정렬되는 경우 동일하다는 것을 언급한다. 서열 동일성의 백분율이 단백질과 관련하여 사용될 때, 동일하지 않은 잔기 위치들이 흔히 보존적 아미노산 치환에 의해 차이가 있는 것으로 인지되고, 여기서 아미노산 잔기들은 유사한 화학적 성질 (예를 들어, 전하 또는 소수성)을 갖는 다른 아미노산 잔기들로 치환되므로 분자의 기능적 성질들을 변화시키지 않는다. 서열이 보존적 치환으로 차이가 있을 때, %서열 동일성은 치환의 보존적 성질에 대해 보정하기 위해 상향 조정될 수 있다. 그러한 보존적 치환에 의해 차이가 있는 서열은 "서열 유사성" 또는 "유사성"을 갖는다고 한다. 이러한 조정을 하기 위한 수단은 당업자에게 널리 공지되어 있다. 전형적으로 이것은 보존적 치환을 완전 미스매치 (mismatch)라기보다는 부분 미스매치로서 스코어링하는 것을 포함하여, %서열 동일성을 증가시킨다. 따라서, 예를 들어, 동일한 아미노산에는 1의 스코어가 주어지고 비-보존적 치환에는 0의 스코어가 주어지고, 보존적 치환에는 0과 1 사이의 스코어가 주어진다. 보존적 치환의 스코어링은, 예를 들어, 프로그램 PC/GENE (미국 캘리포니아주 마운틴 뷰 소재의 인텔리제네틱스 (Intelligenetics))에서 구현되는 바와 같이 계산된다.

[0231] 본 명세서에서 사용되는 바와 같이, "서열 동일성의 백분율"은 비교 창에 걸쳐 2개의 최적으로 정렬된 서열을 비교함으로써 결정된 값을 의미하고, 여기서 비교 창에서 폴리뉴클레오티드 서열의 일부는 2개의 서열의 최적의 정렬에 대해 참조 서열 (이는 부가 또는 결실을 포함하지 않음)과 비교하여 부가 또는 결실 (즉, 갭)을 포함할 수 있다. 백분율은 동일한 핵산 염기 또는 아미노산 잔기들이 양 서열에 존재하는 위치의 수를 결정하여 매칭된 위치의 수를 산출하고, 매칭된 위치의 수를 비교 창 내의 위치의 총수로 나누고, 결과에 100을 곱하여 서열 동일성의 백분율을 산출함으로써 계산된다.

[0232] 달리 언급되지 않는다면, 본 명세서에 제공된 서열 동일성/유사성 값은 하기의 파라미터를 사용하여 GAP 버전 10을 사용하여 획득된 값을 말한다: 50의 GAP 중량 및 3의 길이 중량, 및 nwsgapdna.cmp 스코어링 매트릭스를 사용한 뉴클레오티드 서열에 대한 %동일성 및 %유사성; 8의 GAP 중량 및 2의 길이 중량, 및 BLOSUM62 스코어링

매트릭스를 사용한 아미노산 서열에 대한 %동일성 및 %유사성; 또는 이와 동등한 임의의 프로그램. "동등한 프로그램"은 대상이 되는 임의의 2개의 서열에 대하여 GAP 버전 10에 의해 생성된 상응하는 정렬과 비교할 때 동일한 뉴클레오타이드 또는 아미노산 잔기 매치 및 동일한 %서열 동일성을 갖는 정렬을 생성하는 임의의 서열 비교 프로그램을 의미한다.

[0233] 비제한적인 실시 형태는 다음을 포함한다:

[0234] 1. (a) 세포를 제공하는 단계로서, 세포는 세포에서 활성인 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 제1 폴리뉴클레오타이드를 포함하는 표적 좌를 포함하고, 제1 폴리뉴클레오타이드는 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 추가로 포함하는, 상기 세포를 제공하는 단계, (b) (i) 제1 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도하는 제1 뉴클레아제 작용제; 및 (ii) 제1 인식 부위에 충분히 가깝게 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹된 제1 삽입 폴리뉴클레오타이드를 포함하는 제1 표적화 벡터를 세포 내로 도입하는 단계; 및 (c) 표적 좌에 통합된 제1 삽입 폴리뉴클레오타이드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함하는, 세포에서 표적 좌를 변형시키는 방법.

[0235] 2. (a) 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 제1 폴리뉴클레오타이드를 포함하는 제1 표적 좌를 포함하는 세포를 제공하는 단계로서, 제1 폴리뉴클레오타이드는 제1 뉴클레아제 작용제에 대한 제1 인식 부위를 추가로 포함하는, 상기 세포를 제공하는 단계, (b) (i) 세포에서 활성인 프로모터에 작동가능하게 연결된 제1 뉴클레아제 작용제를 인코딩하는 하나 이상의 발현 작제물로서, 제1 뉴클레아제 작용제는 제1 폴리뉴클레오타이드에서의 제1 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도하여 제1 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 발현 작제물; 및 (ii) 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 제2 폴리뉴클레오타이드를 포함하는 제1 삽입 폴리뉴클레오타이드를 포함하는 제1 표적화 벡터로서, 제1 삽입 핵산은 제1 표적 좌에 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹되는, 상기 제1 표적화 벡터를 세포 내로 도입하는 단계; 및 (c) 제1 표적 좌에 제1 삽입 핵산을 포함하는 변형된 세포를 확인하는 단계로서, 변형된 세포는 제2 선택 마커의 활성은 갖지만 제1 선택 마커의 활성은 갖지 않는, 상기 변형된 세포를 확인하는 단계를 포함하며, 제1 및 제2 선택 마커는 상이한, 세포에서 표적 좌를 변형시키기 위한 방법.

[0236] 3. 실시 형태 1 또는 실시 형태 2에 있어서, 표적 좌는 세포의 게놈 내에 있는, 방법.

[0237] 4. 실시 형태 1 또는 실시 형태 2에 있어서, 표적 좌는 세포 내의 벡터에 위치하는, 방법.

[0238] 5. 실시 형태 1 내지 실시 형태 4 중 어느 하나에 있어서, 제1 인식 부위에서의 Nick 또는 이중-가닥 브레이크는 제1 선택 마커의 활성을 방해하는, 방법.

[0239] 6. 실시 형태 1, 실시 형태 2, 실시 형태 3, 실시 형태 4, 또는 실시 형태 5에 있어서, 확인 단계 (c)는 제1 선택 마커의 활성을 갖지 않는 세포의 확인을 가능하게 하는 조건 하에서 세포를 배양하는 단계를 포함하는, 방법.

[0240] 7. 실시 형태 1 내지 실시 형태 6 중 어느 하나에 있어서, 제1 선택 마커를 포함하는 제1 폴리뉴클레오타이드는 제1 표적 부위 및 제2 표적 부위에 의해 플랭킹되는, 방법.

[0241] 8. 실시 형태 7에 있어서, 확인 단계 (c)는 제1 및 제2 표적 부위에 통합된 제1 삽입 폴리뉴클레오타이드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함하는, 방법.

[0242] 9. 실시 형태 1 내지 실시 형태 8 중 어느 하나에 있어서, 제1 삽입 폴리뉴클레오타이드는 (a) 제1 관심 폴리뉴클레오타이드; 및 (b) 세포에서 활성인 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 제2 폴리뉴클레오타이드로서, 제2 폴리뉴클레오타이드는 제2 뉴클레아제 작용제에 대한 제2 인식 부위를 포함하는, 상기 제2 폴리뉴클레오타이드를 포함하는, 방법.

[0243] 10. 실시 형태 9에 있어서, 방법은 (a) 표적 좌에 통합된 제1 삽입 폴리뉴클레오타이드를 포함하는 세포 내로 (i) 제2 인식 부위에 Nick 또는 이중-가닥 브레이크를 유도하는 제2 뉴클레아제 작용제; 및 (ii) 제2 인식 부위에 충분히 가깝게 위치한 제3 및 제4 표적 부위에 상응하는 제3 및 제4 상동성 아암에 의해 플랭킹된 제2 삽입 폴리뉴클레오타이드를 포함하는 제2 표적화 벡터를 도입하는 단계; 및 (b) 표적 좌에 통합된 제2 삽입 폴리뉴클레오타이드를 포함하는 적어도 하나의 세포를 확인하는 단계를 추가로 포함하는, 방법.

[0244] 11. 실시 형태 10에 있어서, 제2 인식 부위에서의 Nick 또는 이중 가닥 브레이크는 제2 선택 마커의 활성을 방해

하는, 방법.

- [0245] 12. 실시 형태 11에 있어서, 확인 단계 (b)는 제2 선택 마커의 활성을 갖지 않는 세포의 확인을 가능하게 하는 조건 하에서 세포를 배양하는 단계를 포함하는, 방법.
- [0246] 13. 실시 형태 10, 실시 형태 11, 또는 실시 형태 12에 있어서, 제2 선택 마커를 포함하는 제2 폴리뉴클레오티드는 제3 표적 부위 및 제4 표적 부위에 의해 플랭킹되는, 방법.
- [0247] 14. 실시 형태 13에 있어서, 확인 단계 (b)는 제3 및 제4 표적 부위에 통합된 제2 삽입 폴리뉴클레오티드를 포함하는 적어도 하나의 세포를 확인하는 단계를 포함하는, 방법.
- [0248] 15. 실시 형태 10 내지 실시 형태 14 중 어느 하나에 있어서, 제2 삽입 폴리뉴클레오티드는 (a) 제2 관심 폴리뉴클레오티드; 및 (b) 세포에서 활성인 제3 프로모터에 작동가능하게 연결된 제3 선택 마커를 인코딩하는 제3 폴리뉴클레오티드로서, 제3 폴리뉴클레오티드는 제3 뉴클레아제 작용제에 대한 제3 인식 부위를 포함하는, 상기 제3 폴리뉴클레오티드를 포함하는, 방법.
- [0249] 16. 실시 형태 9 내지 실시 형태 15 중 어느 하나에 있어서, 제1 뉴클레아제 작용제는 제2 뉴클레아제 작용제와 상이한, 방법.
- [0250] 17. 실시 형태 9 내지 실시 형태 16 중 어느 하나에 있어서, 제1 선택 마커는 제2 선택 마커와 상이한, 방법.
- [0251] 18. 실시 형태 15에 있어서, 제1 및 제3 뉴클레아제 인식 부위는 서로 동일하고 제2 뉴클레아제 인식 부위와 상이하고; 제1 및 제3 뉴클레아제 작용제는 서로 동일하고 제2 뉴클레아제 작용제와 상이한, 방법.
- [0252] 19. 실시 형태 15에 있어서, 제1 및 제3 선택 마커는 동일한, 방법.
- [0253] 20. 실시 형태 1 내지 실시 형태 19 중 어느 하나에 있어서, 제1, 제2 또는 제3 선택 마커 중 하나는 항생제에 대한 저항성을 부여하는, 방법.
- [0254] 21. 실시 형태 20에 있어서, 항생제는 G418, 하이그로마이신, 블라스토미딘, 네오마이신, 또는 푸로마이신을 포함하는, 방법.
- [0255] 22. 실시 형태 1 내지 실시 형태 19 중 어느 하나에 있어서, 제1, 제2 또는 제3 선택 마커 중 하나는 유도성 프로모터에 작동가능하게 연결되고, 선택 마커의 발현은 세포에 독성인, 방법.
- [0256] 23. 실시 형태 22에 있어서, 제1, 제2 또는 제3 선택 마커는 하이포잔틴-구아닌 포스포리보실트랜스퍼라제 (HGPRT) 또는 단순 포진 바이러스 (HSV-TK)의 티미딘 키나제를 포함하는, 방법.
- [0257] 24. 실시 형태 1 내지 실시 형태 23 중 어느 하나에 있어서, 상기 세포는 원핵 세포인, 방법.
- [0258] 25. 실시 형태 1 내지 실시 형태 23 중 어느 하나에 있어서, 세포는 진핵 세포인, 방법.
- [0259] 26. 실시 형태 25에 있어서, 진핵 세포는 포유류 세포인, 방법.
- [0260] 27. 실시 형태 26에 있어서, 포유류 세포는 비-인간 포유류 세포인, 방법.
- [0261] 28. 실시 형태 27에 있어서, 포유류 세포는 설치류 유래인, 방법.
- [0262] 29. 실시 형태 28에 있어서, 설치류는 래트 또는 마우스인, 방법.
- [0263] 30. 실시 형태 26 내지 실시 형태 29 중 어느 하나에 있어서, 세포는 다능성 세포인, 방법.
- [0264] 31. 실시 형태 26에 있어서, 포유류 세포는 인간 유도 다능성 줄기 (iPS) 세포인, 방법.
- [0265] 32. 실시 형태 30에 있어서, 다능성 세포는 비-인간 배아 줄기 (ES) 세포인, 방법.
- [0266] 33. 실시 형태 30에 있어서, 다능성 세포는 마우스 배아 줄기 (ES) 세포 또는 래트 배아 줄기 (ES) 세포인, 방법.
- [0267] 34. 실시 형태 30 내지 실시 형태 33 중 어느 하나에 있어서, 다능성 세포는 조혈 줄기 세포인, 방법.
- [0268] 35. 실시 형태 30 내지 실시 형태 33 중 어느 하나에 있어서, 다능성 세포는 신경 줄기 세포인, 방법.
- [0269] 36. 실시 형태 26에 있어서, 포유류 세포는 인간 섬유아세포인, 방법.
- [0270] 37. 실시 형태 1 또는 실시 형태 2에 있어서, 제1 표적화 벡터와 제1 뉴클레아제 작용제의 조합 사용은 제1 표

적화 벡터의 단독 사용과 비교하여 증가된 표적화 효율을 가져오는, 방법.

- [0271] 38. 실시 형태 37에 있어서, 제1 표적화 벡터의 표적화 효율은 제1 표적화 벡터의 단독 사용과 비교하여 2배 이상 증가되는, 방법.
- [0272] 39. 실시 형태 1 내지 실시 형태 38 중 어느 하나에 있어서, 제1 또는 제2 뉴클레아제 작용제는 뉴클레아제 작용제를 인코딩하는 핵산 서열을 포함하는 발현 작제물을 포함하고, 핵산은 세포에서 활성화된 제4 프로모터에 작동가능하게 연결되는, 방법.
- [0273] 40. 실시 형태 1 내지 실시 형태 39 중 어느 하나에 있어서, 제1 또는 제2 뉴클레아제 작용제는 뉴클레아제를 인코딩하는 mRNA인, 방법.
- [0274] 41. 실시 형태 1 내지 실시 형태 39 중 어느 하나에 있어서, 제1 또는 제2 뉴클레아제 작용제는 징크 핑거 뉴클레아제 (ZFN)인, 방법.
- [0275] 42. 실시 형태 1 내지 실시 형태 39 중 어느 하나에 있어서, 제1 또는 제2 뉴클레아제 작용제는 전사 활성화인자-유사 이펙터 뉴클레아제 (TALEN)인, 방법.
- [0276] 43. 실시 형태 1 내지 실시 형태 39 중 어느 하나에 있어서, 제1 또는 제2 뉴클레아제 작용제는 메가뉴클레아제인, 방법.
- [0277] 44. 실시 형태 1 내지 실시 형태 43 중 어느 하나에 있어서, 제1 또는 제2 뉴클레아제 작용제는 클러스터화된 규칙적으로 사이 공간을 둔 짧은 회문성 반복 (CRISPR)-관련 (Cas) 단백질 및 가이드 RNA (gRNA)를 포함하는, 방법.
- [0278] 45. 실시 형태 44에 있어서, 가이드 RNA (gRNA)는 (a) 제1, 제2, 또는 제3 인식 부위를 표적으로 하는 클러스터화된 규칙적으로 사이 공간을 둔 짧은 회문성 반복 (CRISPR) RNA (crRNA); 및 (b) 트랜스-활성화 CRISPR RNA (tracrRNA)를 포함하는, 방법.
- [0279] 46. 실시 형태 45에 있어서, 제1 또는 제2 인식 부위는 프로토스페이서 인접 모티프 (PAM) 서열에 의해 바로 플랭킹되는, 방법.
- [0280] 47. 실시 형태 44, 실시 형태 45 또는 실시 형태 46에 있어서, 관심 게놈 좌는 서열 번호 1의 뉴클레오티드 서열을 포함하는, 방법.
- [0281] 48. 실시 형태 44, 실시 형태 45, 실시 형태 46 또는 실시 형태 47에 있어서, Cas 단백질은 Cas9인, 방법.
- [0282] 49. 실시 형태 44 내지 실시 형태 46 중 어느 하나에 있어서, gRNA는 (a) 서열 번호 2의 핵산 서열의 키메라 RNA; 또는 (b) 서열 번호 3의 핵산 서열의 키메라 RNA를 포함하는, 방법.
- [0283] 50. 실시 형태 44 내지 실시 형태 46 중 어느 하나에 있어서, crRNA는 서열 번호 4; 서열 번호 5; 또는 서열 번호 6을 포함하는, 방법.
- [0284] 51. 실시 형태 44 내지 실시 형태 46 중 어느 하나에 있어서, tracrRNA는 서열 번호 7 또는 서열 번호 8을 포함하는, 방법.
- [0285] 52. 실시 형태 1 내지 실시 형태 51 중 어느 하나에 있어서, 제1, 제2, 및/또는 제3 인식 부위는 제1, 제2, 또는 제3 선택 마커의 인트론, 엑손, 프로모터, 프로모터 조절 영역, 또는 인핸서 영역에 위치하는, 방법.
- [0286] 53. 실시 형태 1 내지 실시 형태 52 중 어느 하나에 있어서, 제1 표적 부위 및 제2 표적 부위는 제1 인식 부위에 바로 인접한, 방법.
- [0287] 54. 실시 형태 10 내지 19 중 어느 하나의 방법에 있어서, 제1 표적 부위 및 제2 표적 부위는 제1 인식 부위로부터 약 10개 뉴클레오티드 내지 약 14 kb에 위치하는, 방법.
- [0288] 55. 실시 형태 10 내지 실시 형태 19 중 어느 하나에 있어서, 제3 표적 부위 및 제4 표적 부위는 제2 인식 부위에 바로 인접한, 방법.
- [0289] 56. 실시 형태 10 내지 19 중 어느 하나의 방법에 있어서, 제3 표적 부위 및 제4 표적 부위는 제2 인식 부위로부터 약 10개 뉴클레오티드 내지 약 14 kb에 위치하는, 방법.
- [0290] 57. 실시 형태 1 내지 실시 형태 56 중 어느 하나에 있어서, 제1 상동성 아암 및 제2 상동성 아암의 총 합계는



약 10 kb 이상인, 방법.

- [0291] 58. 실시 형태 10 내지 실시 형태 57 중 어느 하나에 있어서, 제3 상동성 아암 및 제4 상동성 아암의 총 합계는 약 10 kb 이상인, 방법.
- [0292] 59. 실시 형태 1 내지 실시 형태 58 중 어느 하나에 있어서, 제1 삽입 폴리뉴클레오티드의 길이는 약 5 kb 내지 약 300 kb 범위인, 방법.
- [0293] 60. 실시 형태 10 내지 실시 형태 59 중 어느 하나에 있어서, 제2 삽입 폴리뉴클레오티드의 길이는 약 5 kb 내지 약 300 kb 범위인, 방법.
- [0294] 61. 실시 형태 1 내지 실시 형태 60 중 어느 하나에 있어서, 제1 삽입 폴리뉴클레오티드의 표적 좌 내로의 통합은 녹아웃, 녹-인, 점 돌연변이, 도메인 교체, 엑손 교체, 인트론 교체, 조절 서열 교체, 유전자 교체, 또는 이들의 조합을 야기하는, 방법.
- [0295] 62. 실시 형태 10 내지 실시 형태 61 중 어느 하나에 있어서, 제2 삽입 폴리뉴클레오티드의 표적 좌 내로의 통합은 녹아웃, 녹-인, 점 돌연변이, 도메인 교체, 엑손 교체, 인트론 교체, 조절 서열 교체, 유전자 교체, 또는 이들의 조합을 야기하는, 방법.
- [0296] 63. 실시 형태 1 내지 실시 형태 62 중 어느 하나에 있어서, 제1 삽입 폴리뉴클레오티드는 인간 폴리뉴클레오티드를 포함하는 관심 폴리뉴클레오티드를 포함하는, 방법.
- [0297] 64. 실시 형태 8 내지 실시 형태 63 중 어느 하나에 있어서, 제2 삽입 폴리뉴클레오티드는 인간 폴리뉴클레오티드를 포함하는 관심 폴리뉴클레오티드를 포함하는, 방법.
- [0298] 65. 실시 형태 1 내지 실시 형태 64 중 어느 하나에 있어서, 제1 삽입 폴리뉴클레오티드는 T 세포 수용체 알파 좌의 영역을 포함하는 관심 폴리뉴클레오티드를 포함하는, 방법.
- [0299] 66. 제8항 내지 제65항 중 어느 하나에 있어서, 제2 삽입 폴리뉴클레오티드는 T 세포 수용체 알파 좌의 영역을 포함하는 관심 폴리뉴클레오티드를 포함하는, 방법.
- [0300] 67. 실시 형태 65 또는 실시 형태 66에 있어서, 제1 또는 제2 삽입 폴리뉴클레오티드는 T 세포 수용체 알파 좌의 적어도 하나의 가변 영역 유전자 분절 및/또는 결합 영역 유전자 분절을 포함하는 관심 폴리뉴클레오티드를 포함하는, 방법.
- [0301] 68. 실시 형태 65 내지 실시 형태 67 중 어느 하나에 있어서, T 세포 수용체 알파 좌의 영역은 인간 유래인, 방법.
- [0302] 69. 실시 형태 1 내지 실시 형태 64 중 어느 하나에 있어서, 제1 삽입 폴리뉴클레오티드는 비-인간 면역글로불린 중쇄 불변 영역 핵산 서열에 작동가능하게 연결된 비재배열된 인간 면역글로불린 중쇄 가변 영역 핵산 서열을 포함하는 관심 폴리뉴클레오티드를 포함하는, 방법.
- [0303] 70. 실시 형태 1 내지 실시 형태 69 중 어느 하나에 있어서, 확인 단계는 대립유전자의 변형 (MOA) 검정을 통해 수행되는, 방법.
- [0304] 71. 실시 형태 1 내지 실시 형태 65 중 어느 하나에 있어서, 제1 삽입 폴리뉴클레오티드는 세포의 게놈 내의 핵산 서열에 상동성 또는 이중상동성인 핵산 서열을 포함하는 관심 폴리뉴클레오티드를 포함하는, 방법.
- [0305] 72. 실시 형태 10 내지 실시 형태 71 중 어느 하나에 있어서, 제2 삽입 폴리뉴클레오티드는 세포의 게놈 내의 핵산 서열에 상동성 또는 이중상동성인 핵산 서열을 포함하는, 방법.
- [0306] 73. 실시 형태 1 내지 실시 형태 70 중 어느 하나에 있어서, 제1 삽입 폴리뉴클레오티드는 외인성 핵산 서열을 포함하는 관심 폴리뉴클레오티드를 포함하는, 방법.
- [0307] 74. 실시 형태 10 내지 실시 형태 70 또는 실시 형태 73 중 어느 하나에 있어서, 제2 삽입 폴리뉴클레오티드는 외인성 핵산 서열을 포함하는 관심 폴리뉴클레오티드를 포함하는, 방법.
- [0308] 기타 비제한적인 실시 형태는 다음을 포함한다:
- [0309] 1. (a) 제1 프로모터에 작동가능하게 연결된 제1 선택 마커를 인코딩하는 핵산을 포함하는 제1 표적 좌를 포함하는 세포를 제공하는 단계, (b) (i) Cas 단백질 및 제1 가이드 RNA (gRNA)를 인코딩하는 하나 이상의 발현 재질로서, 이들 각각은 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제1 핵산에서의 제1



gRNA 표적 부위에 Nick 또는 이중-가닥 브레이크를 유도하여 제1 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 발현 작제물; 및 (ii) 제2 프로모터에 작동가능하게 연결된 제2 선택 마커를 인코딩하는 제2 핵산을 포함하는 제1 삽입 핵산을 포함하는 제1 표적화 벡터로서, 제1 삽입 핵산은 제1 표적 좌에 위치한 제1 및 제2 표적 부위에 상응하는 제1 및 제2 상동성 아암에 의해 플랭킹되는, 상기 제1 표적화 벡터를 세포 내로 도입하는 단계; 및 (c) 제1 표적 좌에 제1 삽입 핵산을 포함하는 변형된 세포를 확인하는 단계로서, 변형된 세포는 제2 선택 마커의 활성은 갖지만 제1 선택 마커의 활성은 갖지 않는, 상기 변형된 세포를 확인하는 단계를 포함하며, 제1 및 제2 선택 마커는 상이한, 세포에서 관심 표적 좌를 변형시키는 방법.

- [0310] 2. 실시 형태 1에 있어서, 제1 gRNA는 제1 삽입 핵산에 혼성화되지 않는, 방법.
- [0311] 3. 실시 형태 1에 있어서, 관심 표적 좌는 세포의 게놈에 위치하는, 방법.
- [0312] 4. 실시 형태 1에 있어서, 관심 표적 좌는 세포 내의 벡터에 위치하는, 방법.
- [0313] 5. 실시 형태 1에 있어서, 확인 단계 (c)는 제2 선택 마커의 활성은 갖지만 제1 선택 마커의 활성은 갖지 않는 변형된 세포의 확인을 가능하게 하는 조건 하에서 세포를 배양하는 단계를 포함하는, 방법.
- [0314] 6. 실시 형태 1에 있어서, (d) 제1 표적 좌에 제1 삽입 핵산을 포함하는 변형된 세포 내로 (i) Cas 단백질 및 제2 gRNA를 인코딩하는 하나 이상의 핵산으로서, 이들 각각은 변형된 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제2 핵산을 포함하는 제1 삽입 핵산에서 제2 gRNA 표적 부위에 Nick 또는 이중-가닥 브레이크를 유도하여 제2 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 핵산; 및 (ii) 제3 프로모터에 작동가능하게 연결된 제3 선택 마커를 인코딩하는 제3 핵산을 포함하는 제2 삽입 핵산을 포함하는 제2 표적화 벡터로서, 제2 삽입 핵산은 제2 표적 좌에 위치한 제3 및 제4 표적 부위에 상응하는 제3 및 제4 상동성 아암에 의해 플랭킹되는, 상기 제2 표적화 벡터를 도입하는 단계; 및 (e) 제2 표적 좌에 제2 삽입 핵산을 포함하는 제2 변형된 세포를 확인하는 단계로서, 제2 변형된 세포는 제3 선택 마커의 활성은 갖지만 제2 선택 마커의 활성은 갖지 않는, 상기 제2 변형된 세포를 확인하는 단계를 추가로 포함하며, 제2 및 제3 선택 마커는 상이한, 방법.
- [0315] 7. 실시 형태 6에 있어서, 제1 및 제2 표적 좌는 서로 바로 인접하게 위치하는, 방법.
- [0316] 8. 실시 형태 6에 있어서, 제1 또는 제2 표적 좌는 제1 또는 제2 gRNA 표적 부위로부터 약 10개 뉴클레오타이드 내지 약 14 kb에 위치하는, 방법.
- [0317] 9. 실시 형태 8에 있어서, 제2 gRNA는 제2 삽입 핵산에 혼성화되지 않는, 방법.
- [0318] 10. 실시 형태 6에 있어서, 확인 단계 (e)는 제3 선택 마커의 활성은 갖지만 제2 선택 마커의 활성은 갖지 않는 제2 변형된 세포의 확인을 가능하게 하는 조건 하에서 변형된 세포를 배양하는 단계를 포함하는, 방법.
- [0319] 11. 실시 형태 6에 있어서, (f) 제2 표적 좌에 제2 삽입 핵산을 포함하는 제2 변형된 세포 내로 (i) Cas 단백질 및 제3 gRNA를 인코딩하는 하나 이상의 발현 작제물로서, 이들 각각은 제2 변형된 세포에서 활성인 프로모터에 작동가능하게 연결되고, Cas 단백질은 제3 핵산을 포함하는 제2 삽입 핵산에서 제3 gRNA 표적 부위에 Nick 또는 이중-가닥 브레이크를 유도하여 제3 선택 마커의 발현 또는 활성을 방해하는, 상기 하나 이상의 발현 작제물; 및 (ii) 제4 프로모터에 작동가능하게 연결된 제4 선택 마커를 인코딩하는 제4 핵산을 포함하는 제3 삽입 핵산을 포함하는 제3 표적화 벡터로서, 제3 삽입 핵산은 제3 표적 좌에 위치한 제5 및 제6 표적 부위에 상응하는 제5 및 제6 상동성 아암에 의해 플랭킹되는, 상기 제3 표적화 벡터를 도입하는 단계; 및 (g) 제3 표적 좌에 제3 삽입 핵산을 포함하는 제3 변형된 세포를 확인하는 단계로서, 제3 변형된 세포는 제4 선택 마커의 활성은 갖지만 제3 선택 마커의 활성은 갖지 않는, 상기 제3 변형된 세포를 확인하는 단계를 추가로 포함하며, 제3 및 제4 선택 마커는 상이한, 방법.
- [0320] 12. 실시 형태 11에 있어서, 제2 및 제3 표적 좌는 서로 바로 인접하게 위치하는, 방법.
- [0321] 13. 실시 형태 11에 있어서, 제2 또는 제3 표적 좌는 제1 또는 제2 gRNA 표적 부위로부터 약 10개 뉴클레오타이드 내지 약 14 kb에 위치하는, 방법.
- [0322] 14. 실시 형태 1 내지 실시 형태 13 중 어느 하나에 있어서, 제1, 제2, 제3, 또는 제4 마커는 항생제에 대해 저항성을 부여하는, 방법.
- [0323] 15. 실시 형태 14에 있어서, 항생제는 G418, 하이그로마이신, 블라스토시딘, 네오마이신, 또는 푸로마이신을 포함하는, 방법.

- [0324] 16. 실시 형태 1 내지 실시 형태 13 중 어느 하나에 있어서, 제1, 제2, 제3, 또는 제4 선택 마커는 하이포잔틴-구아닌 포스포리보실트랜스퍼라제 (HGPRT) 또는 단순 포진 바이러스 (HSV-TK)의 티미딘 키나제를 포함하는, 방법.
- [0325] 17. 실시 형태 1, 실시 형태 6, 또는 실시 형태 11에 있어서, 제1, 제2, 또는 제3 gRNA는 (i) 제1, 제2 또는 제3 gRNA 표적 부위에 혼성화되는 뉴클레오티드 서열 및 (ii) 트랜스-활성화 CRISPR RNA (tracrRNA)를 포함하는, 방법.
- [0326] 18. 실시 형태 1, 실시 형태 6, 또는 실시 형태 11에 있어서, 제1, 제2, 또는 제3 표적 좌는 제1, 제2 또는 제3 gRNA 표적 부위에 아주 가깝게 위치해서 gRNA 표적 부위에서의 Nick 또는 이중-가닥 브레이크가 표적 좌에서 표적화 벡터의 상동성 재조합을 촉진하도록 하는, 방법.
- [0327] 19. 실시 형태 1, 실시 형태 6, 또는 실시 형태 11에 있어서, Cas 단백질은 Cas9인, 방법.
- [0328] 20. 실시 형태 19에 있어서, 제1, 제2, 또는 제3 gRNA 표적 부위는 프로토스페이서 인접 모티프 (PAM) 서열에 의해 바로 플랭킹되는, 방법.
- [0329] 21. 실시 형태 1, 실시 형태 6, 또는 실시 형태 11에 있어서, 상기 세포는 원핵 세포인, 방법.
- [0330] 22. 실시 형태 1, 실시 형태 6, 및 실시 형태 11에 있어서, 세포는 진핵 세포인, 방법.
- [0331] 23. 실시 형태 22에 있어서, 진핵 세포는 포유류 세포인, 방법.
- [0332] 24. 실시 형태 23에 있어서, 포유류 세포는 섬유아세포인, 방법.
- [0333] 25. 실시 형태 23에 있어서, 포유류 세포는 비-인간 포유류 세포인, 방법.
- [0334] 26. 실시 형태 23에 있어서, 포유류 세포는 설치류 유래인, 방법.
- [0335] 27. 실시 형태 26에 있어서, 설치류는 래트, 마우스, 또는 햄스터인, 방법.
- [0336] 28. 실시 형태 22에 있어서, 진핵 세포는 다능성 세포인, 방법.
- [0337] 29. 실시 형태 28에 있어서, 다능성 세포는 조혈 줄기 세포 또는 신경 줄기 세포인, 방법.
- [0338] 30. 실시 형태 28에 있어서, 다능성 세포는 인간 유도 다능성 줄기 (iPS) 세포인, 방법.
- [0339] 31. 실시 형태 28에 있어서, 다능성 세포는 마우스 배아 줄기 (ES) 세포 또는 래트 배아 줄기 (ES) 세포인, 방법.
- [0340] 32. 실시 형태 1, 실시 형태 6, 및 실시 형태 11 중 어느 하나에 있어서, 제1, 제2, 또는 제3 gRNA 표적 부위는 제1, 제2, 또는 제3 선택 마커를 인코딩하는 제1, 제2, 또는 제3 핵산에서의 인트론, 엑손, 프로모터, 또는 프로모터 조절 영역에 위치하는, 방법.
- [0341] 33. 실시 형태 1, 실시 형태 6, 또는 실시 형태 11에 있어서, 제1, 제2, 또는 제3 표적화 벡터는 약 10 kb 이상인, 방법.
- [0342] 34. 실시 형태 1, 실시 형태 6, 또는 실시 형태 11에 있어서, 제1, 제2, 또는 제3 삽입 핵산은 약 5 kb 내지 약 300 kb의 범위인, 방법.
- [0343] 35. 실시 형태 1, 실시 형태 6, 또는 실시 형태 11에 있어서, 제1, 제2, 또는 제3 삽입 핵산은 인간 T 세포 수용체 알파 좌의 게놈 영역을 포함하는, 방법.
- [0344] 36. 실시 형태 35에 있어서, 게놈 영역은 인간 T 세포 수용체 알파 좌의 적어도 하나의 가변 영역 유전자 분절 및/또는 결합 영역 유전자 분절을 포함하는, 방법.
- [0345] 37. 실시 형태 6에 있어서, 제1 및 제3 선택 마커는 동일한, 방법.
- [0346] 38. 실시 형태 11에 있어서, 제1 및 제3 선택 마커는 동일하고 제2 및 제4 선택 마커는 동일한, 방법.
- [0347] 39. 실시 형태 38에 있어서, 제1 및 제3 gRNA는 동일한, 방법.
- [0348] 40. 실시 형태 1, 실시 형태 6, 실시 형태 37, 실시 형태 38 또는 실시 형태 39의 방법에 있어서, gRNA는 하이그로마이신 또는 네오마이신 저항성 유전자에 특이적인, 방법.

- [0349] 41. 실시 형태 40에 있어서, 네오마이신 저항성 유전자에 특이적인 gRNA는 서열 번호 13, 14, 15, 또는 16에 제시된 뉴클레오타이드 서열을 포함하는 핵산에 의해 인코딩되는, 방법.
- [0350] 42. 실시 형태 40에 있어서, 하이그로마이신 저항성 유전자에 특이적인 gRNA는 서열 번호 17, 18, 19, 또는 20에 제시된 뉴클레오타이드 서열을 포함하는 핵산에 의해 인코딩되는, 방법.
- [0351] 43. 실시 형태 6, 실시 형태 37, 실시 형태 38 또는 실시 형태 39에 있어서, a) 제1 gRNA는 서열 번호 13, 14, 15, 또는 16에 제시된 뉴클레오타이드 서열을 포함하는 핵산에 의해 인코딩되고 제2 gRNA는 서열 번호 17, 18, 19, 또는 20에 제시된 뉴클레오타이드 서열을 포함하는 핵산에 의해 인코딩되거나; 또는 b) 제1 gRNA는 서열 번호 17, 18, 19, 또는 20에 제시된 뉴클레오타이드 서열을 포함하는 핵산에 의해 인코딩되고 제2 gRNA는 서열 번호 13, 14, 15, 또는 16에 제시된 뉴클레오타이드 서열을 포함하는 핵산에 의해 인코딩되는, 방법.
- [0352] 하기 실시예들은 제한으로서가 아니라 예로서 제공된다.
- [0353] 실시예
- [0354] 도 1 및 도 2에 묘사된 순차적인 유전자 표적화 실험은 약물 선택 카세트에서 표적 서열을 인식하고 절단하도록 설계된 징크 핑거 뉴클레아제 (ZFN)와 대형 BAC-기반 표적화 벡터 (LTVEC)를 조합하는 것의 가치를 입증하였다.
- [0355] 순차적인 표적화에서의 제1 단계 (도 1)에 있어서는, 인간 T-세포 수용체 알파 (TCR  $\alpha$ )의 11개의 가변 (V) 도메인을 인코딩하는 136 kb의 DNA를 상응하는 마우스 TCR  $\alpha$  좌 내로 삽입하는 변형 (TCR  $\alpha$  B-hyg 대립유전자)을 만들도록 LTVEC를 작제하였다. 마우스 가변 (V) 및 결합 (J) 유전자 분절을 인간 V 및 J로 대체한, TCR  $\alpha$  좌에 미리 만들어진 변형 (TCR  $\alpha$  A-neo 대립유전자)을 가지고 있는 천만 개의 마우스 배아 줄기 (ES) 세포 내로, 작제된 LTVEC 0.02 mg을 전기천공시켰다. 성장 배지에서 전기천공된 ES 세포를 회수한 후, 하이그로마이신을 첨가하여, 그의 계능 내로 LTVEC가 혼입되었던 세포로부터 유래한 콜로니들을 선택하였다. 단리된 콜로니들의 대립유전자의 변형 (MOA) 스크리닝 결과, 2.9%의 표적화 효율에 대해, 스크리닝된 136개의 하이그로마이신 저항성 콜로니들 중 4개의 올바르게 표적화된 클론들이 확인되었다 (표 1, 실험 1). 11개의 추가적인 V의 삽입에 더하여, 올바르게 표적화된 클론들은 네오마이신 (G418) 저항성 카세트 (neo<sup>r</sup>)를 대체하였던 하이그로마이신 저항성 카세트 (hyg<sup>r</sup>)를 가졌다.
- [0356] 실험 2는 neor 유전자에서 인식 서열에 결합하고 DNA에서 이중-가닥 브레이크를 촉매하는 Neo-ZFN(1,2)의 각 반쪽을 발현한 2개의 플라스미드를 각각 0.02 mg 첨가한 것을 제외하고는 실험 1과 동일하였다. Neo-ZFN(1,2)의 포함으로 인해, 9.7%의 표적화 효율에 대해, 스크리닝된 568개의 하이그로마이신 저항성 클론들 중 55개의 올바르게 표적화된 클론들을 얻었으며, 이는 LTVEC 단독으로의 전기천공과 비교하여 3.3배 더 높은 표적화 효율을 나타낸다 (표 1, 실험 1과 2를 비교).
- [0357] 실험 3은 Neo-ZFN(3,4)를 인코딩하는 플라스미드로 Neo-ZFN(1,2)에 대한 플라스미드를 대체한 것을 제외하고는 실험 2와 동일하였다. Neo-ZFN(3,4)의 포함으로 인해, 11.7%의 표적화 효율에 대해, 스크리닝된 360개의 하이그로마이신 저항성 클론들 중 42개의 올바르게 표적화된 클론들을 얻었으며, 이는 LTVEC 단독으로의 전기천공과 비교하여 4배 더 높은 표적화 효율을 나타낸다 (표 1, 실험 1과 3을 비교).
- [0358] 순차적인 표적화에서의 제2 단계 (도 2)에 있어서는, TCR  $\alpha$  A-neo 또는 B-hyg 대립유전자에서의 것과 상이한 11개의 추가적인 인간 TCR  $\alpha$  가변 (V) 도메인을 인코딩하는 157 kb의 DNA를 순차적 표적화의 제1 단계 (도 1)에서 만들어진 TCR  $\alpha$  B-hyg 대립유전자를 가지고 있는 천만 개의 마우스 배아 줄기 (ES) 세포 내로 삽입하는 변형 (TCR  $\alpha$  C-neo 대립유전자)을 만들도록 설계된 LTVEC 0.002 mg을 전기천공시켜 ES 세포 내로 LTVEC를 도입하였다. 성장 배지에서 전기천공된 ES 세포를 회수한 후, G418를 첨가하여, 그의 계능 내로 LTVEC가 혼입되었던 세포로부터 유래한 콜로니들을 선택하였다. 단리된 콜로니들의 MOA 스크리닝 결과, 1.0%의 표적화 효율에 대해, 스크리닝된 192개의 G418 저항성 콜로니들 중 2개의 올바르게 표적화된 클론들이 확인되었다 (표 1, 실험 4). 11개의 추가적인 V의 삽입에 더하여, 올바르게 표적화된 클론들은 hyg<sup>r</sup> 카세트를 대체하였던 neor 카세트를 가졌다.
- [0359] 실험 5는 hyg<sup>r</sup>에서 인식 서열에 결합하고 DNA에서 이중-가닥 브레이크를 촉매하는 Hyg-ZFN(1,2)의 각 반쪽을 발현한 2개의 플라스미드를 각각 0.02 mg 첨가한 것을 제외하고는 실험 4와 동일하였다. Hyg-ZFN(1,2)의 포함으로 인해, 21%의 표적화 효율에 대해, 스크리닝된 192개의 G418 저항성 클론들 중 40개의 올바르게 표적화된 클론들을 얻었으며, 이는 LTVEC 단독으로의 전기천공과 비교하여 21배 더 높은 표적화 효율을 나타낸다 (표 1, 실험 5).

험 4와 5를 비교).

[0360] 실험 6은 Hyg-ZFN(3,4)를 인코딩하는 플라스미드로 Hyg-ZFN(1,2)에 대한 플라스미드를 대체한 것을 제외하고는 실험 5와 동일하였다. Hyg-ZFN(3,4)의 포함으로 인해, 22%의 표적화 효율에 대해, 스크리닝된 192개의 하이그로마이신 저항성 클론들 중 42개의 올바르게 표적화된 클론들을 얻었으며, 이는 LTVEC 단독으로의 전기천공과 비교하여 22배 더 높은 표적화 효율을 나타낸다 (표 1, 실험 4와 6을 비교).

[0361] 표 1에 요약된 실험들은 순차적 표적화 실험에서 LTVEC와 함께,  $neo^r$  또는  $hyg^r$  선택 카세트를 표적으로 하는 ZFN의 포함은 단지 LTVEC만을 포함하는 표적화 실험과 비교하여 3 내지 20배로 표적화 효율을 향상시킬 수 있다는 것을 확증하였다. 순차적 표적화 실험에서 ZFN의 포함에 의해 얻어진 향상된 표적화 효율은 인간 DNA의 매우 큰 조각 (136 kb 및 157 kb)의 미리 변형된 대립유전자의 원하는 염색체 위치에서의 정확한 올바른 의도된 삽입을 촉진하였다. ZFN-향상된 표적화는 표적화 프로젝트에서 성공 가능성을 크게 증가시키고 ES 세포 스크리닝에 대해 시간, 노동, 및 재료비에서의 유의한 절감을 산출한다.



[0362]

[표 1]

*neo<sup>r</sup>* 및 *hyg<sup>r</sup>* 서열을 인식하는 징크 핑거 뉴클레아제에 의한 순차적 유전자 표적화의 향상

실험	전기천공된 DNA	수용자 ES 세포	표적화 효율 <sup>1</sup> (%)	표적화의 향상 <sup>2</sup>
1	i. 인간화된 TCR $\alpha$ 대립유전자 B-hyg LTVEC <sup>3</sup>	인간화된 TCR $\alpha$ A-neo 대립유전자 <sup>3</sup>	2.9 (4/136)	해당 없음
2	i. 인간화된 TCR $\alpha$ 대립유전자 B-hyg LTVEC ii. Neo-ZFN1 플라스미드 <sup>4</sup> iii. Neo-ZFN2 플라스미드 <sup>4</sup>	인간화된 TCR $\alpha$ A-neo 대립유전자	9.7 (55/568)	3.3X
3	i. 인간화된 TCR $\alpha$ 대립유전자 B-hyg LTVEC ii. Neo-ZFN3 플라스미드 <sup>5</sup> iii. Neo-ZFN4 플라스미드 <sup>5</sup>	인간화된 TCR $\alpha$ A-neo 대립유전자	11.7 (42/360)	4.0X
4	i. 인간화된 TCR $\alpha$ 대립유전자 C-neo LTVEC <sup>6</sup>	인간화된 TCR $\alpha$ B-hyg 대립유전자 <sup>6</sup>	1.0 (2/192)	해당 없음
5	i. 인간화된 TCR $\alpha$ 대립유전자 C-neo LTVEC ii. Hyg-ZFN1 플라스미드 <sup>7</sup> iii. Hyg-ZFN2 플라스미드 <sup>7</sup>	인간화된 TCR $\alpha$ B-hyg 대립유전자	21 (40/192)	21X
6	i. 인간화된 TCR $\alpha$ 대립유전자 B-hyg LTVEC ii. Hyg-ZFN3 플라스미드 <sup>8</sup> iii. Hyg-ZFN4 플라스미드 <sup>8</sup>	인간화된 TCR $\alpha$ B-hyg 대립유전자	22 (42/192)	22X

<sup>1</sup> 백분율로 표현된, 스크리닝된 총 클론 수에 대한 올바르게 표적화된 클론 수의 비 (괄호 안에 표시됨)

<sup>2</sup> 표적화 백터를 단독으로 사용한 실험에 대한 표적화 백터와 ZFN 플라스미드를 조합한 실험의 표적화 효율의 비

<sup>3</sup> 도 1 참조

<sup>4</sup>Neo-ZFN(1,2) (도 3 참조) 뉴클레아제 결합 부위/절단 부위:  
GGGCGCCCGTTCTTTTT/gtcaag/ACCGACCTGTCCGGTG (서열 번호 9)

<sup>5</sup>Neo-ZFN(3,4) (도 3 참조) 뉴클레아제 결합 부위/절단 부위:  
CCGTTCTTTTTTGTC/aagacc/GACCTGTCCGGTGCC (서열 번호 10)

<sup>6</sup> 도 2 참조

<sup>7</sup>Hyg-ZFN(1,2) (도 3 참조) 뉴클레아제 결합 부위/절단 부위:  
TGCGATCGCTGCGGCCGA/tcttag/CCAGACGAGCGGGTTCGG (서열 번호 11)

<sup>8</sup>Hyg-ZFN(3,4) (도 3 참조) 뉴클레아제 결합 부위/절단 부위:  
CGCTGCGGCCGATCT/tagcca/GACGAGCGGGTTCGG (서열 번호 12)

[0363]

[0364] [표 2]

네오마이신 및 하이그로마이신 저항성 마커를 표적화하는 데 사용하기 위한 샘플 gRNA

서열 번호		gRNA
13	Neo Crispr#1	UGC GCAAGGAACGCCCCGUCG
14	Neo Crispr#2	GGCAGCGCGGCUAUCGUGGC
15	Neo Crispr#3	ACGAGGCAGCGCGGCUAUCG
16	Neo Crispr#4	GCUCUGAUGCCGCCGUGUUC
17	Hyg Crispr#1	ACGAGCGGGUUCGGCCCAUU
18	Hyg Crispr#6	CUUAGCCAGACGAGCGGGUU
19	Hyg Crispr#10	GCCGAUCUUAGCCAGACGAG
20	Hyg Crispr#16	CGACCUGAUGCAGCUCUCGG

[0365]

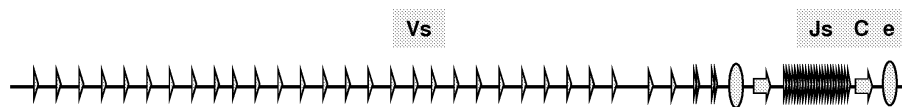
[0366]

본 명세서에서 언급된 모든 간행물 및 특허 출원은 본 발명이 속하는 기술분야의 당업자의 수준을 나타낸다. 모든 간행물 및 특허 출원은 마치 각 개별 간행물 또는 특허 출원이 구체적으로 그리고 개별적으로 참고로 포함되는 것처럼 그와 동일한 정도로 본 명세서에 참고로 포함된다. 달리 내용으로부터 자명하지 않는 한, 본 발명의 임의의 실시 형태, 태양, 단계 또는 특징은 임의의 다른 것과 조합되어 사용될 수 있다. 범위에 대한 언급은 범위 내의 임의의 정수, 범위 내의 임의의 하위 범위를 포함한다. 다수의 범위에 대한 언급은 그러한 범위의 복합을 포함한다.

## 도면

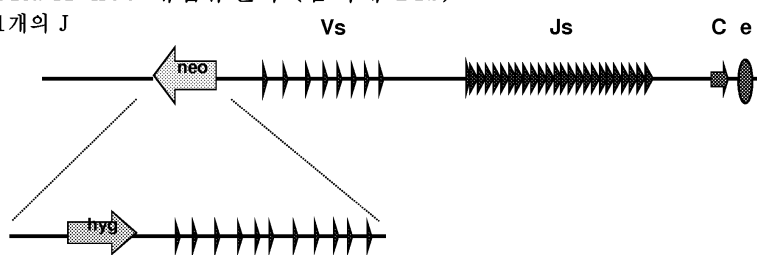
### 도면1

마우스 T-세포 수용체 알파 좌 (염색체 14a)



인간화된 TCRα A-neo 대립유전자 (염색체 14b)

8개의 V 및 61개의 J



인간화된 TCRα 대립유전자 B-hyg 표적화 벡터

11개의 추가적인 V

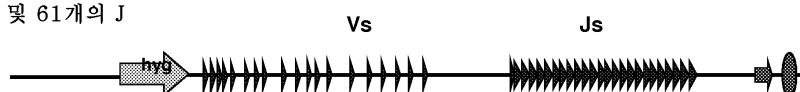
전기천공 + neo ZFN

마우스 T-세포 수용체 알파 좌 (염색체 14a)



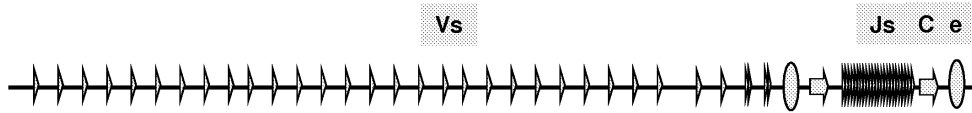
인간화된 TCRα B-hyg 대립유전자 (염색체 14b)

19개의 V 및 61개의 J



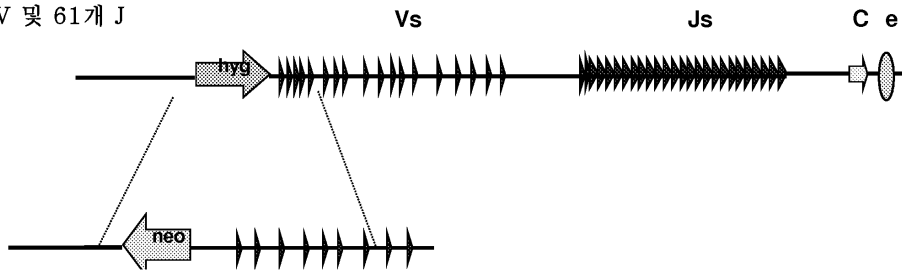
도면2

마우스 T-세포 수용체 알파 좌 (염색체 14a)



인간화된 TCRα B-hyg 대립유전자 (염색체 14b)

19개 V 및 61개 J



인간화된 TCRα 대립유전자 C-neo 표적화 벡터

11개의 추가적인 V

전기천공 + hyg ZFN

마우스 T-세포 수용체 알파 좌 (염색체 14a)



인간화된 TCRα C-neo 대립유전자 (염색체 14b)

30개 V 및 61개 J



### 도면3



### 서열 목록

#### SEQUENCE LISTING

<110> Auerbach, Wojtek

Frendewey, David

Droguett, Gustavo

Gagliardi, Anthony

Kuno, Junko

Valenzuela, David M.

<120> METHODS AND COMPOSITIONS

FOR MODIFYING A TARGETED LOCUS

<130> 057766-461003

<150> US 62/008,832

<151> 2014-06-06

<150> US 62/017,916

<151> 2014-06-27

<160> 21

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 23

<212> DNA



<213> Artificial Sequence

<220>

<223> a target locus that is linked to a guide RNA

(gRNA)

<220>

<221> misc\_feature

<222> (2)...(21)

<223> n = A, T, C, or G

<400> 1

gnnnnnnnnnn nnnnnnnnnn ngg

23

<210> 2

<211> 80

<212> RNA

<213> Artificial Sequence

<220>

<223> a guide RNA (gRNA)

<400> 2

guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc cguaaucaac uugaaaaagu 60

ggcaccgagu cgugcuuuu

80

<210> 3

<211> 42

<212> RNA

<213> Artificial Sequence

<220>

<223> a guide RNA (gRNA)

<400> 3

guuuuagagc uagaaauagc aaguuaaaau aaggcuaguc cg

42

<210> 4

<211> 30

<212> RNA

<213> Artificial Sequence

<220>

<223> a crRNA

<400> 4	
guuuuagagc uagaaauagc aaguuaaaau	30
<210> 5	
<211> 33	
<212> RNA	
<213> Artificial Sequence	
<220>	
<223> a crRNA	
<400> 5	
guuuuagagc uagaaauagc aaguuaaaau aag	33
<210> 6	
<211> 26	
<212> RNA	
<213> Artificial Sequence	
<220>	
<223> a crRNA	
<400> 6	
gaguccgagc agaagaagaa guuuua	26
<210> 7	
<211> 12	
<212> RNA	
<213> Artificial Sequence	
<220>	
<223> a tracrRNA	
<400> 7	
aaggcuaguc cg	12
<210> 8	
<211> 50	
<212> RNA	
<213> Artificial Sequence	
<220>	
<223> a tracrRNA	
<400> 8	
aaggcuaguc cguuaucaac uugaaaaagu ggcaccgagu cggugcuuuu	50

<210> 9

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Neo-ZFN(1,2): NUCLEASE BINDING SITE/cut site

<400> 9

gggcgcgccgg ttctttttgt caagaccgac ctgtccggtg 40

<210> 10

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> ZFN(3,4): NUCLEASE BINDING SITE/cut site

<400> 10

ccggtttcttt ttgtcaagac cgacctgtcc ggtgcc 36

<210> 11

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> ZFN(1,2): NUCLEASE BINDING SITE/cut site

<400> 11

tgcgatcgct gcggccgac ttagccagac gagcgggttc gg 42

<210> 12

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Hyg-ZFN(3,4): NUCLEASE BINDING SITE/cut site

<400> 12

cgctgcggcc gatcttagcc agacgagcgg gttcgg 36

<210> 13

<211> 20  
 <212> RNA  
 <213> Artificial Sequence  
 <220>  
 <223> gRNA Neo Crispr#1  
 <400> 13  
 ugcgcaagga acgcccguug 20  
 <210> 14  
 <211> 20  
 <212> RNA  
 <213> Artificial Sequence  
 <220>  
 <223> gRNA Neo Crispr#2  
 <400> 14  
 ggacgacggg cuaucguggc 20  
 <210> 15  
 <211> 20  
 <212> RNA  
 <213> Artificial Sequence  
 <220>  
 <223> gRNA Neo Crispr#3  
 <400> 15  
 acgaggcagc gcggcuauug 20  
 <210> 16  
 <211> 20  
 <212> RNA  
 <213> Artificial Sequence  
 <220>  
 <223> gRNA Neo Crispr#4  
 <400> 16  
 gcucugaugc cgccguguuc 20  
 <210> 17  
 <211> 20  
 <212> RNA

<213> Artificial Sequence

<220>

<223> gRNA Hyg Crispr#1

<400> 17

acgagcgggu ucggcccauu

20

<210> 18

<211> 20

<212> RNA

<213> Artificial Sequence

<220>

<223> gRNA Hyg Crispr#2

<400> 18

cuuagccaga cgagcggguu

20

<210> 19

<211> 20

<212> RNA

<213> Artificial Sequence

<220>

<223> gRNA Hyg Crispr#3

<400> 19

gccgaucuua gccagacgag

20

<210> 20

<211> 20

<212> RNA

<213> Artificial Sequence

<220>

<223> gRNA Hyg Crispr#4

<400> 20

cgaccugaug cagcucucgg

20

<210> 21

<211> 25

<212> DNA

<213> Artificial Sequence



<220>

<223> a target locus that is linked to a guide RNA  
(gRNA)

<220>

<221> misc\_feature

<222> (3)...(23)

<223> n = A, T, C, or G

<400> 21

ggnnnnnnnnn nnnnnnnnnn nnngg

25