



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0054445
(43) 공개일자 2017년05월17일

(51) 국제특허분류(Int. Cl.)
C12N 5/0786 (2010.01) *C07K 14/805* (2006.01)
C12N 15/11 (2006.01) *C12N 9/22* (2006.01)

(52) CPC특허분류
C12N 5/0645 (2013.01)
C07K 14/805 (2013.01)

(21) 출원번호 10-2017-7009275

(22) 출원일자(국제) 2015년09월16일
심사청구일자 없음

(85) 번역문제출일자 2017년04월05일

(86) 국제출원번호 PCT/US2015/050411

(87) 국제공개번호 WO 2016/044416
국제공개일자 2016년03월24일

(30) 우선권주장
62/051,159 2014년09월16일 미국(US)

(71) 출원인
상가모 테라퓨틱스, 인코포레이티드
미국 캘리포니아 94804 리치몬드 스위트 에이100
커넬 불러바드 501

(72) 발명자
코스트 그레고리 제이.

미국 캘리포니아 94804 리치몬드 스위트 에이100
커넬 불러바드 501 상가모 테라퓨틱스, 인코포레이티드 내

밀러 제프리 씨.

미국 캘리포니아 94804 리치몬드 스위트 에이100
커넬 불러바드 501 상가모 테라퓨틱스, 인코포레이티드 내

장 레이

미국 캘리포니아 94804 리치몬드 스위트 에이100
커넬 불러바드 501 상가모 테라퓨틱스, 인코포레이티드 내

(74) 대리인
송봉식, 정삼영

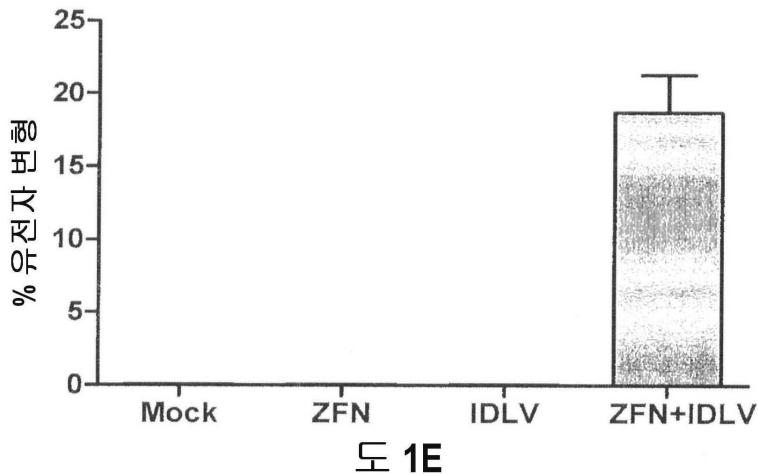
전체 청구항 수 : 총 26 항

(54) 발명의 명칭 뉴클레아제-매개된 게놈 엔지니어링 및 조혈 줄기세포에서의 교정을 위한 방법 및 조성물

(57) 요 약

본 발명은 게놈 엔지니어링 분야, 특히 조혈 줄기세포의 게놈의 표적화된 변형의 분야에 있다. 본원에는 발현을 변경시키기 위한 또는 겸상 적혈구 질환과 같은 유전자 질환에 포함된 단백질 (예컨대 질환의 생성 단백질 부족, 결핍 또는 일탈 및/또는 이를 단백질을 조절하는 단백질)을 코드화하는 하나 이상의 유전자를 교정하기 위한 방법 및 조성물이 개시된다. 본 발명은 유전자 요법 및 유전자 엔지니어링에 사용하기 위한 조성물 및 방법을 기술한다.

대 표 도



(52) CPC특허분류

C12N 15/111 (2013.01)

C12N 9/22 (2013.01)

C12N 2510/00 (2013.01)

명세서

청구범위

청구항 1

내인성 인간 베타-헤모글로빈 (Hbb) 유전자의 SEQ ID NO:23 또는 SEQ ID NO:24에서 또는 가까이에서 게놈 변형을 포함하는 유전자 변형된 세포로서, 게놈 변형이 삽입, 결실 및 그것들의 조합으로 구성되는 군으로부터 선택되는 유전자 변형된 세포.

청구항 2

제1 항에 있어서, 세포는 줄기세포인 것을 특징으로 하는 유전자 변형된 세포.

청구항 3

제2 항에 있어서, 줄기세포는 조혈 줄기세포인 것을 특징으로 하는 유전자 변형된 세포.

청구항 4

제3 항에 있어서, 조혈 줄기세포는 CD34+ 세포인 것을 특징으로 하는 유전자 변형된 세포.

청구항 5

제1 항 내지 제4 항 중 어느 한 항의 줄기세포로부터 파생된 유전자 변형된 분화된 세포.

청구항 6

제5 항에 있어서, 세포는 적혈구 (RBC)인 것을 특징으로 하는 유전자 변형된 세포.

청구항 7

제1 항 내지 제6 항 중 어느 한 항에 있어서, 유전자 변형은 뉴클레아제를 사용하여 만들어지는 것을 특징으로 하는 유전자 변형된 세포.

청구항 8

제7 항에 있어서, 뉴클레아제는 징크 평거 뉴클레아제인 것을 특징으로 하는 유전자 변형된 세포.

청구항 9

제8 항에 있어서, 뉴클레아제는 징크 평거 뉴클레아제를 포함하고, 징크 평거 뉴클레아제는 인식 나선을 포함하는 5 또는 6 징크 평거 도메인을 포함하며 추가로 징크 평거 단백질이 표 1의 한 개의 열에 나타낸 순서로 인식 나선 영역을 포함하는 것을 특징으로 하는 유전자 변형된 세포.

청구항 10

제1 항 내지 제9 항 중 어느 한 항에 있어서, 삽입은 전이유전자를 코드화하는 도너 폴리뉴클레오티드의 통합을 포함하는 것을 특징으로 하는 유전자 변형된 세포.

청구항 11

제10 항에 있어서, 도너 폴리뉴클레오티드는 겹상 적혈구 돌연변이를 교정하는 것을 특징으로 하는 유전자 변형된 세포.

청구항 12

제1 항 내지 제11 항 중 어느 한 항의 유전자 변형된 세포를 포함하는 약학 조성물.

청구항 13

인식 나선 영역을 포함하는 5 또는 6 징크 평거 도메인을 포함하는 징크 평거 단백질로서, 징크 평거 단백질은 표 1의 한 개의 열에 나타낸 순서로 인식 나선 영역을 포함하는 징크 평거 단백질.

청구항 14

제13 항의 징크 평거 단백질 및 야생형 또는 엔지니어링된 절단 하프-도메인을 포함하는 융합 단백질.

청구항 15

제13 항 또는 제14 항의 하나 이상의 단백질을 코드화하는 폴리뉴클레오티드.

청구항 16

제14 항의 하나 이상의 융합 단백질 또는 제15 항의 하나 이상의 폴리뉴클레오티드를 포함하는 분리된 세포.

청구항 17

제16 항에 있어서, 세포는 적혈구 (RBC) 또는 전구 세포 (예컨대 CD4+ 조혈 줄기세포)로 구성되는 군으로부터 선택되는 것을 특징으로 하는 세포.

청구항 18

i) 제15 항에 따르는 단백질을 코드화하는 폴리뉴클레오티드 또는 ii) 제13 항 또는 제14 항에 따르는 단백질 중 적어도 하나를 포함하는 키트.

청구항 19

세포에서 Hbb를 변경시키는 방법으로서,

세포에 제15 항에 따르는 하나 이상의 폴리뉴클레오티드를, 하나 이상의 단백질이 발현되고 Hbb 유전자의 발현이 변경되는 조건하에서 도입하는 단계를 포함하는 방법.

청구항 20

제19 항에 있어서, Hbb 유전자의 발현은 증가되는 것을 특징으로 하는 방법.

청구항 21

제19 항 또는 제20 항에 있어서, 도너 서열을 세포의 게놈 안으로 통합시키는 단계를 더 포함하는 것을 특징으로 하는 방법.

청구항 22

제19 항 내지 제21 항 중 어느 한 항에 있어서, 세포는 적혈구 (RBC) 전구 세포 및 조혈 줄기세포로 구성되는 군으로부터 선택되는 것을 특징으로 하는 방법.

청구항 23

제19 항 내지 제22 항 중 어느 한 항에 있어서, 도너 서열은 Hbb 유전자의 mRNA 프로세싱을 위한 스플라이스 도너 부위를 변경시키는 돌연변이를 교정하는 올리고뉴클레오티드를 포함하는 것을 특징으로 하는 방법.

청구항 24

내인성 Hbb 유전자 내에서의 게놈 변형을 포함하는 유전자 변형된 세포의 제조 방법으로서,

- a) 세포를 제14 항에 따르는 융합 단백질을 코드화하는 폴리뉴클레오티드와 접촉시키는 단계;
- b) 세포를 폴리뉴클레오티드로부터 융합 단백질을 발현시키기에 좋은 조건하에 두는 단계; 및
- c) 내인성 Hbb 유전자를 유전자 변형된 세포를 생성하기에 충분한 발현된 융합 단백질로 변형시키는 단계를 포함하는 방법.

청구항 25

제24 항에 있어서, 방법이 적어도 하나의 사이토카인으로 세포를 자극하는 단계를 더 포함하는 것을 특징으로 하는 방법.

청구항 26

β -지중해 빈혈에 걸린 환자의 치료 방법으로서,

제12 항에 따르는 약학 조성물을 환자에서 Hbb 유전자 발현을 증가시키기에 충분한 양으로 환자에게 투여하는 단계를 포함하는 방법.

발명의 설명

기술 분야

[0001] 관련 출원에 대한 상호-참조

본 출원은 2014년 9월 16일자로 출원된 미국 임시 출원 번호 62/051,159의 유익을 주장하며, 그것의 내용은 전체적으로 참조로 본원에 포함된다.

[0003] 기술분야

[0004] 본 발명은 게놈 엔지니어링 분야, 특히 조혈 세포의 게놈의 표적화된 변형에 있다.

배경 기술

[0005] 대다수의 질환에 대한 유전자 요법에서 가장 전도유망한 접근법 중 한 가지는 줄기세포의 시험관 내 유전자 변형과 이어서 변형된 세포의 환자에서의 이식 및 접목의 사용을 포함한다. 특히 유망한 것은 도입된 줄기세포가 장기간 지속성 및 다중-혈통 세포분화를 나타낼 때이다. 가장 통상적으로 CD34 세포 표면 마커의 발현을 기반으로 한 풍부해진 세포 형태의 조혈 줄기세포는, 그것들이 쉽게 얻어질 수 있고 장기간 조혈 줄기세포 (LT-HSC)를 함유할 수 있고, 그것은 이식 후 완전한 조혈 혈통을 복원할 수 있기 때문에 특히 유용한 세포 집단이다.

[0006] 게놈 DNA의 표적화된 절단을 위한 다양한 방법 및 조성물들이 기술되어 있다. 그런 표적화된 절단 사건은, 예를 들면 표적화된 돌연변이 생성을 유도하고, 세포 DNA 서열의 표적화된 결실을 유도하며, 임의의 유기체로부터의 세포에서 예정된 염색체 유전자좌의 표적화된 재조합을 촉진시키기 위해 사용될 수 있다. 예컨대 미국 특허 번호 9,045,763; 9,005,973; 8,956,828; 8,945,868; 8,586,526; 6,534,261; 6,599,692; 6,503,717; 6,689,558; 7,067,317; 7,262,054; 7,888,121; 7,972,854; 7,914,796; 7,951,925; 8,110,379; 8,409,861; 미국 특허 공개 공보 20030232410; 20050208489; 20050026157; 20050064474; 20060063231; 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983 및 20130177960 및 20150056705 및 미국 출원 번호 14/706,747를 참조하는데, 모든 문서의 내용은 모든 목적에 대해 전체 내용이 참조로 포함된다. 이를 방법은 표적 DNA 서열에 이중 가닥 브레이크 (DSB) 또는 닉 (nick)을 유도하여서 비-상동성 단부 결합 (NHEJ)과 같은 오류 발생 과정에 의한 브레이크의 수복 또는 수복 주형을 사용한 수복 (상동성 지시된 수복 또는 HDR)이 유전자의 녹아웃 또는 관심 서열의 삽입 (표적화된 통합)을 초래할 수 있는 엔지니어링된 절단 시스템의 사용을 포함한다. 뒤따르는 수복 경로 (NHEJ 대 HDR 또는 둘 다)는 전형적으로 수복 주형의 존재 및 여러 경합하는 수복 경로의 활성에 좌우된다.

[0007] 외부에서 공급되는 수복 주형 (예컨대 도너 (donor))의 부재시 이중 가닥 브레이크의 도입은 통상적으로 세포 NHEJ 경로에 의해 도입된 돌연변이를 통한 표적화된 유전자의 비활성화를 위해 사용된다. NHEJ 경로는 미세상동 성-매개된 단부 결합을 기반으로 표준 Ku-의존성 경로 및 대체 Ku-무관 경로로 나누어질 수 있는데, 절단 부위 가까이에 직접 반복부의 짧은 트랙의 장점을 취한다. 이를 두 가지 NHEJ 경로를 통한 유전자 편집에 이어지는 삽입 및/또는 결실 ('indel')의 패턴은 용도에 따라 상이하고, 돌연변이의 기능적 결과의 차이를 초래할 수 있다.

[0008] 이중 가닥 브레이크의 양옆의 서열들에 상동하는 외부에서 공급된 도너 운반 신장부의 존재하에, 도너 분자를 사용하는 상동성 지시된 유전자 수복 (HDR)은 단일 염기 또는 DNA의 작은 신장부의 서열을 변화시키기 위해 ('유전자 교정'), 또는 다른 한편으로, 예정된 게놈 위치 안으로의 전체 발현 카세트의 표적화된 삽입 ('유전자

첨가')을 위해 사용될 수 있다.

[0009] 절단은 엔지니어링된 징크 평거 뉴클레아제 (ZFN), 전사-활성화제 유사 이펙터 뉴클레아제 (TALEN)과 같은 특이적인 뉴클레아제의 사용을 통해, 또는 특이적인 절단을 안내하기 위해 엔지니어링된 crRNA/tracr RNA ('단일 가이드 RNA')를 가지는 CRISPR/Cas 시스템을 사용하여 일어날 수 있다. 예컨대 미국 특허 번호 8,697,359 및 8,932,814 및 미국 특허 공보 번호 20150056705를 참조한다. 나아가, 표적화된 뉴클레아제는 Argonaute 시스템을 기반으로 발달되었고 (예컨대 *T. thermophilus*로부터 유래하고, 'TtAgo'로서 알려짐, Swarts *et al* (2014) *Nature* 507(7491):258-261 참조), 게놈 편집 및 유전자 요법에 사용하기 위한 가능성을 가질 수 있다.

[0010] 상기 언급된 뉴클레아제 시스템 중 하나를 사용하는 표적화된 절단은 HDR 또는 NHEJ-매개된 과정 중 어느 하나를 사용하여 특이적인 표적 위치 안으로 핵산을 삽입시키기 위해 사용될 수 있다. 그러나, 뉴클레아제 시스템과 도너 둘 다를 세포에 전달하는 것은 문제의 소지가 될 수 있다. 예를 들어, 플라스미드의 세포 안으로의 형질도입을 통한 도너 또는 뉴클레아제의 전달은 수용체 세포, 특히 일차 세포인 세포에 독성이 될 수 있고 그러므로 세포주로부터의 세포와 같이 왕성하진 않다.

[0011] CD34+ 줄기 또는 선구 세포는 림프종 혈통의 세포 (예컨대 T 세포, B 세포, NK 세포) 및 골수종 혈통의 세포 (예컨대 단핵세포, 적혈구, 호산성 세포, 호염기성 세포 및 호중구)로 개신 및/또는 분화하는 능력을 특징으로 하는 세포의 이종성 세트이다. 그것들의 이종성 성질은 CD34+ 줄기세포 집단 내에, 자주 특이적인 그룹의 다능성 (혈통이 어디에 전용되든간에)을 반영하는 다수의 하위그룹이 있다는 사실로부터 발생한다. 예를 들어, CD38-인 CD34+ 세포는 보다 원시적이고, 미성숙한 CD34+ 선구 세포 (또한 장기간 조혈 선구체로도 언급됨)인 한편, CD34+CD38+인 것들 (단기간 조혈 선구체)은 혈통 전용이다 (Stella *et al* (1995) *Hematologica* 80:367-387 참조). 다음에 이 집단이 추가로 분화 경로를 다운 진행시킬 때, CD34 마커는 손실된다. CD34+ 줄기세포는 임상 세포 요법에서 많은 가능성을 나타낸다.

[0012] 적색 혈액 세포 (RBC), 또는 적혈구는 혈액의 주요 세포 성분이다. 실제로, RBC는 인간에서 세포의 1/4을 차지 한다. 성숙한 RBC는 인간에서 핵 및 많은 다른 소기관이 결핍되어 있고, 조직 밖으로 이산화탄소를 운반하고 재거를 위해 폐로 되돌리는 것뿐만 아니라 조직에 산소를 운반하는 기능을 하는, RBC에서 발견되는 금속단백질인 헤모글로빈으로 가득하다. 그 단백질은 RBC의 건조 중량의 대략 97%를 구성하고 혈액의 산소 운반 능력을 약 70 배 증가시킨다. 헤모글로빈은 2개의 α -형 글로빈 사슬과 2개의 β -형 글로빈 사슬 및 4개의 헴 그룹을 포함하는 이종사량체이다. 성인에서 $\alpha_2\beta_2$ 사량체는 헤모글로빈 A (HbA) 또는 성인 헤모글로빈으로 언급된다. 전형적으로, 알파 및 베타 글로빈 사슬은 대략 1:1 비율로 합성되고 이 비율은 헤모글로빈 및 RBC 안정화의 관점에서 결정적인 것으로 보인다. 실제로, 한 유형의 글로빈 유전자가 부적절하게 발현된 일부 경우에 (하기 참조), 다른 유형의 글로빈의 감소하는 발현 (예컨대 특이적 siRNA를 사용함)은, 이 1:1 비율을 되돌리면서 돌연변이체 세포 표현형의 일부 측면을 완화시킨다 (Voon *et al* (2008) *Haematologica* 93(8):1288 참조). 발생기 태아에서, 헤모글로빈 A보다 산소에 대한 더 큰 결합 친화성을 가져서 산소가 산모의 혈류를 통해 아기의 시스템으로 전달될 수 있는, 헤모글로빈의 상이한 형태인 태아 헤모글로빈 (HbF)이 생성된다. 태아 헤모글로빈은 또한 2개의 α 글로빈 사슬을 함유하지만, 성인 β -글로빈 사슬 대신 2개의 태아 γ -글로빈 사슬을 함유한다 (즉 태아 헤모글로빈은 $\alpha_2\gamma_2$ 이다). 대략 임신 30주에는, 태아에서의 γ 글로빈의 합성이 떨어지기 시작하는 한편 β 글로빈의 생성이 증가한다. 대략 10개월이 되면, 일부 (총 헤모글로빈의 대략 1 내지 3%) HbF가 성인기로 계속되긴 하지만 신생아의 헤모글로빈은 거의 전부 $\alpha_2\beta_2$ 가 된다. γ 의 β 로의 생성으로부터의 스위치의 조절은 꽤 복잡하고, 주로 γ 글로빈의 발현상의 하향-조절과 동시에 β 글로빈 발현의 상향-조절을 포함한다.

[0013] 헤모글로빈 사슬을 코드화하는 서열의 유전자는 결핍은 겸상 적혈구 빈혈 및 지중해 빈혈을 포함하여, 헤모글로빈 혈증으로 알려져 있는 많은 질환에 기여할 수 있다. 헤모글로빈 혈증을 가진 환자들의 대부분에서, γ 글로빈을 코드화하는 유전자가 여전히 존재하지만, 발현은 상기 기술된 것과 같이 출산 즈음에 발생하는 정상 유전자 억압으로 인해 상대적으로 낮다.

[0014] 미국에서 5000명 중 1명이 겸상 적혈구 질환 (SCD)을 가지고 있는 것으로 추산되는데, 그들 대부분이 사하라 사막 이남의 아프리카 태생이다. 말라리아에 대한 보호를 위한 겸상 적혈구 이형접합성의 장점인 것으로 나타나고, 그로써 이런 특성은 시간이 경과함에 따라 선택될 수 있어서, 사하라 사막 이남의 아프리카에서는, 인구의 1/3이 겸상 적혈구 특성을 가는 것으로 추정된다. 겸상 적혈구 질환은 발린이 아미노산 #6에서 글루탐산에 대해 치환되어 있는 (DNA 수준에서는 GTG에 대해 GAG) β -글로빈 유전자에서의 돌연변이에 의해 유발되고, 그 결과의 헤모글로빈은 "헤모글로빈 S" 또는 "HbS"로 언급된다. 저산소 조건하에서, HbS의 데옥시 형태에서 형태상의 시프트는 E와 F 나선 사이의 단백질 상의 소수성 폐치를 노출시킨다. 헤모글로빈의 베타 사슬의 위치 6

에서 발린의 소수성 잔기는 소수성 패치와 회합할 수 있어서 HbS 분자의 응집 및 섬유상 침전의 형성을 유발한다. 이들 응집체는 계속해서 RBC의 비정상 또는 '겸상 세포화'를 유발하여 세포의 가요성의 손실을 초래한다. 겸상 세포화된 RBC는 더 이상 모세혈관계로 들어갈 수 없고 겸상 적혈구 환자에게서 혈관-폐색 위기를 초래할 수 있다. 또한, 겸상 적혈구 RBC는 정상 RBC보다 더 깨지기 쉽고, 용혈되는 경향이 있으며, 궁극적으로 환자에서 빈혈을 유발한다.

[0015] 겸상 적혈구 환자의 치료 및 관리는 항생물질 치료, 통증 관리 및 급성 에피소드 중의 주입을 포함하는 평생 과제이다. 한 가지 접근법은 γ 글로빈의 생성을 증가시킴으로써 효과를 부분적으로 나타내는 하이드록시우레아의 사용이다. 만성 하이드록시우레아 치료법의 장기 부작용은 여전히 알려져 있지 않지만, 치료는 원치 않은 부작용을 유발하고 환자마다 가변적인 효능을 나타낼 수 있다. 겸상 적혈구 치료의 효능의 증가에도 불구하고, 환자들의 기대수명은 여전히 50대 중반부터 후반에 머물러 있고 질환의 관련된 이병률은 환자의 삶의 질에 지대한 영향을 미친다.

발명의 내용

해결하려는 과제

[0016] 그러므로, 게놈 편집을 위해, 일탈 유전자를 교정하거나 다른 유전자들의 발현을 변경하기 위해, 예를 들면 겸상 적혈구 질환과 같은 혜모글로빈증을 치료하기 위해 사용될 수 있는 추가의 방법 및 조성물에 대한 요구가 여전히 존재한다.

과제의 해결 수단

[0017] 본원에는 겸상 적혈구 질환과 같은 유전자 질환 (예컨대 질환에서 부족하거나, 결핍되거나 또는 일탈된 단백질 및/또는 그런 단백질을 조절하는 단백질을 생성하는)에 포함된 단백질을 코드화하는 하나 이상의 유전자를 교정하거나 유전자의 발현을 변경하기 위한 방법 및 조성물이 개시된다. 본 발명은 유전자 요법 및 게놈 엔지니어링에 사용하기 위한 조성물 및 방법을 기술한다.

[0018] 그런 단백질의 변경은 이들 유전자 질환의 치료를 초래할 수 있다. 특히, 게놈 편집은 일탈 유전자를 교정하거나, 야생형 유전자를 삽입하거나, 또는 내인성 유전자의 발현을 변화시키기 위해 사용된다. 한 가지 접근법은 잘못된 내인성 β 글로빈 유전자가 표적화되고 돌연변이 서열이 대체되는 유전자 교정의 사용을 포함한다. 한 가지 접근법은 추가로 줄기세포 (예컨대 조혈 줄기세포 또는 RBC 전구체)의 변형의 사용을 포함하는데, 그 줄기세포는 다음에 환자로의 이식을 위해, 혜모글로빈증의 치료를 위해 사용될 수 있다.

[0019] 한 측면으로, 본원에는 예를 들면 세포 또는 세포주의 동일 유형의 야생형 게놈 서열과 비교하여 유전자 변형된 세포 또는 세포주가 기술된다. 특정 구체예에서, 세포는 유전자 변형된 RBC 전구체 ("HSC"로 알려져 있는 조혈 줄기세포)를 포함한다. 세포 또는 세포주는 변형에 대해 이형접합성이거나 동형접합성일 수 있다. 변형은 삽입, 결실 및/또는 그것들의 조합을 포함할 수 있다. 특정 구체예에서, HSC는 엔지니어링된 뉴클레아제 및 도너 핵산으로 변형되어서 야생형 유전자 (예컨대 글로빈 유전자)가 삽입되고 발현되며 및/또는 내인성 일탈 유전자가 교정된다. 특정 구체예에서, 변형 (예컨대 삽입)은, 제한은 아니지만, 절단 및/결합 부위(들)의 상류 또는 하류의 1 내지 300 (또는 그 사이의 임의의 수의 염기쌍) 염기쌍 내에 있는 서열에 대한 변형; 결합 및/또는 절단 부위(들)의 어느 한 부위의 1 내지 100 염기쌍 (또는 그 사이의 임의의 수의 염기쌍) 내에 있는 변형; 결합 및/또는 절단 부위(들)의 어느 한 측에 있는 1 내지 50 염기쌍 (또는 그 사이의 임의의 수의 염기쌍) 내에 있는 변형; 및/또는 뉴클레아제 결합 및/또는 절단 부위의 하나 이상의 염기쌍에 대한 변형을 포함하여, 뉴클레아제 (들) 결합 및/또는 절단 부위(들)에 또는 가까이에 있다. 특정 구체예에서, 변형은 SEQ ID NO:23 또는 24에서 또는 가까이에서 (예컨대 1 내지 300 염기쌍 또는 그 사이에 있는 임의의 염기쌍 수) 일어난다. 다른 구체예에서, 변형은 SEQ ID NO:23 또는 24의 1 내지 100 염기쌍에서 (또는 그 사이에 있는 임의의 염기쌍 수) 일어난다. 특정 구체예에서, 변형은 SEQ ID NO:23 및/또는 SEQ ID NO:24 내에서 일어나고, 예를 들면 SEQ ID NO:23 또는 24중 어느 하나에서 하나 이상의 염기쌍이 변형된다. 일부 경우에, 삽입을 위해 야생형 유전자 서열이 야생형 β 글로빈을 코드화한다. 다른 경우에, 내인성 일탈 유전자는 β 글로빈 유전자, 예를 들면 내인성 일탈 인간 베타-혜모글로빈 (Hbb) 유전자의 적어도 하나의 돌연변이를 교정하는 하나 이상의 게놈 변형이다. 일부 측면으로, 베타 글로빈 유전자의 변형은 전사된 RNA의 적절한 스플라이싱을 허용한다. 다른 측면으로, 베타 글로빈 유전자의 변형은 겸상 대립형질의 돌연변이를 교정하여서 폴리펩티드 서열의 6번째 위치에 있는 부정확한 발린 아미노산을 코딩하는 코돈 GTG가 야생형 폴리펩티드에서 발견되는 것과 같은 글루탐산을 코드화하는 GAG로

바뀐다. 본원에서 기술되는 것과 같은 유전자 변형된 줄기세포로부터 파생된 부분적으로 또는 전체적으로 분화된 세포들 또한 제공된다 (예컨대 RBS 또는 RBC 전구체 세포). 본원에서 기술되는 유전자 변형된 세포를 포함하는 약학 조성물과 같은 조성물 또한 제공된다.

[0020] 다른 측면으로, 본원에는 게놈의 관심의 영역 (예컨대 β 글로빈 유전자)에서 표적 부위에 결합하는 징크-핑거 단백질 (ZEP)이 기술되는데, ZEP는 하나 이상의 엔지니어링된 징크-핑거 결합 도메인을 포함한다. 한 구체예에서, ZEP는 관심의 표적 게놈 영역을 절단하는 징크-핑거 뉴클레아제 (ZFN)이고, ZFN은 하나 이상의 엔지니어링된 징크-핑거 결합 도메인 및 뉴클레아제 절단 도메인 또는 절단 하프-도메인을 포함한다. 절단 도메인 및 절단 하프 도메인은 예를 들면, 다양한 제한 엔도뉴클레아제 및/또는 호밍 (homing) 엔도뉴클레아제로부터 얻어질 수 있다. 한 구체예에서, 절단 하프-도메인은 타입 IIS 제한 엔도뉴클레아제 (예컨대 *Fok I*)로부터 유도된다. 특정 구체예에서, 징크 펑거 도메인은 글로빈 또는 안전한 은닉 (harbor) 유전자에서 표적 부위를 인식한다. 특정 구체예에서, 징크 펑거 도메인은 5 또는 6 징크 펑거 도메인을 포함하고 글로빈 유전자에서 표적 부위를 인식한다 (예컨대 표 1에 나타낸 인식 나선 영역을 가지는 5 또는 6 펑거를 가지는 징크 펑거 단백질).

[0021] 다른 측면으로, 본원에는 게놈의 관심 영역 (예컨대 β 글로빈 유전자)에서 표적 부위에 결합하는 TALE 단백질이 기술되는데, TALE은 하나 이상의 엔지니어링된 TALE 결합 도메인을 포함한다. 한 구체예에서, TALE은 관심의 표적 게놈 영역을 절단하는 뉴클레아제 (TALEN)이고, TALEN은 하나 이상의 엔지니어링된 TALE DNA 결합 도메인 및 뉴클레아제 절단 도메인 또는 절단 하프-도메인을 포함한다. 절단 도메인 및 절단 하프 도메인은, 예를 들면 다양한 제한 엔도뉴클레아제 및/또는 호밍 엔도뉴클레아제로부터 얻어질 수 있다. 한 구체예에서, 절단 하프-도메인은 타입 IIS 제한 엔도뉴클레아제 (예컨대 *Fok I*)로부터 유도된다.

[0022] 또 다른 측면으로, 본원에는 게놈의 관심 영역 (예컨대 질환 관련 유전자)의 표적 부위에 결합하는 CRISPR/Cas 또는 tAgo 시스템이 기술되며, CRISPR/Cas 시스템은 CRIPSR/Cas 뉴클레아제 및 엔지니어링된 crRNA/tracrRNA (또는 단일 가이드 RNA)를 포함한다. 특정 구체예에서, CRISPR/Cas 또는 tAgo 시스템은 글로빈 유전자에서 표적을 인식한다.

[0023] 일부 구체예에서, 세포의 게놈 엔지니어링을 촉진하는 뉴클레아제는 웹티드로서 전달되는 한편, 다른 구체예에서 그것은 뉴클레아제를 코드화하는 핵산으로서 전달된다. 일부 구체예에서, 하나 이상의 뉴클레아제가 사용되고 핵산 형태, 단백질 형태 또는 그것들의 조합으로 전달될 수 있다. 일부 바람직한 구체예에서, 뉴클레아제를 코드화하는 핵산(들)은 mRNA이고, 일부 경우에, mRNA는 보호된다. 추가의 바람직한 구체예에서, mRNA는 ARCA cap을 포함할 수 있고 및/또는 변형 및 미변형 뉴클레오티드의 혼합물을 포함할 수 있다. 뉴클레아제는 징크 펑거 뉴클레아제 (ZFN), TALE-뉴클레아제 (TALEN), TtAgo 또는 CRISPR/Cas 뉴클레아제 시스템 또는 그것들의 조합을 포함할 수 있다. 바람직한 구체예에서, 뉴클레아제(들)을 코드화하는 핵산은 전기천공법을 통해 전달된다. 다른 측면으로, 본원에는 세포에서 뉴클레아제-매개된 절단 후에 (예컨대 HDR을 통해) 표적화된 통합을 증가시키는 방법이 기술된다. 특정 구체예에서, 방법은 (i) 하나 이상의 뉴클레아제 (및/또는 mRNA 또는 그 뉴클레아제(들) 및 필요하다면 하나 이상의 단일 가이드 RNA를 발현하는 발현 구성물)을 하나 이상의 도너 분자와 함께 숙주 세포 안에 도입하는 단계 및 (ii) 줄기성의 손실 또는 세포 및/또는 그 세포를 함유하는 배양 배지 안으로의 분화 없이 줄기세포 팽창에 영향을 미치고 및/또는 증가시키는 하나 이상의 인자를 뉴클레아제(들)의 도입 전에, 중에 및/또는 후에 도입하는 단계를 포함한다. 도너는 뉴클레아제(들)을 코드화하는 핵산보다 앞서, 후에 또는 함께 전달될 수 있다.

[0024] 도너는 바이러스 및/또는 비-바이러스 유전자 전달 방법에 의해 전달될 수 있다. 다른 구체예에서, 도너는 아데노 관련 바이러스 (AAV)를 통해 세포에 전달된다. 일부 경우에, AAV는 캡시드 혈청형과 비교하여 이종성 혈청형의 것인 LTR을 포함한다. 다른 구체예에서, 도너는 렌티바이러스를 통해 세포에 전달된다. 일부 경우에, 렌티바이러스는 인테그라제 결핍 렌티바이러스 (IDLV)이다.

[0025] 특정 구체예에서, 도너 핵산은 비-상동성 의존 방법 (예컨대 NHEJ)을 통해 통합된다. 상기에서 주지된 것과 같이, 생체 내 절단시 도너 서열은 표적화된 방식으로 DSB의 위치에서 세포의 게놈 안으로 통합될 수 있다. 도너 서열은 DSB를 생성하기 위해 사용된 하나 이상의 뉴클레아제에 대해 하나 이상의 동일한 표적 부위를 포함할 수 있다. 그러므로, 도너 서열은 그 안으로의 통합이 바람직한 내인성 유전자를 절단하기 위해 사용된 하나 이상의 동일한 뉴클레아제에 의해 절단될 수 있다. 특정 구체예에서, 도너 서열은 DSB를 유도하기 위해 사용된 뉴클레아제로부터의 상이한 뉴클레아제 표적 부위들을 포함한다. 표적 세포의 게놈의 DSB는 임의의 메커니즘에 의해 생성될 수 있다. 특정 구체예에서, DSB는 관심의 영역 내의 서열에 결합하도록 엔지니어링된 징크 펑거 결합 도메인 및 절단 도메인 또는 절단 하프-도메인을 포함하는 융합 단백질인 하나 이상의 징크-핑거 뉴클레아제

(ZFNs)에 의해 생성된다. 다른 구체예에서, DSB는 뉴클레아제 도메인 (TALEN)에 융합된 하나 이상의 TALE DNA-결합 도메인 (자연적으로 발생하는 또는 자연적으로 발생하지 않는)에 의해 생성된다. 또 추가의 구체예에서, DSB는 엔지니어링된 단일 가이드 RNA 또는 그것이 기능적 동등물이 뉴클레아제를 계놈의 표적화된 부위로 안내하기 위해 사용되는 CRISPR/Cas 뉴클레아제 시스템을 사용하여 생성된다. 또 다른 추가의 구체예에서, DSB는 TtAgo 시스템을 사용하여 생성된다.

[0026] 또 다른 측면으로, 발명은 세포 계놈의 뉴클레아제-매개된 절단 후에 유전자 파괴 및/또는 표적화된 통합 (예컨대 유전자 교정)을 증가시키기에 유용한 키트를 제공한다 (예컨대 ZFNs, TAL-이팩터 도메인 뉴클레아제 융합 단백질, TtAgo 시스템, 또는 CRISPR/Cas 시스템으로 엔지니어링된 호밍 엔도뉴클레아제 또는 엔지니어링된 가이드 RNA). 키트는 전형적으로 표적 부위에 결합하는 하나 이상의 뉴클레아제, 줄기세포 팽창 및/또는 분화에 영향을 미치는 하나 이상의 인자 및 뉴클레아제 및 줄기세포-영향 인자들을 세포에 도입시키기 위한 지시를 포함하여서 뉴클레아제-매개된 유전자 파괴 및/또는 표적화된 통합이 증대된다. 선택적으로, 뉴클레아제의 표적 부위(들)을 함유하는 세포 또한 본원에 기술된 키트에 포함될 수 있다. 특정 구체예에서, 키트는 표적 유전자 및 표적 유전자 내에서 절단할 수 있는 공지의 뉴클레아제를 가지는 적어도 하나의 구성물을 포함한다. 그런 키트는 다양한 달라지는 숙주 세포 유형에서 절단 조건의 최적화에 유용하다. 발명에 의해 고려된 다른 키트는 계놈 내의 공지의 표적 유전자와 내에서 절단할 수 있는 공지의 뉴클레아제를 포함할 수 있고, 추가로 도너 핵산을 포함할 수 있다. 일부 측면으로, 도너 DNA는 폴리펩티드, 조절 영역 또는 구조적 핵산을 코드화할 수 있다. 일부 구체예에서, 폴리펩티드는 리포터 유전자 (예컨대 GFP)이다. 그런 키트는 도너 통합을 위한 또는 유전자 파괴 또는 표적화된 삽입을 함유하고 있는 특이적으로 변형된 세포, 세포주 및 동물의 구성을 위한 최적화에 유용하다.

[0027] 다른 측면으로, 본원에서 기술되는 유전자 변형된 줄기세포를 대상에게 투여하는 방법이 기술된다. 유전자 변형된 혈구 전구체 ("HSC/PC")는 전형적으로 골수 이식으로 얻어지고 HSC/PC는 생체 내에서 분화하고 성숙한다. 일부 구체예에서, HSC/PC는 G-CSF-유도된 고정에 이어서 분리되고, 다른 구체예에서는 세포는 인간 골수 또는 텫 줄로부터 분리된다. 일부 측면으로, HSC/PC는 특이적 유전자 또는 조절 서열을 녹아웃하도록 디자인된 뉴클레아제로의 처리에 의해 편집된다. 다른 측면으로, HSC/PC는 엔지니어링된 뉴클레아제 및 도너 핵산으로 변형되어서 야생형 유전자 또는 관심의 다른 유전자가 삽입되고 발현되며 및/또는 내인성 일탈 유전자가 교정된다. 일부 구체예에서, 변형된 HSC/PC는 약한 골수소멸성 사전-조건화에 이어서 환자에게 투여된다. 다른 측면으로, HSC/PC는 전체 골수소멸 후에 투여되어서 이식 후에, 100%의 조혈세포가 변형된 HSC/PC로부터 유도된다. 나아가, 세포는 세포 주기의 G2 단계에서 정지될 수 있다.

[0028] 일부 구체예에서, 유전자 도입 HSC/PC 세포 및/또는 동물은 인간 유전자를 코드화하는 도입 유전자를 포함한다. 일부 경우에, 유전자 도입 동물은 외인성 도입 유전자에 상응하는 내인성 유전자와에서 녹아웃을 포함하고, 그로써 인간 단백질이 분리상태로 연구될 수 있는 생체 내 시스템의 개발이 허용된다. 그런 유전자 도입 모델은 관심의 인간 단백질과 상호작용하거나 그것을 변형시킬 수 있는 작은 분자 또는 큰 생체분자 또는 다른 실체를 확인하기 위한 스크리닝 목적에 대해 사용될 수 있다. 일부 측면으로, 도입 유전자는 선택된 유전자와 안으로 (예컨대 안전한-은닉), 줄기세포 (예컨대 배아 줄기 세포, 유도된 전능 줄기세포, 조혈 줄기세포 등) 안으로 또는 본원에 기술된 방법들 중 임의의 방법에 의해 얻어진 동물 배아 안으로 통합된 후, 배아는 이식되어서 살아 있는 동물이 태어난다. 그런 다음 동물은 성적으로 성숙해지고 적어도 자손의 일부가 편집된 내인성 유전자 서열 또는 통합된 도입 유전자를 포함하는 자손을 생산하도록 허용된다.

[0029] 이들 및 다른 측면들은 전체 발명을 고려하여 숙련자들에게 쉽게 드러날 것이다.

도면의 간단한 설명

[0030] **도 1A** 내지 **1E**는 IDLV 도너를 사용하는 CD34+ 세포의 베타-글로빈 유전자와에서의 절단 및 교정을 도시한다. **도 1A** (SEQ ID NO:3)는 출발 코돈 (진한 색) 및 겹상 돌연변이 (진한 색, 이탈릭체) 상부의 ZFN 표적 부위 (밑줄)를 보여주는 베타-글로빈의 엑손 I DNA 서열의 일부를 도시한다. **도 1B**는 인간 제대혈 CD34+ 세포의 베타-글로빈 유전자와에서의 표적화된 절단을 보여주는 겜이다. 세포들은 ZFN을 코드화하는 시험관 내 전사된 mRNA를 사용한 전기천공 후 3일 후 분석되었다. 'Mock'은 미처리 CD34+ 세포를 나타낸다. 화살표는 PCR 증폭 및 Surveyor® 뉴클레아제 (Transgenomic®)로의 소화 후의 절단 밴드를 나타낸다. **도 1C**는 엑손 1 및 2와 관련하여 베타 글로빈 유전자의 겹상 돌연변이에서의 부위-특이적 유전자 교정의 개략도이다. 또한 도너 구성물 및 ZFN에 의한 절단 및 HDR에 의한 수복시 그 결과의 계놈 DNA가 상세하게 예시된다. 겹상 적혈구 돌연변이 및 *Hha*I RFLP의 위치 (별표)가 표시된다. Surveyor 뉴클레아제에 대한 프라이머 위치 및 RFLP 분석이 아래에 도시된다. 일정 비율의 엑손: 비스듬한 해시 마크를 일정 비율이 아닌 구역을 나타낸다. **도 1D**는 ZFN 및 도너의 존재하에 베타-글로

빈의 표적화된 유전자 변형에 대한 RELP 젤이다. 제대혈 (CB) CD34+ 세포는 젤의 상부에 표시된 바와 같이 시험관 내 전사된 ZFN mRNA로 전기천공되었고 및/또는 IDLV를 포함하는 도너로 형질도입되었다. 세포들은 처리 후 4일 후에 수확되었고, 도너 영역 외부로부터 PCR 증폭되었으며, *HhaI* 효소로 소화되었고, 아가로오스 젤 상에서 분석되었다. 화살표는 절단된 생성물을 보여주고, 표적 부위에서 게놈 안으로의 RFLP의 통합을 나타낸다. **도 1E**는 표시된 조건하에서 CD34+ 세포에서의 유전자 변형 백분율을 나타내는 그래프이다. CB CD34+ 세포는 시험관 내 전사된 ZFN mRNA로 전기천공되었고 도너 IDLV로 형질도입되었다. 세포들은 처리 후 3일 후에 수확되었고, 도너 영역 외부로부터 PCR 증폭되었으며 qPCR이 *HhaI* RFLP를 생성하는 사일런트 염기 변화의 통합을 특이적으로 검출하기 위해 디자인된 프라이머들로 완료되었고 암플리콘의 베타-글로빈 유전자좌의 엑손 2에서 결합하는 프라이머들에 대해 표준화되었다, (모든 조건에 대해 n=4). 유전자 변형은 ZFN 및 도너 둘 다로 처리된 세포의 대략 18%에서 발견되었다.

도 2A 내지 2D는 형질도입에 대해 사용된 ZFN 및 IDLV 용량의 최적화의 결과를 도시한다. **도 2A, B 및 C**는 ZFN mRNA의 적정 후 결과를 보여주는 그래프이다. CD34+ 세포가 증가하는 양의 시험관 내 전사된 ZFN mRNA로 전기천공되었고 일정한 양의 도너 IDLV로 형질도입되었다. Mock는 미처리 세포를 나타내며, ZFN 단독 샘플은 30-30ug/mL의 각각의 ZFN으로 전기천공되었고, IDLV 단독 샘플은 펄스되고 2E+07 TU/ml의 도너 IDLV로 형질도입되었다. **도 2A**는 RFLP에 대해 qPCR에 의해 측정되는 바 전기천공 후 3일째의 유전자 변형률을 보여준다. **도 2B**는 전기천공 후 1일째의 팽창 배수를 보여주고 **도 2C**는 트립판 축출 염료에 의해 측정되는 바 전기천공 후 1일째의 세포 생존율을 보여준다, (모든 조건에 대해 n=8). **도 2D**는 도너 인테그라제-결핍 렌티바이러스 벡터 (IDLV)의 개량도를 보여준다: 1.1 kb 영역의 베타-글로빈이 긴 말단 반복부 (PTR)의 프로모터 요소들에 대하여 역으로, 3' LTR이 프로바이러스 DNA의 5' LTR로 역전사되는 동안 전달되는 자기-비활성화 (SIN) 결실을 함유하고 있는 렌티바이러스 골격에 클론되었다. UTR: 미번역 영역, RRE: 역-반응 요소, Int: 인트론.

도 3A 내지 3C는 도너 IDLV의 적정을 보여주는 그래프이다. CD34+ 세포들은 일정량의 시험관 내 전사된 ZFN mRNA로 전기천공되었고 증가하는 양의 도너 IDLV로 형질도입되었다. Mock는 미처리 세포, 10 ug/mL의 각각의 ZFN으로 전기천공된 ZFN만의 샘플, 펄스되고 (pulsed) 6E+07 TU/ml의 도너 IDLV로 형질도입된 IDLV만의 샘플을 나타낸다. **도 3A**는 RFLP에 대한 qPCR에 의해 측정되는 바 전기천공 후 3일째의 유전자 변형률을 보여준다. **도 3B**는 전기천공 후 1일째의 팽창 배수를 보여주고 **도 3C**는 트립판 축출 염료에 의해 측정되는 바 전기천공 후 1일째의 세포 생존률을 보여준다, (모든 조건에 대해 n=6).

도 4A 내지 4F는 올리고뉴클레오티드 도너를 사용하여 CD34+ 세포에서 베타-글로빈 유전자좌에서 ZFN 쌍 33488/33501을 사용한 교정을 도시한다. **도 4A**는 올리고뉴클레오티드-지시된 유전자 변형을 개량적으로 보여준다 (SEQ ID NO:4 및 5). **도 4B**는 베타-글로빈 유전자좌의 *AvrII*-소화된 PCR 암플리콘의 젤이다. 단편은 천연 *AvrII* 부위를 함유하고, 그것의 절단은 *AvrII* 소화에 대한 내부 대조표준으로서 작용한다 (젤 상의 하부 밴드). 화살표는 *AvrII* 절단 생성물을 나타낸다. 올리고 도너 및 ZFN의 존재하에, 유전자 교정률은 15%이다. **도 4C**는 SBS#33501 ZFN 결합 부위에서 사일런트 돌연변이의 가능한 6개의 부위를 나타낸다 (WT에 대해 SEQ ID NO:53 및 SMS에 대해 SEQ ID NO:54). 겸상 돌연변이는 이탈릭체로, 가능한 사일런트 돌연변이 부위 (SMS)는 진하게 나타낸다. **도 4D**는 사일런트 돌연변이가 베타-글로빈에서 유전자 교정을 증가시켰음을 보여준다. mPB CD34+ 세포가 ZFN 및 표시된 도너 올리고뉴클레오티드로 트랜스펙션되었다. 관련 사일런트 돌연변이의 도입은 대용량 서열화를 통해 평가되었다. 흰색 막대는 삽입 및 결실 (indel)을 나타내고; 회색 막대는 유전자 교정을 나타낸다. **도 4E**는 사일런트 돌연변이가 ZFN 재-절단을 차단하는 것을 보여준다. Indel이 있는 대립유전자가 상동성-매개된 변형의 증거에 대해 조사되었다. NHEJ-유도된 indel의 증거를 또한 가지는 유전자 변형을 가진 대립유전자의 백분율이 도시된다. **도 4F**는 ZFN 농도의 최적화 및 도너 유형을 보여준다. NHEJ-유도 indel (흰색 막대) 및 유전자 교정 (회색 막대)은 대용량 서열화에 의해 평가되었다. 대용량 DNA 서열화의 깊이가 제공되면, 측정 오류는 매우 낮을 것으로 예상된다. WT: 야생형; SMS: 사일런트 돌연변이 부위; 올리고: 올리고뉴클레오티드.

도 5는 도너 길이 및 가타의 최적화를 도시한다. mPB CD34+ 세포가 ZFN과 3개의 염기쌍 서열을 베타-글로빈 유전자좌에 도입하도록 디자인된 올리고뉴클레오티드의 조합으로 트랜스펙션되었다. 세포가 수확되고 NHEJ-유도된 indel의 수준 및 이 짧은 서열의 표적화된 통합 ("TI")이 대용량 DNA 서열화 (HTS)에 의해 분석되었다. 샘플당 대략 4 내지 9,000 서열 판독을 얻었다.

도 6A 및 6B는 부분적인 HDR 사건을 도시하고 합성이 올리고-매개된 유전자 변형 중에 좌측에서 우측으로 진행되는 것을 나타낸다. **도 6A**는 부분적인 HDR 생성물 (SMS1, SMS12, SMS24, 및 SMS4 대립유전자)의 증거에 대해 분석된, ZFN과 도 3F로부터의 SMS124 도너로부터의 대용량 DNA 서열화 데이터를 보여준다. 이 실험에서 ~41%의 유전자 변형 중에 6%의 변형은 SMS12 사건으로부터 유래한 반면 SMS24 대립유전자는 거의 완전하게 부재하였다

(0.03%). 이 비대칭성은 DNA 합성이 좌측에서 우측으로 진행되는 것을 시사한다. 도 6B는 3개의 모든 ZFN 농도로부터의 데이터의 평균을 사용한 것이 도시되는데, 30 및 60 $\mu\text{g/mL}$ 의 샘플로부터의 사건이 과열 (Burst) 15 $\mu\text{g/mL}$ 의 샘플에서 볼 수 있는 ~41% 유전자 변형을 매칭하기 위해 상향 표준화되었다.

도 7A 내지 7D는 베타-글로빈 유전자 클러스터에서의 온-표적 절단 분석을 도시한다. 도 7A는 다음과 같이 GFP 포획을 예시한다: K562 적혈구성 백혈병 세포가 시험관 내 전사된 mRNA로 전기천공되고 GFP를 발현하는 IDLV 벡터로 형질도입되었다. 세포는 60일 동안 배양되어 모든 미-트랩핑된 GFP가 희석되었고, 그런 후 GFP 포지티브 세포가 형광-활성화 세포 분류 (FACS)를 사용하여 분류되었다. 샘플에 대해 nrLAM-PCR이 수행되었고 벡터 통합 부위 분석이 완료되었다. 클러스터 분석을 사용하여 측정된 베타-글로빈 및 엘타-글로빈에서의 대부분의 통상적인 통합 부위들이 도시된다. HBB: 야생형 인간 혜모글로빈 베타; HBD: 인간 혜모글로빈 엘타. 도 7B는 ZFN 절단 부위 주변의 베타-글로빈 유전자 클러스터에서 염색체 11 상의 부분적인 글로빈 유전자 서열의 일렬배열을 보여준다 (SEQ ID NO:6 내지 11). ZFN 결합 부위는 밀줄이 있고, 겸상 돌연변이는 굵은 이탤릭체이다. HBB: 야생형 인간 혜모글로빈 베타; HBD: 인간 혜모글로빈 엘타; HBE1: 인간 혜모글로빈 엡실론; HBBP1: 인간 혜모글로빈 베타 위유전자 1; HBG1: 인간 혜모글로빈 감마 A; HBG2: 인간 혜모글로빈 감마 G. 도 7C는 ZFN 시험관 내 전사된 mRNA로 처리된 또는 미처리된 (Mock) CD34+ 세포의 Surveyor® 뉴클레아제 분석을 보여주는 겔이다. 표시된 베타 글로빈 클러스터 유전자들의 각각은 표적 부위와 가장 높은 상동성의 영역 주변에서 증폭되었고 Surveyor® 뉴클레아제로 절단되었다. 밴드들은 밀도계에 의해 정량되었다. 도 7D는 도 7B에서와 같이 SCD 환자 골수 샘플의 Surveyor® 뉴클레아제 분석을 보여주는 겔이다. 감마-글로빈 유전자의 단일 뉴클레오티드 다형태에 대한 이 형접합체는 ZFN에 노출되지 않은 샘플에서도 배경 절단을 유발하였다 (가는 화살표).

도 8은 ZFN 및 올리고뉴클레오티드 변형된 HSPC의 콜로니-형성 가능성을 보여주는 그래프이다. mPB CD34+ 세포는 mock 트랜스펙션되거나, 10 $\mu\text{g/mL}$ ZFN으로 트랜스펙션되거나, 또는 ZFN 및 3 μM 의 올리고뉴클레오티드 도너로 트랜스펙션되었다. 배양 후 5일 후, 세포는 메틸셀룰로오스 및 사이토카인-함유 배지에 플레이팅되었고 분화되었다. 15일 동안 성장된 후 콜로니-형성 유닛 (CFU)이 카운팅되었고 형태학적으로 적혈구 (E), 과열-형성 유닛 - 적혈구 (BFU-E), 과립구/마크로파지 (GM) 또는 과립구/적혈구/마크로파지/거대핵세포 (GEMM) 중 어느 하나로 분류되었다.

도 9A 및 9B는 적혈구 분화 중에 변형된 세포의 안정성을 도시한다. 도 9A는 mPB CD34+ 세포가 0, 3.75, 7.5, 15 또는 30 $\mu\text{g/mL}$ 의 ZFN 및 3 μM 의 올리고뉴클레오티드 도너로 트랜스펙션되거나 또는 올리고뉴클레오티드 도너 단독으로 트랜스펙션되었음을 보여준다. 세포는 적혈구로 분화하도록 유도되었고 0, 6, 12, 15, 및 18일 후에 부분표본이 제거되었다. 계놈 DNA가 세포로부터 제조되었고 유전자-변형된 세포의 빈도가 대용량 서열화를 통해 분석되었다. 대략 10 내지 50,000 서열 판독이 샘플당 시점당 얻어졌다. 도 9B는 엔지니어링된 세포에서 글로빈 단백질의 HPLC 분석을 보여준다.

도 10A 내지 10E는 ZFN 및 도너-처리된 세포의 NSG 마우스 안으로의 이식을 도시한다. 도 10A는 전기천공 후 7 일째에 RFLP에 대한 qPCR에 의해 측정되는 바 ZFN+IDLV로 처리되고 시험관 내에서 배양된 벌크 이식된 세포의 유전자 변형률을 보여준다. Mock 세포는 미처리된다 (n=3 독립적인 실험). 도 10B는 ZFN+IDLV 변형된 CD34+ 세포에 대한 대용량 서열화에 의해 평가된 겸상 염기에서의 변형을 보여주는 그래프이다. 베타-글로빈 유전자좌의 서열화의 결과는 절단 부위에서의 삽입 및 결실 (indel)뿐 아니라 겸상 염기 (T)로 변화된 야생형을 함유하는 배열된 총 판독의 백분율을 보여준다. 샘플은 도 10A에서와 같다. 변화된 염기, 백색; indel, 회색. 도 10C는 이식 후 5주 및 8주째에 이식된 마우스의 말초혈에서의 이식을 보여주는 그래프이다. 인간 이식은 mock- 또는 ZFN+IDLV-처리된 세포 중 어느 하나를 받은 마우스로부터의 세포의 유동 세포분석에 의해 총 hCD45+ 및 mCD45+ 세포 중 hCD45+ 세포의 백분율로서 측정되었다. Mock, 흰색 다이아몬드; ZFN+IDLV, 검은색 다이아몬드. (n=3 독립적인 실험; mock n=5, ZFN+IDLV n=12). 도 10D는 CD34+ 세포가 올리고, ZFN 또는 ZFN+올리고로 전기천공되고 시험관 내 배양된 후 NSG 마우스로 이식된 것을 보여주는 그래프이다. 겸상 염기에서의 변형률 (n=1) 및 indel이 도 10C에서와 같이 도시된다. 도 10E는 도 10C에서와 같이 말초혈에서, 올리고-, ZFN- 또는 ZFN+올리고-처리된 세포 중 어느 하나를 받은 마우스로부터의 세포의 이식을 보여주는 그래프이다. 올리고, 원형; ZFN, 삼각형; ZFN+올리고, 다이아몬드. (올리고 n=8, ZFN n=7, ZFN+올리고 n=9), n.s.: 유의미하지 않음, * p<0.05, **p<0.01.

도 11A 및 11B는 8주째에 이식된 NSG 마우스의 혈통 분석을 도시한다. mock 및 ZFN 및 도너 처리된 세포로 이식된 NSG 마우스의 말초혈의 이식후 8주째의 면역표현형 분석. B-세포 (CD19), T-세포 (CD3), 조혈 선구체 (CD34), 골수 선구체 (CD33), 및 천연 킬러 세포 (CD56)의 마커에 대해 포지티브인 인간 CD45+ 세포의 백분율이 유동 세포분석을 사용하여 나열되었다. 도 11A는 mock- 또는 ZFN+IDLV-처리된 세포로 이식된 마우스로부터의 세

포를 보여주는 그래프이다, (n=3 독립적 실험; mock n=5, ZFN+IDLV n=12). **도 11B**는 올리고-, ZFN- 또는 ZFN+올리고-처리된 세포로 이식된 마우스로부터의 세포를 보여주는 그래프이다, (올리고 n=8, ZFN n=7, ZFN+올리고 n=9). n.s.: 유의미하지 않음; 별표는 유의미성을 나타낸다, * p<0.05, **p<0.01.

도 12A 내지 12D는 최종 말초혈 이식 및 혈통 분석을 도시한다. **도 12A**는 이식후 16주째에 이식된 마우스의 말초혈에서의 이식을 보여주는 그래프이다. 인간 이식은 mock- 또는 ZFN+IDLV-처리된 세포 중 어느 하나를 받은 마우스로부터의 세포의 유동 세포분석에 의해 총 hCD45+ 및 mCD45+ 세포 중 hCD45+ 세포의 백분율로서 측정되었다. Mock, 흰색 다이아몬드; ZFN+IDLV, 검은색 다이아몬드. **도 12B**는 이식후 16주째에 이식된 NSG 마우스의 말초혈의 면역표현형 분석을 보여주는 그래프이다. B-세포 (CD19), T-세포 (CD3), 조혈 선구체 (CD34), 골수 선구체 (CD33), 및 천연 킬러 세포 (CD56)의 마커에 대해 포지티브인 hCD45+ 세포의 백분율이 mock- 또는 ZFN+IDLV-처리된 세포 중 어느 하나를 받은 마우스로부터의 세포의 유동 세포분석을 사용하여 나열되었다. Mock, 흰색 다이아몬드; ZFN+IDLV, 검은색 다이아몬드; (n=3 독립적인 실험; mock n=5, ZFN+IDLV n=12). **도 12C**는 **도 12A**에서와 같이 올리고-, ZFN- 또는 ZFN+올리고-처리된 세포 중 어느 하나를 받은 마우스로부터의 세포의 말초혈에서의 이식을 보여주는 그래프이다. 올리고, 원형; ZFN, 삼각형; ZFN+올리고, 다이아몬드. **도 12D**는 **도 12B**에서와 같이 올리고-, ZFN- 또는 ZFN+올리고-처리된 세포 중 어느 하나를 받은 마우스로부터의 세포의 말초혈의 혈통 분석을 보여주는 그래프이다. 올리고, 원형; ZFN, 삼각형; ZFN+올리고, 다이아몬드; (올리고 n=8, ZFN n=7, ZFN+올리고 n=9). n.s.: 유의미하지 않음, * p<0.05.

도 13A 내지 13C는 유전자 변형된 세포가 NSG 마우스에서 지속적인 것을 보여준다. **도 13A**는 mock- 또는 ZFN+IDLV-처리된 세포 중 어느 하나를 받은 마우스로부터의 세포에서 16주째에 이식된 마우스의 골수 및 비장에서의 유전자 변형률을 보여주는 그래프이다. 조직이 수확되었고, 계놈 DNA가 추출되고 증폭되었으며, 통합된 RFLP에 대해 qPCR로 분석되었다. Mock, 흰색 다이아몬드; ZFN+IDLV, 검은색 다이아몬드. **도 13B**는 **도 13A**에서 기술된 샘플의 대용량 서열화에 의해 평가된 겹상 돌연변이에서의 교정을 보여주는 그래프이다. 베타-글로빈 유전자좌의 서열화의 결과는 겹상 돌연변이에서의 변형된 염기를 함유하는 배열된 총 판독의 백분율을 보여준다. Mock, 흰색 다이아몬드; ZFN+IDLV, 검은색 다이아몬드. (n=3 독립적인 실험; mock n=5, ZFN+IDLV n=12). **도 13C**는 **도 13B**에서 기술된 것과 같이 올리고-, ZFN- 또는 ZFN+올리고-처리된 세포 중 어느 하나를 받은 마우스로부터의 세포에서 겹상 돌연변이에서의 교정을 보여주는 그래프이다. 올리고, 원형; ZFN, 삼각형; ZFN+올리고, 다이아몬드. (올리고 n=8, ZFN n=7, ZFN+올리고 n=9). n.s.: 유의미하지 않음; 별표는 유의미성을 나타낸다, * p<0.05, **p<0.01.

도 14A 내지 14D는 최종 골수 이식 및 혈통 분석을 도시한다. **도 14A**는 이식후 16주째에 이식된 마우스의 골수에서의 이식을 보여주는 그래프이다. 인간 이식은 mock- 또는 ZFN+IDLV-처리된 세포 중 어느 하나를 받은 마우스로부터의 세포의 유동 세포분석에 의해 총 hCD45+ 및 mCD45+ 세포 중 hCD45+ 세포의 백분율로서 측정되었다. Mock, 흰색 다이아몬드; ZFN+IDLV, 검은색 다이아몬드. **도 14B**는 이식후 16주째에 이식된 NSG 마우스의 말초혈의 면역표현형 분석을 보여주는 그래프이다. B-세포 (CD19), T-세포 (CD3), 조혈 선구체 (CD34), 골수 선구체 (CD33), 및 천연 킬러 세포 (CD56)의 마커에 대해 포지티브인 hCD45+ 세포의 백분율이 mock- 또는 ZFN+IDLV-처리된 세포 중 어느 하나를 받은 마우스로부터의 세포의 유동 세포분석을 사용하여 나열되었다. Mock, 흰색 다이아몬드; ZFN+IDLV, 검은색 다이아몬드; (n=3 독립적인 실험; mock n=5, ZFN+IDLV n=12). **도 14C**는 **도 14A**에서와 같이 올리고-, ZFN- 또는 ZFN+올리고-처리된 세포 중 어느 하나를 받은 마우스로부터의 세포의 골수에서의 이식을 보여주는 그래프이다. 올리고, 원형; ZFN, 삼각형; ZFN+올리고, 다이아몬드. **도 14D**는 **도 14B**에서와 같이 올리고-, ZFN- 또는 ZFN+올리고-처리된 세포 중 어느 하나를 받은 마우스로부터의 세포의 골수의 혈통 분석을 보여주는 그래프이다. 올리고, 원형; ZFN, 삼각형; ZFN+올리고, 다이아몬드; (올리고 n=8, ZFN n=7, ZFN+올리고 n=9). n.s.: 유의미하지 않음, * p<0.05.

도 15A 내지 15C는 겹상 골수 CD34+ 세포의 분화를 도시한다. **도 15A**는 시험관 내 전사된 ZFN mRNA로 전기천공되고 겹상 위치에서 WT 염기를 운반하는 도너 IDLV로 형질도입되고 적혈구 조건하에서 성장한 SCD 환자 골수 CD334+ 세포의 배수 팽창을 보여주는 그래프이다. **도 15B**는 플레이팅된 200 세포당 확인된 조혈 콜로니의 각 유형에 대한 총 콜로니 형성 유닛이 나타난 그래프이다. **도 15C**는 플레이팅된 총 수에 관련하여 형성된 콜로니의 백분율을 보여주는 그래프이다. CFU: 콜로니-형성 유닛; BFU: 발아-형성 유닛; GEMM: 과립구, 적혈구, 단핵세포 및 마크로파지; E: 적혈구계, GM: 과립구 및 마크로파지; G: 과립구; M: 단핵세포, 별표는 유의미성을 나타낸다, * p<0.05, **p<0.01, (n=2 독립적인 실험; Mock n=3; ZFN 단독 n=4; IDLV 단독 n=3, ZFN+IDLV n=6).

도 16A 내지 16D는 겹상 골수 CD34+ 세포의 기능적 교정을 도시한다. 겹상 적혈구 질환 환자 골수 CD34+ 세포가 시험과 내 전사된 ZFN mRNA로 전기천공되었고 겹상 위치에서 WT 염기를 운반하는 도너 IDLV로 형질도입되고 적

혈구 조건하에서 성장되었다. 도 16A는 전기천공 후 12일째에 통합된 RFLP에 대해 qPCR에 의해 분석된 유전자 변형률을 보여주는 그래프이다. 도 16B는 대용량 서열화에 의해 평가된 겸상 돌연변이에서의 교정을 보여주는 그래프이다. 베타-글로빈 유전자좌의 서열화의 결과는 절단 부위에서의 삽입 및 결실 (indel)뿐 아니라 겸상 돌연변이에서의 교정된 WT 염기 (A)를 함유하는 배열된 총 판독의 백분율을 보여준다. 교정된 염기, 백색; indel, 회색. 도 16C는 배양의 종결시 분화된 적혈구계 세포의 HPLC 분석을 보여준다. 세포는 펠릿화되고, 용해되었고, 상층액이 고성능 액체 크로마토그래피 (HPLC)에 의해 분석되었다. 좌측 패널은 SCD mock 샘플을 보여주고 우측 패널은 SCD ZFN+IDLV 샘플을 보여준다. 쉐이딩은 HbA를 나타낸다: WT 성인 혈모글로빈 피크. 도 16D는 주요 피크에 의해 표시된 곡선 아래의 총 면적 중 HbA의 퍼센트의 정량 그래프이다. HbA: WT 성인 혈모글로빈, HbF: 태아 혈모글로빈, HbS: 겸상 혈모글로빈. n.s.: 유의미하지 않음, (n=2 독립적인 실험; Mock n=3; ZFN 단독 n=4; IDLV 단독 n=3, ZFN+IDLV n=6).

도 17A 및 17B는 올리고뉴클레오티드 도너 및 달라지는 농도의 징크-핑거 뉴클레아제 쌍 47773/47817을 사용한 CD34+ 세포의 베타-글로빈 유전자좌에서의 교정을 보여준다. 도 17A는 NHEJ DNA 수복의 빈도를 보여준다. 도 17B는 상동성-의존성 DNA 수복의 빈도를 보여준다. 겸상-적혈구 돌연변이에 더불어, 올리고뉴클레오티드 도너는 도 4C에서 나타낸 SMS12 돌연변이 (검은색 막대) 또는 SNS012 돌연변이 (회색 막대) 중 어느 하나를 포함한다.

발명을 실시하기 위한 구체적인 내용

[0031]

본원에는 유전자 요법 또는 게놈 엔지니어링에 사용하기 위한 세포의 형질도입을 위한 조성물 및 방법이 개시된다. 특히, 표적화된 절단에 의한 외인성 서열 또는 게놈 변경의 뉴클레아제-매개된 (즉 ZFN, TALEN, TtAgo 또는 CRISPR/Cas 시스템) 표적화된 통합은 세포에서 효과적으로 이루어진다. HSC/PC의 형질도입 및 엔지니어링에 대해 특히 유용한 방법 및 조성물은 또한 글로빈 유전자에 삽입 및/또는 결실을 포함하고 있는 유전자 변형된 세포를 제공하기 위해 다른 세포 유형에 대해서도 사용될 수 있다. 구체적으로, 발명의 방법 및 조성물은 겸상 적혈구 질환의 치료 및 방지를 위해 질환 관련 대립유전자를 교정함으로써 글로빈 유전자의 편집에 유용하다.

[0032]

일반적 사항

[0033]

방법의 실행, 뿐만 아니라 본원에 개시된 조성물의 제조 및 사용은, 다르게 표시되지 않는 한, 기술분야의 숙련도 내에 있는 분자 생물학, 생화학, 크로마틴 구조 및 분석, 컴퓨터 화학, 세포 배양, 재조합 DNA 및 관련 분야에서의 종래 기법들을 사용한다. 이들 기법은 문헌에 전체적으로 설명되어 있다. 예를 들면, Sambrook *et al.* MOLECULAR CLONING: A LABORATORY MANUAL, Second edition, Cold Spring Harbor Laboratory Press, 1989 및 Third edition, 2001; Ausubel *et al.*, CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1987 및 periodic updates; the series METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe, CHROMATIN STRUCTURE 및 FUNCTION, Third edition, Academic Press, San Diego, 1998; METHODS IN ENZYMOLOGY, Vol. 304, "Chromatin" (P.M. Wassarman 및 A. P. Wolffe, eds.), Academic Press, San Diego, 1999; 및 METHODS IN MOLECULAR BIOLOGY, Vol. 119, "Chromatin Protocols" (P.B. Becker, ed.) Humana Press, Totowa, 1999를 참조한다.

[0034]

정의

[0035]

용어 "핵산", "폴리펩티드" 및 "올리고뉴클레오티드"는 상호교환적으로 사용되며 선형 또는 원형 형태의, 및 단일- 또는 이중-가닥 중 어느 한 형태의 데옥시리보뉴클레오티드 또는 리보뉴클레오티드 중합체를 나타낸다. 본 발명의 목적에 대해, 이들 용어는 중합체의 길이와 관련하여 제한하는 것으로 해석되지 않아야 한다. 이 용어는 천연 뉴클레오티드의 공지의 유사체뿐 아니라, 염기, 당 및/또는 포스페이트 모이어티에서 변형된 뉴클레오티드 (예컨대 포스포로티오에이트 골격)를 포함할 수 있다. 일반적으로, 특정 뉴클레오티드의 유사체는 동일한 염기-쌍 형성 특이성을 가진다; 즉 A의 유사체는 T와 염기-쌍을 이룰 것이다.

[0036]

용어 "폴리펩티드", "펩티드" 및 "단백질"은 아미노산 잔기의 중합체를 나타내기 위하여 상호교환적으로 사용된다. 이 용어는 또한 하나 이상의 아미노산이 상응하는 자연적으로-발생하는 아미노산의 화학적 유사체 또는 변형된 유도체인 아미노산 중합체에도 적용된다.

[0037]

"결합"은 거대분자들 사이의 (예컨대 단백질과 핵산 사이의) 서열-특이적, 비-공유 상호작용을 나타낸다. 결합 상호작용의 모든 성분은, 전체로서의 상호작용이 서열-특이적인 한, 서열-특이적일 필요는 없다 (예컨대 DNA 골격의 포스페이트 잔기와 접촉함). 그런 상호작용은 일반적으로 10^{-6} M^{-1} 또는 그 아래의 해리 상수 (K_d)에 의해

특성화된다. "친화성"은 결합 강도를 나타낸다: 증가된 결합 친화성은 더 낮은 K_d 와 상관된다.

[0038] "결합 단백질"은 다른 분자에 결합할 수 있는 단백질이다. 결합 단백질은 예를 들면, DNA 분자 (DNA-결합 단백질), RNA 분자 (RNA-결합 단백질) 및/또는 단백질 분자 (단백질-결합 단백질)에 결합할 수 있다. 단백질-결합 단백질의 경우에, 그것은 자체에 결합할 수 있고 (동종이량체, 동종삼량체 등을 형성함) 및/또는 상이한 단백질 또는 단백질들의 하나 이상의 분자에 결합할 수 이TEK. 결합 단백질은 한 가지 이상의 유형의 결합 활성을 가질 수 있다. 예를 들어, 징크 핑거 단백질은 DNA-결합, RNA-결합 및 단백질-결합 활성을 가진다.

[0039] "징크 핑거 DNA 결합 단백질" (또는 결합 도메인)은 하나 이상의 징크 핑거를 통해 서열-특이적 방식으로 DNA에 결합하고, 그 구조가 아연 이온의 조정을 통해 안정화되는 결합 도메인 내의 아미노산 서열의 영역인 더 큰 단백질 내에 있는 단백질 또는 도메인이다. 용어 징크 핑거 DNA 결합 단백질은 자주 징크 핑거 단백질 또는 ZFP로 약칭된다.

[0040] "TALE DNA 결합 도메인" 또는 "TALE"은 하나 이상의 TALE 반복 도메인/유닛을 포함하는 폴리펩티드이다. 반복 도메인은 TALE의 그것의 동족 표적 DNA 서열에의 결합에 포함된다. 단일 "반복 유닛" (또한 "반복부"로서 언급됨)은 전형적으로 33 내지 35 아미노산 길이이고 자연적으로 발생하는 TALE 단백질 내에서 다른 TALE 반복부와 적어도 약간의 서열 상동성을 나타낸다.

[0041] 징크 핑거 및 TALE 결합 도메인은 예정된 뉴클레오티드 서열에, 예를 들면 자연적으로 발생하는 징크 핑거 또는 TALE 단백질의 인식 나선 영역의 엔지니어링 (하나 이상의 아미노산의 변경)을 통해 결합하기 위해 "엔지니어링"될 수 있다. 그러므로, 엔지니어링된 DNA 결합 단백질 (징크 핑거 또는 TALEs)은 자연적으로 발생하지 않는 단백질이다. DNA-결합 단백질을 엔지니어링하는 방법의 비-제한적인 실례는 디자인 및 선택이다. 디자인된 DNA 결합 단백질은 그것의 디자인/조성이 원칙적으로 합리적 기준으로부터 유래되는 자연에서 발생하지 않는 단백질이다. 디자인을 위한 합리적 기준은 치환 규칙 및 기존의 ZFP 및/또는 TALE 디자인 및 결합 데이터의 데이터베이스 저장 정보에서의 프로세싱 정보에 대한 컴퓨터화된 알고리즘의 적용을 포함한다. 예를 들면 미국 특허 6,140,081; 6,453,242; 6,534,261 및 8,586,526; 또한 WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 및 WO 03/016496을 참조한다.

[0042] "선택된" 징크 핑거 단백질 또는 TALE은 그것의 구조가 원칙적으로 과거 디스플레이, 상호작용 트랩 또는 하이브리드 선택과 같은 실험 과정으로부터 유래하는, 자연에서 발견되지 않는 단백질이다. 예컨대 미국 특허 번호 5,789,538; 5,925,523; 6,007,988; 6,013,453; 6,200,759; 8,586,526; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197, WO 02/099084를 참조한다.

[0043] "TtAgo"는 유전자 침묵에 포함되는 것으로 여겨지는 원핵 아르고노트 (Argonaute) 단백질이다. TtAgo는 박테리아 써머스 써모필러스 (*Thermus thermophilus*)로부터 유도된다. 예컨대 Swarts *et al.*, *ibid*, G. Sheng *et al.*, (2013) *Proc. Natl. Acad. Sci. U.S.A.* 111, 652를 참조한다. "TtAgo 시스템"은 TtAgo 효소에 의한 절단을 위해 필요한, 예컨대 가이드 RNA를 포함하는 모든 성분이다.

[0044] "재조합"은 제한은 아니지만, 비-상동성 단부 결합 (NHEJ) 및 상동성 재조합에 의한 도너 포획을 포함하여, 두 개의 폴리뉴클레오티드 사이의 유전자 정보의 교환 과정을 나타낸다. 본 발명의 목적에 대하여, "상동성 재조합 (HR)"은 예를 들면 상동성-지시된 수복 메커니즘을 통해 세포에서 이중-가닥 파괴를 수복하는 중에 일어나는 그런 교환의 특수한 형태를 나타낸다. 이 과정은 뉴클레오티드 서열 상동성을 필요로 하고, "표적" 분자 (즉 이중-가닥 파괴를 경험한 분자)의 주형 수복을 위해 "도너" 분자를 사용하며, 그것이 도너로부터 표적으로의 유전자 정보의 전달을 유도하기 때문에 "비-교차형 유전자 전환" 또는 "짧은 트랙 유전자 전환"으로 알려진다. 임의의 특정 이론에 의해 구속되는 것을 바라지 않으면서, 그런 전달은 파괴된 표적과 도너 사이에 형성되는 이중듀플렉스 DNA의 미스매치 교정, 및/또는 도너가 표적의 일부가 될 유전자 정보를 재합성하기 위해 사용되는 "합성-의존성 가닥 아닐링" 및/또는 관련된 과정들을 포함할 수 있다. 그런 특수화된 HR은 자주 표적 분자의 서열의 변경을 초래하여서 도너 폴리뉴클레오티드의 서열의 일부 또는 전부가 표적 폴리뉴클레오티드 안에 통합된다.

[0045] 발명의 방법에서, 본원에서 기술되는 하나 이상의 표적화된 뉴클레아제는 예정된 부위에서 표적 서열에 (예컨대 세포 크로마틴) 이중-가닥 파괴 (DSB)를 생성한다. DSB는 상동성-지시된 수복에 의해 또는 비-상동성-지시된 수복 메커니즘에 의해 결실 및/또는 삽입을 초래할 수 있다. 결실은 임의의 수의 염기쌍을 포함할 수 있다. 유사하게, 삽입은 예를 들면, 파괴 영역에서 뉴클레오티드 서열에 대한 상동성을 선택적으로 가진는 "도너" 폴리뉴클레오티드의 통합을 포함하여, 임의의 수의 염기쌍을 포함할 수 있다. 도너 서열은 물리적으로 통합될 수 있고, 또는, 다르게는, 도너 폴리뉴클레오티드는 상동성 재조합을 통해 파괴의 수복을 위한 주형으로서 사용되

어, 세포 크로마틴 안으로의 도너에서와 같이 뉴클레오티드 서열의 전부 또는 일부의 통합을 초래한다. 그러므로, 세포 크로마틴의 제1 서열은 변경될 수 있고, 특정 구체예에서는, 도너 폴리뉴클레오티드에 존재하는 서열로 전환될 수 있다. 그러므로, 용어 "대체하다" 또는 "대체"의 사용은 한 뉴클레오티드 서열의 다른 것에 의한 대체(즉 정보 차원에서 서열의 대체)를 나타내는 것으로 인지될 수 있고, 한 폴리뉴클레오티드의 다른 것에 의한 물리적 또는 화학적 대체를 반드시 필요로 하지는 않는다.

[0046] 본원에 기술된 방법들 중 임의의 방법에서, 징크-핑거 단백질, TALEN, TtAgo 및/또는 CRISPR/Cas 시스템의 추가의 쌍이 세포 내에서 추가의 표적 부위의 추가의 이중-가닥 절단을 위해 사용될 수 있다.

[0047] 본원에 기술된 방법들 중 어떠한 것이든지 관심의 유전자(들)의 발현을 파괴하는 도너 서열의 표적화된 통합에 의해 세포에서 임의의 크기의 도너의 삽입 및/또는 하나 이상의 표적 서열의 부분적인 또는 완전한 비활성화에 대해 사용될 수 있다. 부분적으로 또는 완전하게 비활성화된 유전자들을 가진 세포주 또한 제공된다.

[0048] 본원에 기술된 방법들 중 임의의 방법에서, 외인성 뉴클레오티드 서열 ("도너 서열" 또는 "전이유전자")은 관심 영역의 게놈 서열들에 상동하지만 동일하지는 않은 서열들을 함유할 수 있고, 따라서 관심 영역의 비-동일 서열을 삽입하기 위한 상동성 재조합을 자극할 수 있다. 그러므로, 특정 구체예에서, 관심 영역의 서열들에 상동하는 도너 서열의 부분들은 대체되는 게놈 서열에 대해 약 80 내지 99% (또는 그 사이의 임의의 정수)의 서열 동일성을 나타낸다. 다른 구체예에서, 예를 들어 만약 도너와 100 연속 염기쌍이 넘는 게놈 서열 사이에서 단지 한 뉴클레오티드만이 다르다면 도너와 게놈 서열 사이의 상동성은 99%보다 높다. 특정 경우에, 도너 서열의 비-상동성 부분은 관심 영역에 존재하지 않은 서열들을 함유할 수 있어서, 새로운 서열이 관심 영역에 도입된다. 이런 경우에, 비-상동성 서열은 일반적으로 관심 영역의 서열들에 상동하거나 동일한 50 내지 1,000 염기쌍 (또는 그 사이의 임의의 정수)의 서열 또는 1,000보다 큰 임의의 수의 염기쌍이 양 옆에 있다. 다른 구체예에서, 도너 서열은 제1 서열에 비-상동성이고, 비-상동성 재조합 메커니즘에 의해 게놈 안으로 삽입된다.

[0049] "절단"은 DNA 분자의 공유 골격의 파괴를 나타낸다. 절단은 제한은 아니지만, 포스포다이에스테르 결합의 효소적 또는 화학적 가수분해를 포함하는 다양한 방법들에 의해 개시될 수 있다. 단일-가닥 절단 및 이중-가닥 절단들 다 가능하며, 이중-가닥 절단은 두 개의 구별되는 단일-가닥 절단 사건의 결과로서 일어날 수 있다. DNA 절단은 평활(blunt) 단부 또는 점착(staggered) 단부 중 어느 하나의 생성을 초래할 수 있다. 특정 구체예에서, 융합 폴리펩티드가 표적화된 이중-가닥 절단에 사용된다.

[0050] "절단 하프-도메인"은 제2 폴리펩티드 (동일하거나 상이함)와 함께 절단 활성 (바람직하게는 이중-가닥 절단 활성)을 가지는 복합체를 형성한다. 용어 "제1 및 제2 절단 하프-도메인", "+ 및 - 절단 하프-도메인" 및 "우측 및 좌측 절단 하프-도메인"은 이량체화하는 절단 하프-도메인의 쌍을 나타내기 위해 상호교환적으로 사용된다.

[0051] "엔지니어링된 절단 하프-도메인"은 다른 절단 하프-도메인 (예컨대 또 다른 엔지니어링된 절단 하프-도메인)과 부득이한 이종이량체를 형성하기 위해 변형되어 있는 절단 하프-도메인이다. 미국 특허 번호 7,888,121; 7,914,796; 8,034,598 및 8,823,618 참조 (이것들은 전체 내용이 참조로 본원에 포함된다).

[0052] 용어 "서열"은 DNA 또는 RNA일 수 있고; 선형, 원형 또는 분지형일 수 있으며 단일-가닥 또는 이중 가닥 중 어느 하나일 수 있는, 임의의 길이의 뉴클레오티드 서열을 나타낸다. 용어 "도너 서열"은 게놈 안으로 삽입되는 뉴클레오티드 서열을 나타낸다. 도너 서열은 임의의 길이, 예를 들면 2 내지 100,000,000 뉴클레오티드의 길이 (또는 그 사이의 또는 그것을 초과하는 임의의 정수), 바람직하게는 약 100 내지 100,000 뉴클레오티드 길이 (또는 그 사이의 임의의 정수), 보다 바람직하게는 약 2000 내지 20,000 뉴클레오티드 길이 (또는 그 사이의 임의의 정수), 보다 더 바람직하게는 약 5 내지 15 kb (또는 그 사이의 임의의 정수)의 것일 수 있다.

[0053] "크로마틴"은 세포 게놈을 포함하는 핵단백질 구조이다. 세포 크로마틴은 핵산, 주로 DNA 및, 히스톤 및 비-히스톤 염색체 단백질을 포함하여 단백질을 포함한다. 진핵 세포 크로마틴의 대부분은 뉴클레오솜의 형태로 존재하는데, 뉴클레오솜 코어는 히스톤 H2A, H2B, H3 및 H4 중 각각을 두 개 포함하는 8량체와 회합된 DNA의 대략 150 염기쌍; 및 뉴클레오솜 코어 사이에 뻗어 있는 링커 DNA (유기체에 따라 다양한 길이를 가짐)를 포함한다. 히스톤 H1의 분자는 일반적으로 링커 DNA와 회합된다. 본 발명의 목적에 대해, 용어 "크로마틴"은 원핵 및 진핵 둘 다의 세포 핵단백질의 모든 유형을 포함하는 것을 의미한다. 세포 크로마틴은 염색체 및 에피솜 크로마틴 둘 다를 포함한다.

[0054] "염색체"는 세포의 게놈의 전부 또는 일부를 포함하는 크로마틴 복합체이다. 세포의 게놈은 자주, 세포의 게놈을 포함하는 모든 염색체의 집합인, 그것의 핵형에 의해 특성화된다. 세포의 게놈은 하나 이상의 염색체를 포함할 수 있다.

- [0055] "에피솜"은 복제 핵산, 핵단백질 복합체 또는 세포의 염색체 핵형의 일부가 아닌 핵산을 포함하는 다른 구조이다. 에피솜의 실례로는 플라스미드 및 특정 바이러스 계놈이 있다.
- [0056] "접근 가능한 영역"은 핵산에 존재하는 표적 부위가 그 표적 부위를 인식하는 외인성 분자에 의해 결합될 수 있는 세포 크로마틴에 있는 부위이다. 임의의 특정 이론에 의해 구속되는 것을 원하지 않으면서, 접근 가능한 영역은 뉴클레오솜 구조 안으로 패키징되지 않는 것이라 여겨진다. 접근 가능한 영역의 뚜렷한 구조는 자주 그것의 화학적 및 효소적 프로브, 예를 들면 뉴클레아제에 대한 민감성에 의해 검출될 수 있다.
- [0057] "표적 부위" 또는 "표적 서열"은 결합에 대한 충분한 조건이 존재한다면, 담결합 분자가 결합하게 될 핵산의 부분을 규정하는 핵산 서열이다.
- [0058] "외인성" 분자는 세포에 정상적으로 존재하지는 않지만, 하나 이상의 유전자적, 생화학적 또는 다른 방법에 의해 세포 안으로 도입될 수 있는 분자이다. "세포에서의 정상적인 존재"는 세포의 특정 발달 단계 및 환경 조건과 관련하여 측정된다. 그러므로, 예를 들어, 균육의 배아 발달 중에만 존재하는 분자는 성인 균육 세포와 관련하여 외인성 분자이다. 유사하게, 열 충격에 의해 유도된 분자는 비-열-충격 세포와 관련하여 외인성 분자이다. 외인성 분자는, 예를 들면 기능부전 내인성 분자의 기능화 버전 또는 정상적으로-기능하는 내인성 분자의 기능부전 버전을 포함할 수 있다.
- [0059] 외인성 분자는, 다른 것들 중에서도, 조합적인 화학 과정에 의해 생성된 것과 같은 작은 분자, 또는 단백질, 핵산, 탄수화물, 지질, 당단백질, 리포단백질, 다당류, 상기 분자들의 임의의 변형된 유도체, 또는 상기 분자들의 하나 이상을 포함하는 임의의 복합체와 같은 거대분자일 수 있다. 핵산은 DNA와 RNA를 포함하며, 단일- 또는 이중-가닥일 수 있고; 선형, 분지형 또는 원형일 수 있으며; 임의의 길이의 것일 수 있다. 핵산은 듀플렉스를 형성할 수 있는 것들뿐 아니라 트리플렉스-형성 핵산을 포함한다. 예를 들면 미국 특허 번호 5,176,996 및 5,422,251 참조. 단백질은, 제한은 아니지만, DNA-결합 단백질, 전사 인자, 크로마틴 리모델링 인자, 메틸화된 DNA-결합 단백질, 중합효소, 메틸화제, 탈메틸화제, 아세틸화제, 탈아세틸화제, 키나아제, 포스파타제, 인테그라제, 재조합효소, 리가아제, 토포아이소메라제, 기라아제 및 헬리카아제를 포함한다.
- [0060] 외인성 분자는 내인성 분자와 동일한 유형의 분자, 예컨대 외인성 단백질 또는 핵산일 수 있다. 예를 들어, 외인성 핵산은 감염성 바이러스 계놈, 세포 안으로 도입된 플라스미드 또는 에피솜, 또는 세포에 정상적으로 존재하지 않는 염색체를 포함할 수 있다. 외인성 분자를 세포 안으로 도입시키기 위한 방법들은 기술분야의 숙련자들에게 알려져 있고, 제한은 아니지만, 지질-매개된 전달 (즉 중성 및 양이온성 지질을 포함하여 리포솜), 전기 천공, 직접 주입, 세포 융합, 입자 폭발, 칼슘 포스페이트 공-침전, DEAE-덱스트란-매개 전달 및 바이러스 벡터-매개 전달을 포함한다. 외인성 분자는 또한 내인성 분자와 동일한 유형의 분자일 수 있지만 그 세포가 유도되는 것과 상이한 종으로부터 유도된다. 예를 들어, 인간 핵산 서열은 마우스 또는 햄스터로부터 원래 유도된 세포주 안에 도입될 수 있다. 외인성 분자의 식물 세포 안으로의 도입 방법은 기술분야의 숙련자들에게 알려져 있으며, 제한은 아니지만 원형질 형질전환, 탄화규소 (예컨대 WHISKERS™), 아그로박테리움 (*Agrobacterium*)-매개된 형질전환, 지질-매개된 전달 (즉 중성 및 양이온성 지질을 포함하여 리포솜), 전기천공, 직접 주입, 세포 융합, 입자 폭발 (예컨대 "유전자 총"을 사용함), 칼슘 포스페이트 공-침전, DEAE-덱스트란-매개 전달 및 바이러스 벡터-매개 전달을 포함한다.
- [0061] 대조적으로, "내인성" 분자는 특정 세포에 특정 발달 단계에서 특정 환경 조건 하에 정상적으로 존재하는 분자이다. 예를 들어, 내인성 핵산은 염색체, 미토콘드리아의 계놈, 엽록체 또는 다른 소기관, 또는 자연적으로-발생하는 에피솜성 핵산을 포함할 수 있다. 추가의 내인성 분자로는 단백질, 예를 들면 전사 인자 및 효소를 들 수 있다.
- [0062] 본원에서 사용되는 용어 "외인성 핵산의 생성물"은 폴리뉴클레오티드 및 폴리펩티드 생성물 둘 다, 예를 들면 전사 생성물 (RNA와 같은 폴리뉴클레오티드) 및 변역 생성물 (폴리펩티드)을 포함한다.
- [0063] "융합" 분자는 둘 이상의 하위유닛 분자가, 바람직하게는 공유적으로 연결되어 있는 분자이다. 하위유닛 분자들은 동일한 화학적 유형의 분자일 수 있거나, 또는 상이한 화학적 유형의 분자일 수 있다. 융합 분자의 제1 유형의 실례로는, 제한은 아니지만, 융합 단백질 (예를 들면 ZFP 또는 TALE DNA-결합 도메인과 하나 이상의 활성화 도메인 사이의 융합) 및 융합 핵산 (예를 들면 상기 기술된 융합 단백질을 코드화하는 핵산)을 포함한다. 융합 분자의 제2 유형의 실례로는, 제한은 아니지만, 트리플렉스-형성 핵산과 폴리펩티드 사이의 융합, 및 마이너 그루브 결합제 (minor groove binder)와 핵산 사이의 융합을 포함한다.
- [0064] 융합 단백질의 세포에서의 발현은 그 융합 단백질의 세포에의 전달로부터 또는 그 융합 단백질을 코드화하는 폴

리뉴클레오티드의 세포에의 전달에 의해 유발될 수 있고, 이때 폴리뉴클레오티드는 전사되고, 전사물은 번역되어 융합 단백질이 생성된다. 트랜스-스플라이싱, 폴리펩티드 절단 및 폴리펩티드 결합이 또한 단백질의 세포에서의 발현에 포함될 수 있다. 폴리뉴클레오티드 및 폴리펩티드의 세포에의 전달 방법은 본 발명의 다른 곳에서 제시된다.

[0065] 본 발명의 목적에 대해 "유전자"는 유전자 생성물 (이하 참조)을 코드화하는 DNA 영역뿐 아니라, 그 유전자 생성물의 생성을 조절하는 모든 DNA 영역을, 이때 그런 조절 서열이 코딩 및/또는 전사된 서열에 인접하든 아니든 포함한다. 따라서, 유전자는, 반드시 그것들에 제한될 필요는 않지만, 프로모터 서열, 터미네이터, 리보솜 결합부위 및 내부 리보솜 유입 부위와 같은 번역 조절 서열, 인핸서, 사일런서, 인슬레이터, 경계 요소, 복제 기원, 매트릭스 부착 부위 및 유전자좌 제어 영역을 포함한다.

[0066] "유전자 발현"은 유전자에 함유된 정보의 유전자 생성물로의 전환을 나타낸다. 유전자 생성물은 유전자의 직접적인 전사 생성물 (예컨대 mRNA, tRNA, rRNA, 앤티센스 RNA, 리보자임, 구조 RNA 또는 임의의 다른 유형의 RNA) 또는 mRNA의 번역에 의해 생성된 단백질일 수 있다. 유전자 생성물은 또한 캡핑, 폴리아데닐화, 메틸화, 및 편집과 같은 과정에 의해 변형된 RNA, 및 예를 들면 메틸화, 아세틸화, 인산화, 유비퀴틴화, ADP-리보실화, 미리스틸화 및 글리코실화에 의해 변형된 단백질을 포함한다.

[0067] 유전자 발현의 "조정"은 유전자의 활성화 변화를 나타낸다. 발현의 조정은, 제한은 아니지만 유전자 활성화 및 유전자 억압을 포함할 수 있다. 게놈 편집 (예컨대 절단, 변경, 비활성화, 무작위 돌연변이)이 발현을 조정하기 위해 사용될 수 있다. 유전자 비활성화는 본원에 기술된 ZFP, TALE, TtAgo 또는 CRISPR/Cas 시스템을 포함하지 않는 세포와 비교하여 유전자 발현에서의 임의의 감소를 나타낸다. 그러므로, 유전자 비활성화는 부분적이거나 완전할 수 있다.

[0068] "관심 영역"은 내인성 분자에 결합하는 것이 바람직한, 세포 크로마틴의 임의의 영역, 예컨대, 예를 들면 유전자 또는 유전자 내부의 또는 인접한 비-코딩 서열이다. 결합은 표적화된 DNA 절단 및/또는 표적화된 재조합의 목적에 대한 것일 수 있다. 관심 영역은 염색체, 에피솜, 소기관 게놈 (예컨대 미토콘드리아, 엽록체), 또는 예를 들면 감염성 바이러스 게놈에 존재할 수 있다. 관심 영역은 유전자의 코딩 영역 내부에, 전사된 비-코딩 영역 내부에, 예컨대, 예를 들면 리더 서열, 트레일러 서열 또는 인트론에, 또는 코딩 영역의 상류 또는 하류에 있는 비-전사 영역 내부에 존재할 수 있다. 관심 영역은 단일 뉴클레오티드 쌍으로서 또는 최대 2,000 뉴클레오티드 쌍의 길이, 또는 임의의 정수 값의 뉴클레오티드 쌍의 길이 정도로 작을 수 있다.

[0069] "진핵" 세포는, 제한은 아니지만 줄기 세포 (다능 및 다분화능)를 포함하여, 진균 세포 (예컨대 효모), 식물 세포, 동물 세포, 포유류 세포 및 인간 세포 (예컨대 T-세포)를 포함한다.

[0070] 용어 "작동 가능한 연쇄" 및 "작동 가능하게 연결된" (또는 "작동하도록 연결된")은 둘 이상의 성분 (예컨대 서열 요소들)의 병치와 관련하여 상호교환적으로 사용되고, 이때 성분들은 두 성분이 정상적으로 기능하고 성분들 중 적어도 하나가 다른 성분들 중 적어도 하나에 대해 발휘된 기능을 매개할 수 있는 가능성을 허용하도록 배열된다. 예시에 의하면, 프로모터와 같은 전사 조절 서열은 전사 조절 서열이 하나 이상의 전사 조절 인자들의 존재 또는 부재에 반응하여 코딩 서열의 전사 수준을 제어한다면 코딩 서열에 작동 가능하게 연결된다. 전사 조절 서열은 일반적으로 코딩 서열과 시스 (cis)로 작동 가능하게 연결되지만, 그것에 바로 인접할 필요는 없다. 예를 들어, 인핸서는 연속적으로 위치하진 않을지라도 코딩 서열에 작동 가능하게 연결되어 있는 전사 조절 서열이다.

[0071] 융합 폴리펩티드와 관련하여, 용어 "작동 가능하게 연결된"은 각 성분이 그것이 그렇게 연결되지 않는다면 일어날 수 있는 것과 같이 다른 성분에 대한 연쇄에서 동일한 기능을 수행한다는 사실을 나타낼 수 있다. 예를 들어, ZFP, TALE, TtAgo 또는 Cas DNA-결합 도메인이 활성화 도메인에 융합되어 있는 융합 폴리펩티드와 관련하여, ZFP, TALE, TtAgo 또는 Cas DNA-결합 도메인 및 활성화 도메인은 만약 융합 폴리펩티드에서 ZFP, TALE, TtAgo 또는 Cas DNA-결합 도메인과 그것의 표적 부위 및/또는 그것의 결합 부위에 결합할 수 있는 한편, 활성화 도메인은 유전자 발현을 상향 조절할 수 있다면, 작동 가능한 연쇄에 있는 것이다. ZFP, TALE, TtAgo 또는 Cas DNA-결합 도메인이 절단 도메인에 융합되어 있는 융합 폴리펩티드라면, ZFP, TALE, TtAgo 또는 Cas DNA-결합 도메인 및 절단 도메인은, 만약 융합 폴리펩티드에서, ZFP, TALE, TtAgo 또는 Cas DNA-결합 도메인 부분이 그것의 표적 부위 및/또는 그것의 결합 부위에 결합할 수 있는 한편, 절단 도메인은 표적 부위 가까이에서 DNA를 절단할 수 있다면, 작동 가능한 연쇄에 있는 것이다.

[0072] 단백질, 폴리펩티드 또는 핵산의 "기능성 단편"은 그것의 서열이 전-길이 단백질, 폴리펩티드 또는 핵산에 동일

하지 않지만, 전-길이 단백질, 폴리펩티드 또는 핵산과 동일한 기능을 보유하는 단백질, 폴리펩티드 또는 핵산이다. 기능성 단편은 상응하는 천연 분자보다 많거나, 적거나 또는 그것과 동일한 수의 잔기를 가질 수 있고, 및/또는 하나 이상의 아미노산 또는 뉴클레오티드 치환을 함유할 수 있다. 핵산의 기능(예컨대 코딩 기능, 다른 핵산에 하이브리드를 형성하는 능력)을 측정하는 방법은 기술 분야에 잘 알려져 있다. 유사하게, 단백질 기능을 측정하는 방법도 잘 알려져 있다. 예를 들어, 폴리펩티드의 DNA-결합 기능은 예를 들면 필터-결합, 전기영동적 이동성-시프트 또는 면역침전 분석에 의해 측정될 수 있다. DNA 절단은 겔 전기영동에 의해 분석될 수 있다. 상기 Ausubel *et al.* 참조. 단백질의 다른 단백질과 상호작용하는 능력은 예를 들어 공-면역침전, 2-하이브리드 분석 또는 상보성에 의해 유전자적으로 및 생화학적으로 측정될 수 있다. 예를 들면 Fields *et al.* (1989) *Nature* 340:245-246; 미국 특허 번호 5,585,245 및 PCT WO 98/44350을 참조한다.

[0073] "벡터"는 표적 세포에 유전자 서열을 전달할 수 있다. 전형적으로, "벡터 구성물", "발현 벡터" 및 "유전자 전달 벡터"는 관심의 유전자의 발현을 지시할 수 있고 유전자 서열을 표적 세포에 전달할 수 있는 모든 핵산 구성물을 의미한다. 그러므로, 이 용어는 클로닝, 및 발현 비히클뿐 아니라, 통합 벡터를 포함한다.

[0074] 용어 "대상" 및 "환자"는 상호교환적으로 사용되고, 인간 환자 및 비-인간 환자, 뿐만 아니라 토끼, 개, 고양이, 래트, 마우스 및 다른 동물과 같은 실험 동물과 같은 포유류를 나타낸다. 따라서, 본원에서 사용되는 용어 "대상" 또는 "환자"는 발명의 뉴클레아제, 도너 및/또는 유전자 변형된 세포가 투여될 수 있는 임의의 포유류 환자 또는 대상을 의미한다. 본 발명의 대상은 장애가 있는 대상을 포함한다.

[0075] "줄기성"은 임의의 세포의 줄기세포-유사 방식으로 작용하는 상대적인 능력, 즉 임의의 특정 줄기세포가 가질 수 있는 전체적인-, 다수의- 또는 소량-정도의 능력 및 팽창된 또는 무기한의 자기-재생성을 나타낸다.

줄기세포 팽창을 증대시키는 인자들

[0077] 줄기세포 팽창 및/또는 분화를 증대시키는 임의의 인자 또는 인자들은 본 발명의 실시에 사용될 수 있다. 인자들은 직접 세포 안으로 도입되거나 (예를 들면, 인자(들)을 코드화하는 유전자로서) 및/또는 세포에 영향을 미치기 위해 주변의 배양 배지 (배양보조세포층 및 다른 고체 기질을 포함함) 안으로 도입될 수 있다. 뉴클레아제-매개된 변형이 유도되기 전에, 중에 또는 후에, 예를 들면 배양 조건에 그런 인자들을 사용하는 것은 줄기세포의 뉴클레아제-매개된 변형 속도를 증가시킨다.

[0078] 사용될 수 있는 인자들의 비-제한적인 실례로는, SR1, 아릴 탄화수소 수용체 길항체, dmPGE2, 프로스타글란딘, UM171 및 UM729, 라이브러리 스크린에서 확인된 화합물들 (Pabst *et al.* (2014) *Nat Meth* 11:436-442 참조), 라파마이신 (Wang *et al.* (2014) *Blood*. pii: blood-2013-12-546218 참조), 안지오포이에틴-유사 단백질 ("Angpt1s", 예컨대 Notch/delta/ANGPTL5 (Zhang *et al.* (2008) *Blood*. 111(7):3415-3423 참조), Angpt12, Angpt13, Angpt15, Angpt17, 및 Mfap4), 구리 킬레이터 테트라에틸에테펜타민 (TEPA, de Lima *et al.* (2008) *Bone Mar Trans* 41:771-778 참조), 허스톤 탈아세틸화제 (HPAC) 억제제, 예컨대 밸프로산 (Chaurasia *et al.* (2014) *J Clin Invest* 124:2378-2395 참조), IGF-결합 단백질 2 (IGFBP2), 니코틴아미드 (Horwitz *et al.* (2014) *J. Clin. Invest* 124:3121-3128 참조), Tat-myc (WO2010025421) 및 tat-Bc12 (WO2014015312) 융합 단백질, MAPK14/p38a Ly2228820 (Baudet *et al.* (2012) *Blood* 119 (26):6255-6258 참조), HOXB4, OCT3/4 제대혈 및/또는 MSC 유도된 배양보조세포층 또는 E4+EC로 명명된 생체 외 혈관 틈새 공-배양 시스템 (Butler *et al.* (2012) *Blood*. 120(6):1344-1347 참조)과 같은 자기-재생 유전자들의 생성물, 사이토카인, (비-제한적 실례 StemspanTM CC110, CC100, 및/또는 H3000 (StemcellTM 기술), Flt-3 리간드, SCF, TPO)를 포함한다.

[0079] 일부 구체예에서, 인자들은 StemRegenin (SR1, 예컨대 미국 특허 번호 8,741,640; Boitano *et al.*, (2010) *Science* 329(5997):1345-1348), 본원에서 기술된 방법 및 조성물에 생체 외적으로 사용된 CD34+ 세포의 팽창을 촉진하는 아릴 탄화수소 수용체 (AhR) 길항체를 포함한다. 다른 구체예에서, 인자들은 줄기세포 재생의 아고니스트인 UM171 (Fares *et al.* (2013) *Blood*:122 (21))을 포함한다. 또 다른 측면으로, 인자는 하나 이상의 프로스타글란딘, 예를 들면 dmPGE2를 포함한다. 예컨대 미국 특허 번호 8,168,428; North *et al.* (2007) *Nature* 447(7147):1007-1011 참조. 일부 측면으로, 인자는 안지오포이에틴-유사 단백질과 같은 하나 이상의 호르몬 ("Angpt1s", 예컨대 Angpt12, Angpt13, Angpt15, Angpt17 및 Mfap4)을 포함하고, IGF-결합 단백질 2 (IGFBP2)가 사용된다. 예컨대 미국 특허 번호 7,807,464; Zhang *et al.* (2008) 111(7):3415-3423) 참조. 다른 측면으로 인자들은 사용된 HOXB4 또는 OCT와 같은 자기-재생 유전자들의 하나 이상의 단백질 단백질을 포함한다. 예컨대 미국 특허 번호 8,735,153; Watts *et al.* (2012) *Exp Hematol.* 40(3):187-196 참조. 다르게는 이들 유전자들은 배양 배지에서 및/또는 줄기세포에서 일시적으로 발현될 수 있다.

- [0080] 줄기세포 팽창에 영향을 미치는 인자들은 또한, 제한되는 것은 아니지만, 간질 세포 및/또는 MSC 유도된 세포로부터 유도된 배양보조세포층을 포함하여 세포 지지 방법을 포함할 수 있다. 예컨대 Breems *et al* (1998) *Blood* 91(1):111-117 및 Magin *et al.*, (2009) *Stem Cells Dev.* 2009 Jan-Feb;18(1):173-86을 참조한다.
- [0081] 줄기세포 팽창을 증대시키는 하나 이상의 인자들은 그것이 뉴클레아제 활성 및 뉴클레아제-매개된 계놈의 변형을 증가시키기에 효과적인 한 임의의 적합한 양으로 사용될 수 있다. 사용된 특정 농도는 기술분야의 숙련자에 의해 쉽게 결정될 수 있다. 특정 구체예에서, 0.1 nM 내지 100 μM이 사용되며, 예를 들면 0.5 μM 내지 25 μM 농도가 사용되며, 그 사이의 임의의 양도 포함된다 ((예컨대 1 μM 내지 20 μM, 3 μM 내지 10 μM 등)).
- [0082] **융합 분자**
- [0083] 본원에는 세포, 특히 줄기세포에서 선택된 표적 유전자의 절단에 유용한 조성물, 예를 들면 뉴클레아제가 기술된다. 특정 구체예에서, 융합 분자 (예컨대 뉴클레아제)의 하나 이상의 성분은 자연적으로 발생한다. 다른 구체예에서, 융합 분자 (예컨대 뉴클레아제)의 하나 이상의 성분은 자연적으로 발생하지 않는다, 즉 DNA-결합 도메인(들) 및/또는 절단 도메인(들)에서 엔지니어링된다. 예를 들어, 자연적으로-발생하는 뉴클레아제의 DNA-결합 도메인은 선택된 표적 부위 (예컨대 동족 결합 부위와는 상이한 부위에 결합하도록 엔지니어링되어 있는 메가뉴클레아제)에 결합하도록 변경될 수 있다. 다른 구체예에서, 뉴클레아제는 이종성 DNA-결합 및 절단 도메인 (예컨대 징크 평거 뉴클레아제; TAL-이펙터 도메인 DNA 결합 단백질; 이종성 절단 도메인을 가지는 메가뉴클레아제 DNA-결합 도메인)을 포함한다.
- [0084] A. DNA-결합 분자
- [0085] 특정 구체예에서, 본원에 기술된 조성물 및 방법은 도너 분자에 대한 결합 및/또는 세포의 계놈에서 관심 영역에 대한 결합을 위해 메가뉴클레아제 (호밍 엔도뉴클레아제) DNA-결합 분자 (또한 DNA-결합 도메인으로도 언급됨)를 사용한다. 자연적으로-발생하는 메가뉴클레아제는 15 내지 40 염기쌍 절단 부위를 인식하고 통상적으로 다음의 4개의 패밀리로 구분된다: LAGLIDADG 패밀리, GIY-YIG 패밀리, His-Cyst 박스 패밀리 및 HNH 패밀리. 호밍 엔도뉴클레아제의 예를 들면 I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-TevI, I-TevII 및 I-TevIII을 포함한다. 또한 미국 특허 번호 5,420,032; 미국 특허 번호 6,833,252; Belfort *et al.* (1997) *Nucleic Acids Res.* 25:3379-3388; Dujon *et al.* (1989) *Gene* 82:115-118; Perler *et al.* (1994) *Nucleic Acids Res.* 22, 1125-1127; Jaschinski (1996) *Trends Genet.* 12:224-228; Gimble *et al.* (1996) *J. Mol. Biol.* 263:163-180; Argast *et al.* (1998) *J. Mol. Biol.* 280:345-353 및 New England Biolabs 카탈로그를 참조한다. 또한, 호밍 엔도뉴클레아제 및 메가뉴클레아제의 DNA-결합 특이성은 비-천연 표적 부위에 결합하기 위해 엔지니어링될 수 있다. 예를 들면 Chevalier *et al.* (2002) *Molec. Cell* 10:895-905; Epinat *et al.* (2003) *Nucleic Acids Res.* 31:2952-2962; Ashworth *et al.* (2006) *Nature* 441:656-659; Paques *et al.* (2007) *Current Gene Therapy* 7:49-66; 미국 특허 공개공보 번호 20070117128을 참조한다. 호밍 엔도뉴클레아제 및 메가뉴클레아제의 DNA-결합 도메인은 전체로서 뉴클레아제의 맥락에서 변경될 수 있고 (즉 뉴클레아제가 동족의 절단 도메인을 포함하도록) 또는 이종성 절단 도메인에 융합될 수 있다.
- [0086] 다른 구체예에서, 본원에 기술된 방법 및 조성물에서 사용된 하나 이상의 뉴클레아제의 DNA-결합 도메인은 자연적으로 발생하는 또는 엔지니어링된 (자연적으로 발생하지 않는) TAL 이펙터 DNA 결합 도메인을 포함한다. 예컨대 본원에 전체 내용이 참조로 포함되는 미국 특허 번호 8,586,526 참조. 크산토모나스 속 (genus *Xanthomonas*)의 식물 병원성 박테리아는 중요한 농작물에서 많은 질환을 유발하는 것으로 알려진다. 크산토모나스의 병원성은 25 이상의 상이한 이펙터 단백질을 식물 세포 안으로 주입하는 보존된 유형 III 분비 (T3S) 시스템에 좌우된다. 이들 주입된 단백질 중에 식물 전사 활성화제를 모방하고 식물 전사체를 조작하는 전사 활성화제-유사 (TAL) 이펙터가 있다 (Kay *et al* (2007) *Science* 318:648-651 참조). 이들 단백질은 DNA 결합 도메인 및 전사 활성화 도메인을 함유한다. 특성화가 가장 잘 이루어져 있는 것들 중 한 가지는 크산토모나스 캄페스트그리스 피브이. 베시카토리아 (*Xanthomonas campestris* pv. *Vesicatoria*)로부터의 AvrBs3이다 (Bonas *et al* (1989) *Mol Gen Genet* 218:127-136 및 WO2010079430 참조). TAL-이펙터는 이들 단백질의 DNA 결합 특이성에 대한 핵심인 탠덤 반복부의 집중된 도메인을 함유하고, 각 반복부는 대략 34 아미노산을 함유하고 있다. 또한, 그것들은 핵 정위 서열 및 산성 전사 활성화 도메인을 함유한다 (개괄적으로 Schornack S, *et al* (2006) *J Plant Physiol* 163(3):256-272 참조). 또한, 병원성 박테리아 랄스토니아 솔라나세아룸 (*Ralstonia solanacearum*)에서, brg11 및 hpx17로 표시된 2개의 유전자가 R. 솔라나세아룸 바이오바르 1 스트레인 GMI1000 및 바이오바르 4 스트레인 RS1000에서 크산토모나스의 AvrBs3 패밀리에 상동하는 것으로 밝혀졌다 (Heuer *et al* (2007) *App1* 및

Envir Micro 73(13):4379-4384 참조). 이들 유전자는 서로에 대해 뉴클레오티드 서열에서 98.9% 동일하지만, hpx17의 반복 도메인에서 1,575 염기쌍의 손실만큼 상이하다. 그러나, 두 유전자 생성물은 크산토모나스의 AvrBs3 패밀리 단백질과 40% 미만의 서열 동일성을 가진다. 예컨대 전체 내용이 본원에 참조로 포함되는 미국 특허 번호 8,586,526 참조.

[0087] 이들 TAL 이펙터의 특이성은 텐덤 반복부에서 발견되는 서열들에 좌우된다. 반복된 서열은 대략 102 염기쌍을 포함하고 반복부는 전형적으로 서로와 91 내지 100% 상동한다 (Bonas et al, 상기 동일). 반복부의 다형태는 보통 위치 12 및 13에 위치하고 위치 12 및 13에서의 초가변 2잔기 (RVD)와 TAL-이펙터의 표적 서열의 연속적인 뉴클레오티드의 동일성과는 일-대-일 대응인 것으로 나타난다 (Moscou 및 Bogdanove, (2009) Science 326:1501 및 Boch et al (2009) Science 326:1509-1512 참조). 실험적으로, 이들 TAL-이펙터의 DAN 인지에 대한 천연 코드는 위치 12 및 13에서의 HD 서열이 사이토신 (C)에 대한 결합을 유도하고, NG는 T에 결합하며, NI는 A, C, G 또는 T에 결합하고, NN은 A 또는 G에 결합하며, ING는 T에 결합하도록 결정되었다. 이들 DNA 결합 반복부는 반복부의 새로운 조합 및 수로 단백질로 조립되어 새로운 서열들과 상호작용하고 비-내인성 리포터 유전자의 식물 세포에서의 발현을 활성화할 수 있는 인공 전사 인자들이 만들어진다 (Boch et al, 상기 동일). 엔지니어링된 TAL 단백질은 *FokI* 절단 하프 도메인에 연결되어 TAL 이펙터 도메인 뉴클레아제 융합 (TALEN)이 제조된다. 예컨대 미국 특허 번호 8,586,526; Christian et al ((2010) <Genetics> epub 10.1534/genetics.110.120717)을 참조한다. 특정 구체예에서, TALE 도메인은 미국 특허 번호 8,586,526에 기술된 것과 같이 N-캡 및/또는 C-캡을 포함한다.

[0088] 특정 구체예에서, 세포의 계놈의 생체 내 절단 및/또는 표적화된 절단에 사용된 하나 이상의 뉴클레아제의 DNA 결합 도메인은 징크 평거 단백질을 포함한다. 바람직하게, 징크 평거 단백질은 그것이 선택된 표적 부위에 결합하도록 엔지니어링된다는 점에서 자연적으로 발생하는 것이 아니다. 예를 들어 Beerli et al. (2002) *Nature Biotechnol.* 20:135-141; Pabo et al. (2001) *Ann. Rev. Biochem.* 70:313-340; Isalan et al. (2001) *Nature Biotechnol.* 19:656-660; Segal et al. (2001) *Curr. Opin. Biotechnol.* 12:632-637; Choo et al. (2000) *Curr. Opin. Struct. Biol.* 10:411-416; 미국 특허 번호 6,453,242; 6,534,261; 6,599,692; 6,503,717; 6,689,558; 7,030,215; 6,794,136; 7,067,317; 7,262,054; 7,070,934; 7,361,635; 7,253,273; 및 미국 특허 공개공보 번호 2005/0064474; 2007/0218528; 2005/0267061 (이것들은 모두 전체 내용이 본원에 참조로 포함됨)을 참조한다.

[0089] 엔지니어링된 징크 평거 결합 도메인은 자연적으로-발생하는 징크 평거 단백질과 비교하여 신규한 결합 특이성을 가질 수 있다. 엔지니어링 방법으로는, 제한하는 것은 아니지만, 합리적 디자인 및 다양한 유형의 선택을 포함한다. 합리적 디자인은, 예를 들면 삼중 (또는 사중) 뉴클레오티드 서열 및 개별적인 징크 평거 아미노산 서열을 포함하는 데이터베이스를 사용하는 것을 포함하는데, 여기서 각각의 삼중 또는 사중 뉴클레오티드 서열은 특정 삼중 또는 사중 서열에 결합하는 징크 평거의 하나 이상의 아미노산 서열과 회합된다. 예를 들면 공동-소유의 미국 특허 6,453,242 및 6,534,261 (전체 내용이 본원에 참조로 포함됨)을 참조한다.

[0090] 파지 디스플레이 및 2-하이브리드 시스템을 포함하여, 예시적인 선택 방법은 미국 특허 5,789,538; 5,925,523; 6,007,988; 6,013,453; 6,410,248; 6,140,466; 6,200,759; 및 6,242,568; 또한 WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 및 GB 2,338,237에서 개시된다. 또한, 징크 평거 결합 도메인에 대한 결합 특이성의 증대는, 예를 들면 공동-소유의 WO 02/077227에 기술되어 있다.

[0091] 또한, 이들 및 다른 참조문헌에서 개시된 것과 같이, 징크 평거 도메인 및/또는 다중-평거의 징크 평거 단백질은 어떠한 적합한 링커 서열, 이를테면 예를 들어 5 이상의 아미노산 길이의 링커를 사용하여 함께 연결될 수 있다. 또한 6 이상의 아미노산 길이의 예시적인 링커 서열에 대해서는 미국 특허 번호 6,479,626; 6,903,185; 및 7,153,949를 참조한다. 본원에 기술된 단백질은 단백질의 개별적인 징크 평거 사이의 적합한 링커들의 임의의 조합을 포함할 수 있다.

[0092] 융합 단백질 (및 그것을 코드화하는 폴리뉴클레오티드)의 디자인 및 구성을 위한 표적 부위 및 방법의 선택은 기술분야의 숙련자들에게 알려져 있고 미국 특허 번호 6,140,081; 5,789,538; 6,453,242; 6,534,261; 5,925,523; 6,007,988; 6,013,453; 6,200,759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197; WO 02/099084; WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 및 WO 03/016496에서 상세하게 설명된다.

[0093] 또한, 이들 및 다른 참조문헌에서 개시된 것과 같이, 징크 평거 도메인 및/또는 다중-평거의 징크 평거 단백질은 어떠한 적합한 링커 서열, 이를테면 예를 들어 5 이상의 아미노산 길이의 링커를 사용하여 함께 연결될 수

있다. 또한 6 이상의 아미노산 길이의 예시적인 링커 서열에 대해서는 미국 특허 번호 6,479,626; 6,903,185; 및 7,153,949를 참조한다. 본원에 기술된 단백질은 단백질의 개별적인 징크 평거 사이의 적합한 링커들의 임의의 조합을 포함할 수 있다.

[0094] 특정 구체예에서, DNA-결합 도메인은, 예를 들면 단일 가이드 RNA (sgRNA)를 포함하는, CRISPR/Cas 뉴클레아제 시스템의 일부이다. 예컨대 미국 특허 번호 8,697,359 및 미국 특허 공개공보 번호 20150056705를 참조한다. 시스템의 RNA 성분을 코드화하는 CRISPR (주기적으로 간격을 띠고 분포하는 짧은 회문 구조 반복서열) 유전자좌, 및 단백질을 코드화하는 cas (CRISPR-회합됨) 유전자좌 (Jansen *et al.*, 2002. *Mol. Microbiol.* 43:1565-1575; Makarova *et al.*, 2002. *Nucleic Acids Res.* 30:482-496; Makarova *et al.*, 2006. *Biol. Direct* 1:7; Haft *et al.*, 2005. *PLoS Comput. Biol.* 1:e60)는 CRISPR/Cas 뉴클레아제 시스템의 유전자 서열을 구성한다. 미생물 숙주에서의 CRISPR 유전자좌는 CRISPR-매개된 핵산 절단의 특이성을 프로그래밍할 수 있는 비-코딩 RNA 요소뿐 아니라 CRISPR-회합됨 (Cas) 유전자좌의 조합을 함유한다.

[0095] 타입 II CRISPR은 가장 잘 특성화된 시스템 중 하나이고 표적화된 DNA 이중-가닥 파괴를 4개의 순차적 단계로 수행한다. 첫째, 2개의 비-코딩 RNA, 프리-crRNA 어레이 및 tracrRNA가 CRISPR 유전자좌로부터 전사된다. 둘째, tracrRNA는 프리-crRNA의 반복 영역에 하이브리드를 형성하고 개별적인 스페이서 서열을 함유하는 성숙한 crRNA로의 프리-crRNA의 프로세싱을 매개한다. 셋째, 성숙한 crRNA:tracrRNA 복합체는 crRNA 상의 스페이서와 표적 인식을 위한 추가의 요건인 프로토스페이서 인접 모티프 (PAM) 다음에 있는 표적 DNA 상의 프로토스페이서 사이의 왓슨-크릭 염기쌍 형성을 통해 표적 DNA에 대해 Cas9를 지시한다. 마지막으로, Cas9는 프로토스페이서 내에 이중-가닥 파괴를 생성하기 위해 표적 DNA의 절단을 매개한다. CRISPR/Cas 시스템의 활성은 3가지 단계를 포함한다: (i) '적응'으로 불리는 과정에서, 미래의 공격을 방지하기 위해 외부 DNA 서열의 CRISPR 어레이로의 삽입, (ii) 관련된 단백질의 발현, 뿐만 아니라 어레이의 발현 및 프로세싱, 이어서 (iii) 외부 핵산으로의 RNA-매개된 간섭. 그러므로, 박테리아 세포에서, 여러 개의 소위 'Cas' 단백질은 CRISPR/Cas 시스템의 천연 기능과 포함되고 외부 DNA 등의 삽입과 같은 기능에서 역할을 한다.

[0096] 특정 구체예에서, Cas 단백질은 자연적으로 발생하는 Cas 단백질의 "기능적 유도체"이다. 천연 서열 폴리펩티드의 "기능적 유도체"는 천연 서열 폴리펩티드와 공통의 정성적 생물학적 특성을 가지는 화합물이다. "기능적 유도체"는, 제한은 아니지만, 천연 서열의 단편 및 천연 서열 폴리펩티드 및 그것의 단편의 유도체를 포함하며, 단 그것들은 상응하는 천연 서열 폴리펩티드와 공통의 생물학적 특성을 가진다. 본원에서 고려되는 생물학적 활성이란 DNA 기질을 단편으로 가수분해하는 기능적 유도체의 능력이다. 용어 "유도체"는 폴리펩티드, 공유 변형의 아미노산 서열 변이체, 및 그것들의 융합 돌사를 포함한다. Cas 폴리펩티드 또는 그것의 단편의 적합한 유도체는, 제한하는 것은 아니지만, Cas 단백질 또는 그것의 단편의 돌연변이체, 융합물, 공유 변형을 포함한다. Cas 단백질 또는 그것의 단편, 뿐만 아니라 Cas 단백질 또는 그것의 단편의 유도체를 포함하는 Cas 단백질은 세포로부터 얻을 수 있거나 화학적으로 합성될 수 있고 또는 이 두 가지 과정의 조합에 의해 얻을 수 있다. 세포는 자연적으로 Cas 단백질을 생성하는 세포, 또는 Cas 단백질을 자연적으로 생성하고 더 높은 발현 수준으로 내인성 Cas 단백질을 생성하기 위해 또는 외부적으로 도입된 핵산으로부터 Cas 단백질을 생성하기 위해 유전학적으로 엔지니어링되는 세포일 수 있고, 이때 핵산은 내인성 Ca와 동일하거나 상이한 Cas를 코드화한다. 일부 경우에, 세포는 Cas 단백질을 자연적으로 생성하지 않고 Cas 단백질을 생성하기 위해 유전학적으로 엔지니어링된다.

[0097] 일부 구체예에서, DNA 결합 도메인은 TtAgo 시스템의 일부이다 (Swarts *et al.*, 상기 동일; Sheng *et al.*, 상기 동일, 참조). 진핵생물에서, 유전자 침묵은 단백질의 아르고노트 (Ago) 패밀리에 의해 매개된다. 이 패러다임에서, Ago는 작은 (19 내지 31 nt) RNA에 결합된다. 이 단백질-RNA 침묵 복합체는 작은 RNA와 표적 사이의 왓슨-크릭 염기쌍을 통해 표적 RNA를 인식하고 표적 RNA를 핵산 내에서 가수분해적으로 절단한다 (Vogel (2014) *Science* 344:972-973). 대조적으로, 원핵 Ago 단백질은 작은 단일-가닥 DNA 단편에 결합하고 외래 (자주 바이러스) DNA를 감지하고 제거하는 기능을 하는 것으로 보인다 (Yuan *et al.*, (2005) *Mol. Cell* 19, 405; Olovnikov, *et al.* (2013) *Mol. Cell* 51, 594; Swarts *et al.*, 상기 동일). 예시적인 원핵 Ago 단백질로는 아퀴펙스 아에롤리커스 (*Aquifex aeolicus*), 로도박터 스파에로이데스 (*Rhodobacter sphaeroides*) 및 써모필러스 (*Thermus thermophilus*)로부터의 것들이 있다.

[0098] 특성화가 가장 잘 이루어져 있는 원핵 Ago 단백질 중 한 가지는 T. 써모필러스로부터의 단백질이다 (TtAgo; Swarts *et al.* 상기 동일). TtAgo는 5' 포스페이트 기를 가지는 15 nt 또는 13 내지 25 nt의 단일-가닥 DNA 단편과 회합한다. TtAgo에 의해 결합된 "가이드 DNA"는 단백질-DNA 복합체가 왓슨-크릭 상보성 DNA 서열에 DNA의 제3자 문자로 결합하도록 지시하는 작용을 한다. 일단 이들 가이드 DNA의 서열 정보가 표적 DNA의 확인을 허용

하고 나면, TtAgo-가이드 DNA 복합체가 표적 DNA를 절단한다. 그런 메커니즘은 또한 표적 DNA에 결합되어 있는 동안 TtAgo-가이드 DNA 복합체의 구조에 의해 지지된다 (G. Sheng *et al.*, 상기 동일). 로도박터 스파에로이데스로부터의 Ago (RsAgo)가 유사한 특성을 가진다 (Olivnikov *et al.*, 상기 동일).

[0099] 임의적인 DNA 서열의 외인성 가이드 DNA가 TtAgo 단백질 상에 로딩될 수 있다 (Swarts *et al.*, 상기 동일). TtAgo 절단의 특이성이 가이드 DNA에 의해 지시되기 때문에, 따라서 외인성, 조사자-명시된 가이드 DNA로 형성된 TtAgo-DNA 복합체는 상보하는 조사자-명시된 표적 DNA에 대한 직접적인 TtAgo 표적 DNA 절단을 지시할 것이다. 이런 방법으로, 당업자는 DNA에 표적화된 이중-가닥 파괴를 생성할 수 있다. TtAgo-가이드 DNA 시스템 (또는 다른 유기체로부터의 이종상동성 Ago-가이드 DNA 시스템)은 세포 내에서 게놈 DNA의 표적화된 절단을 허용한다. 그런 절단은 단일- 또는 이중-가닥 중 어느 하나일 수 있다. 포유류 게놈 DNA의 절단에 대해, 포유류 세포에서의 발현에 대해 최적화된 TtAgo 코돈의 버전을 사용하는 것이 바람직할 것이다. 나아가, 시험관 내 형성된 TtAgo-DNA 복합체로 세포를 처리하는 것이 바람직할 것이고, 이때 TtAgo 단백질이 세포-침투 햅틱드에 융합된다. 나아가, 돌연변이생성을 통해 37°C에서 개선된 활성을 가지도록 변경되어 있는 TtAgo 단백질의 버전을 사용하는 것이 바람직할 것이다. Ago-RNA-매개된 DNA 절단은 DNA 파괴의 외삽을 위해 기술분야의 표준 기법을 사용하는 유전자 녹-아웃, 표적화된 유전자 침가, 유전자 교정, 표적화된 유전자 결실을 포함하는 결과의 모음에 영향을 주기 위해 사용될 수 있을 것이다.

[0100] 그러므로, 뉴클레아제는 도너 (전이유전자)를 삽입하는 것이 바람직한 임의의 유전자에서 표적 부위에 특이적으로 결합하는 DNA-결합 도메인을 포함한다.

B. 절단 도메인

[0102] 임의의 적합한 절단 도메인은 DNA-결합 도메인에 작동 가능하게 연결되어 뉴클레아제를 형성한다. 예를 들어, ZFP DNA-결합 도메인은 뉴클레아제 도메인에 융합되어 ZFN - 엔지니어링된 (ZFP) DNA 결합 도메인을 통해 의도된 핵산 표적을 인식할 수 있고 다양한 유기체에서 게놈 변형에 사용하기 위한 것을 포함하여, 뉴클레아제 활성을 통해 ZFP 결합 부위 가까이에서 DNA가 절단되는 것을 유발할 수 있는 기능적 실체임 -을 형성한다. 예를 들어, 미국 특허 공보 20030232410; 20050208489; 20050026157; 20050064474; 20060188987; 20060063231; 및 국제 공보 WO 07/014275를 참조한다. 마찬가지로, TALE DNA-결합 도메인은 뉴클레아제 도메인에 융합되어 TALEN이 형성된다. 예컨대 미국 특허 번호 8,586,526을 참조한다. DNA에 결합하고 절단 도메인 (예컨대 Cas 도메인)과 회합되어 표적화된 절단을 유도하는 단일 가이드 RNA (sgRNAs)를 포함하는 CRISPR/Cas 뉴클레아제 시스템이 또한 기술되어 있다. 예컨대 미국 특허 번호 8,697,359 및 8,932,814 및 미국 특허 공보 번호 20150056705를 참조한다.

[0103] 상기에서 주지한 바, 절단 도메인은 DNA-결합 도메인, 예를 들면: 뉴클레아제로부터의 징크 평거 DNA-결합 도메인 및 절단 도메인; 뉴클레아제로부터의 TALEN DNA-결합 도메인 및 절단 도메인; 뉴클레아제 (CRISPR/Cas)로부터의 sgRNA DNA-결합 도메인 및 절단 도메인; 및/또는 상이한 뉴클레아제로부터의 메가뉴클레아제 DNA-결합 도메인 및 절단 도메인에 대해 이종성일 수 있다. 이종성 절단 도메인은 임의의 엔도뉴클레아제 또는 엑소뉴클레아제로부터 얻어질 수 있다. 절단 도메인이 유도될 수 있는 예시적인 엔도뉴클레아제는, 제한하는 것은 아니지만, 제한 엔도뉴클레아제 및 호밍 엔도뉴클레아제를 포함한다. DNA를 절단하는 추가의 효소들이 알려져 있다 (예컨대 S1 뉴클레아제; 녹두 뉴클레아제; 쿼장 DNase I; 미구균 뉴클레아제; 효모 HO 엔도뉴클레아제. 이들 효소 중 하나 이상 (또는 그것의 기능적 단편)이 절단 도메인 및 절단 하프-도메인의 공급원으로서 사용될 수 있다.

[0104] 유사하게, 상기에서 설명된 것과 같이, 절단 활성을 위해 이량체화가 필요한 절단 하프-도메인은 임의의 뉴클레아제 또는 그것의 부분으로부터 유도될 수 있다. 일반적으로, 융합 단백질이 절단 하프-도메인을 포함한다면 두 개의 융합 단백질이 절단에 필요하다. 다르게는, 두 개의 절단 하프-도메인을 포함하는 단일 단백질이 사용될 수 있다. 두 개의 절단 하프-도메인은 동일한 엔도뉴클레아제 (또는 그것의 기능적 단편)으로부터 유도되거나, 또는 각각의 절단 하프-도메인이 상이한 엔도뉴클레아제 (또는 그것의 기능적 단편)으로부터 유도될 수 있다. 또한, 두 개의 융합 단백질에 대한 표적 부위는 바람직하게는 서로에 대해, 두 개의 융합 단백질의 그것들의 각각의 표적 부위에 대한 결합이 절단 하프-도메인이 예컨대 이량체화에 의해 기능적 절단 도메인을 형성하는 것을 허용하는, 서로에 대해 공간적 배향으로 절단 하프-도메인을 위치시키도록 배치된다. 그러므로, 특정 구체예에서, 표적 부위의 모서리 근처는 5 내지 8 뉴클레오티드 또는 15 내지 18 뉴클레오티드만큼 분리된다. 그러나 임의의 정수 수의 뉴클레오티드 또는 뉴클레오티드 쌍 (예컨대 2 내지 50 뉴클레오티드 또는 그 이상의 쌍)이 두 개의 표적 부위 사이에 개재할 수 있다. 일반적으로, 절단 부위는 표적 부위들 사이에 있다.

- [0105] 제한 엔도뉴클레아제 (제한 효소)는 많은 종들에 존재하고 DNA에 서열-특이적으로 결합할 수 있고 (인식 부위에), 그리고 결합 부위에서 또는 가까이에서 DNA를 절단할 수 있다. 특정 제한 효소 (예컨대 Type IIS)는 인식 부위로부터 제거된 부위들에서 DNA를 절단하고 분리 가능한 결합 및 절단 도메인을 가진다. 예를 들어, 타입 IIS 효소 *Fok* I는 한 가닥 상에서 그것의 인식 부위로부터 9 뉴클레오티드에서 및 다른 가닥 상에서 그것의 인식 부위로부터 13 뉴클레오티드에서, DNA의 이중-가닥의 절단을 촉매한다. 예를 들면 미국 특허 5,356,802; 5,436,150 및 5,487,994; 뿐만 아니라 Li et al. (1992) Proc. Natl. Acad. Sci. USA **89**:4275-4279; Li et al. (1993) Proc. Natl. Acad. Sci. USA **90**:2764-2768; Kim et al. (1994a) Proc. Natl. Acad. Sci. USA **91**:883-887; Kim et al. (1994b) J. Biol. Chem. **269**:31,978-31,982를 참조한다. 그러므로, 한 구체예에서, 융합 단백질은 적어도 하나의 타입 IIS 제한 효소 및 하나 이상의 징크 핑거 결합 도메인으로부터의, 엔지니어링되거나 또는 엔지니어링되지 않을 수 있는 절단 도메인 (또는 절단 하프-도메인)을 포함한다.
- [0106] 절단 도메인이 결합 도메인으로부터 분리 가능한, 예시적인 타입 IIS 제한 효소는 *Fok* I이다. 이 특정 효소는 이량체로서 활성이다. Bitinaite et al. (1998) Proc. Natl. Acad. Sci. USA **95**:10,570-10,575. 따라서, 본 발명의 목적에 대해, 개시된 융합 단백질에서 사용된 *Fok* I 효소의 부분은 절단 하프-도메인인 것으로 여겨진다. 그러므로, 징크 핑거-*Fok* I 융합을 사용하는 세포 서열의 표적화된 이중-가닥 절단 및/또는 표적화된 대체에 대해, 각각이 *FokI* 절단 하프-도메인을 포함하는 2개의 융합 단백질이 촉매적으로 활성 절단 도메인을 복원시키기 위해 사용될 수 있다. 다르게는, 징크 핑거 결합 도메인 및 2개의 *Fok* I 절단 하프-도메인을 함유하는 단일 폴리펩티드 분자가 또한 사용될 수 있다. 징크 핑거-*Fok* I 융합을 사용하는 표적화된 절단 및 표적화된 서열 변경에 대한 변수들이 본 발명의 다른 곳에서 제시된다.
- [0107] 절단 도메인 또는 절단 하프-도메인은 절단 활성을 보유하는, 또는 기능적 절단 도메인을 형성하기 위해 다량체화 (예컨대 이량체화)되는 능력을 보유하는 단백질의 임의의 부분일 수 있다.
- [0108] 예시적인 Type IIS 제한 효소는 전체 내용이 본원에 포함되어 있는, 국제 공보 WO 07/014275에서 기술된다. 추가의 제한 효소는 또한 분리 가능한 결합 및 절단 도메인을 함유하고, 이것들은 본 발명에 의해 고려된다. 예를 들면 Roberts et al. (2003) Nucleic Acids Res. **31**:418-420을 참조한다.
- [0109] 특정 구체예에서, 절단 도메인은 예를 들면 미국 특허 공보 번호 20050064474; 20060188987; 20070305346 및 20080131962 (전체 내용이 본원에 참조로 포함됨)에 기술되어 있는 것과 같이, 동종이량체화를 최소화하거나 방지하는 하나 이상의 엔지니어링된 절단 하프-도메인 (또한 이량체화 도메인 돌연변이체로서 언급됨)을 포함한다. *FokI*의 위치 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537, 및 538에 있는 아미노산 잔기들이 *FokI* 절단 하프-도메인의 이량체화에 영향을 주기 위한 모든 표적이다.
- [0110] 하나 이상의 돌연변이, 예를 들면 "E490K:I538K"로 표시된 엔지니어링된 절단 하프-도메인을 생성하기 위한 한 절단 하프-도메인에서 위치 490 (EK) 및 538 (I→K)에서의 돌연변이 및 "Q486E:I499L"로 표시된 엔지니어링된 절단 하프-도메인을 생성하기 위한 다른 절단 하프-도메인에서 위치 486 (QE) 및 499 (IL)의 돌연변이에 의한 돌연변이; 위치 486에서 야생형 Gln (Q) 잔기를 Glu (E) 잔기로, 위치 499에서 야생형 Iso (I)를 Leu (L) 잔기로 및 위치 496에서 야생형 Asn (N) 잔기를 Asp (D) 또는 Glu (E) 잔기 (또한 각각 "ELD" 및 "ELE" 도메인으로 언급됨)로 대체하는 돌연변이; 위치 490, 538 및 537 (야생형 *FokI*에 대한 숫자)에서의 돌연변이를 포함하는 엔지니어링된 절단 하프-도메인, 예를 들면 위치 490에서 야생형 Glu (E) 잔기를 Lys (K) 잔기로, 위치 538에서 야생형 Iso (I) 잔기를 Lys (K) 잔기로, 및 위치 537에서 야생형 His (H) 잔기를 Lys (K) 잔기 또는 Arg (R) 잔기 (또한 각각 "KKK" 및 "KKR" 도메인으로 언급됨)로 대체하는 돌연변이를 가지는 절단 도메인이 사용될 수 있고; 및/또는 엔지니어링된 절단 하프-도메인은 위치 490 및 537 (야생형 *FokI*에 대한 숫자)에서의 돌연변이, 예를 들면 위치 490에서 야생형 Glu (E) 잔기를 Lys (K) 잔기로 및 위치 537에서 야생형 His (H) 잔기를 Lys (K) 잔기 또는 Arg (R) 잔기 (또한 각각 "KIK" 및 "KIR" 도메인으로 언급됨)로 대체하는 돌연변이를 포함한다. 예컨대 미국 특허 번호 7,914,796; 8,034,598 및 8,623,618을 참조하고, 이것들은 모든 목적에 대해 전체 내용이 참조로 포함된다. 다른 구체예에서, 엔지니어링된 절단 하프 도메인은 "Sharkey" 및/또는 "Sharkey" 돌연변이를 포함한다 (Guo et al., (2010) J. Mol. Biol. 400(1):96-107 참조).
- [0111] 다르게는, 뉴클레아제는 생체 내에서 소위 "스플릿-효소" 기술을 사용하여 핵산 표적 부위에서 조립될 수 있다 (예컨대 미국 특허 공보 번호 20090068164 참조). 그런 스플릿 효소의 성분들은 별도의 발현 구성물에서 발현되거나, 또는 개별적인 성분들이 예를 들어 자기-절단 2A 웨터드 또는 IRES 서열에 의해 분리되는 하나의 오픈 리딩 프레임에서 연결될 수 있다. 성분들은 개별적인 징크 핑거 결합 도메인 또는 메가뉴클레아제 핵산 결합 도메인의 도메인일 수 있다.

- [0112] 뉴클레아제는 예를 들면 미국 특허 번호 8,563,314에서 기술된 것과 같이 효모-기반 염색체 시스템에서 사용 전에 활성에 대해 스크리닝될 수 있다.
- [0113] Cas9 관련 CRISPR/Cas 시스템은 두 개의 RNA 비-코딩 성분: tracrRNA 및 동일한 직접 반복부 (DR)에 의해 사이에 배치된 뉴클레아제 가이드 서열 (스페이서)을 함유하는 프리-crRNA 배열을 포함한다. 게놈 엔지니어링을 이루기 위하여 CRISPR/Cas 시스템을 사용하기 위하여, 이들 RNA의 두 기능이 존재해야 한다 (Cong *et al.*, (2013) *Scienceexpress* 1/10.1126/science 1231143 참조). 일부 구체예에서, tracrRNA 및 pre-crRNA는 별도의 발현 구성물을 통해 또는 별도의 RNA로서 공급된다. 다른 구체예에서, 엔지니어링된 성숙한 crRNA (표적 특이성을 부여함)가 tracrRNA (Cas9와의 상호작용을 공급함)에 융합되어 키메라 cr-RNA-tracrRNA 하이브리드 (또한 단일 가이드 RNA로 언급됨)가 생성되는 키메라 RNA가 구성된다. (Jinek, 상기 동일 및 Cong, 상기 동일, 참조).
- 표적 부위**
- [0115] 상기에서 상세하게 기술된 것과 같이, DNA-결합 도메인은 선택된 임의의 서열을 결합시키기 위해 엔지니어링될 수 있다. 엔지니어링된 DNA-결합 도메인은 자연적으로-발생하는 DNA-결합 도메인에 비교하여 신규한 결합 특이성을 가질 수 있다.
- [0116] 적합한 표적 유전자의 비-제한적인 실례는 베타 (β) 글로빈 유전자 (HBB), 감마 (γ) 글로빈 유전자 (HBG1), B-세포 램프종/백혈병 11A (BCL11A) 유전자, Kruppel-유사 인자 1 (KLF1) 유전자, CCR5 유전자, CXCR4 유전자, PPP1R12C (AAVS1) 유전자, 하이포크산틴 포스포리보실트란스페라제 (HPRT) 유전자, 알부민 유전자, 인자 VIII 유전자, 인자 IX 유전자, 류신-풍부 반복 키나아제 2 (LRRK2) 유전자, 헌팅턴 (Htt) 유전자, 로돕신 (RHO) 유전자, 낭성 섬유증 경막 전도성 조절제 (CFTR) 유전자, 계면활성제 단백질 B 유전자 (SFTPB), T-세포 수용체 알파 (TRAC) 유전자, T-세포 수용체 베타 (TRBC) 유전자, 프로그램된 세포 사망 1 (PD1) 유전자, 세포독성 T-림프구 항원 4 (CTLA-4) 유전자, 인간 백혈구 항원 (HLA) A 유전자, HLA B 유전자, HLA C 유전자, HLA-DPA 유전자, HLA-DQ 유전자, HLA-DRA 유전자, LMP7 유전자, 항원 프로세싱과 관련된 수송체 (TAP) 1 유전자, TAP2 유전자, 타파신 유전자 (TAPBP), 부류 II 주요 조직적합성 복합체 전이활성제 (CIITA) 유전자, 디스트로핀 유전자 (DMD), 글루코코르티코이드 수용체 유전자 (GR), IL2RG 유전자, Rag-1 유전자, RFX5 유전자, FAD2 유전자, FAD3 유전자, ZP15 유전자, KASII 유전자, MDH 유전자, 및/또는 EPSPS 유전자이다.
- [0117] 특정 구체예에서, 뉴클레아제는 인간 세포에서 AAVS1, HPRT, 알부민 및 CCR5, 및 쥐과 세포에서 Rosa26과 같은 유전자좌 (예컨대 미국 특허 번호 7,888,121; 7,972,854; 7,914,796; 7,951,925; 8,110,379; 8,409,861; 8,586,526; 미국 특허 공보 20030232410; 20050208489; 20050026157; 20060063231; 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983 및 20130177960 참조) 및 식물에서 Zp15 유전자좌 (미국 특허 U.S. 8,329,986 참조)와 같은 "안전한 은닉" 유전자좌를 표적으로 한다.
- 도너**
- [0119] 특정 구체예에서, 본 발명은 줄기세포의 게놈의 뉴클레아제-매개된 변형에 관련된다. 상기에서 주지된 것과 같이, 예를 들어 명시된 영역의 결실 및/또는 돌연변이 유전자의 결실 또는 야생형 유전자의 증가된 발현을 위한 외인성 서열 (또한 "도너 서열" 또는 "도너" 또는 "전이유전자"로도 불립)의 삽입. 도너 서열은 전형적으로 그것이 놓이는 게놈 서열에 동일하지 않은 것이 쉽게 드러날 것이다. 도너 서열은 관심의 위치에서 효과적인 HDR을 허용하기 위해 두 개의 상동성 영역이 양 옆에 있는 비-상동성 서열을 함유하거나 또는 비-상동성 지시된 수복 메커니즘을 통해 통합될 수 있다. 추가로, 도너 서열은 세포 크로마틴에서 관심의 영역에 상동하지 않는 서열을 함유하는 벡터 분자를 포함할 수 있다. 도너 분자는 세포 크로마틴에 대한 여러, 불연속적인 상동성 영역을 함유할 수 있다. 나아가, 관심 영역에 정상적으로 존재하지 않는 서열의 표적화된 삽입에 대해, 상기 서열은 도너 핵산 분자에 존재하고 관심 영역의 서열에 대한 상동성 영역이 양 옆에 있다.
- [0120] 뉴클레아제와 같이, 도너는 임의의 형태로 도입될 수 있다. 특정 구체예에서, 도너는 변형된 세포에서 잔류 바이러스를 제거하기 위해 mRNA 형태로 도입된다. 다른 구체예에서, 도너는 기술분야에 공지된 방법들에 의해 DNA 및/또는 바이러스 벡터를 사용하여 도입될 수 있다. 예컨대 미국 특허 공보 번호 20100047805 및 20110207221을 참조한다. 도너는 세포 안으로 원형 또는 선형 형태로 도입될 수 있다. 만약 선형 형태로 도입된다면, 도너 서열의 단부들은 기술분야의 숙련자들에 공지된 방법에 의하여 (예컨대 핵산 외적으로 가부순해적 분해로부터) 보호될 수 있다. 예를 들어, 하나 이상의 디데옥시뉴클레오티드 잔기가 선형 분자의 3' 말단에 첨가되고 및/또는 자기-상보적인 올리고뉴클레오티드가 하나 또는 둘 다의 단부에 결찰된다. 예를 들면 Chang *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:4959-4963; Nehls *et al.* (1996) *Science* 272:886-889를 참조한다. 외인성 폴

리뉴클레오티드를 분해로부터 보호하는 추가의 방법으로는, 제한하는 것은 아니지만, 말단 아미노기(들)의 첨가 및 변형된 뉴클레오티드내 연쇄, 예를 들면 포스포로티오에이트, 포스포라미데이트 및 0-메틸 리보오스 또는 데옥시리보오스 잔기의 사용을 포함한다.

[0121] 특정 구체예에서, 도너는 길이가 1 kb보다 큰, 예를 들면 2 내지 200 kb, 2 내지 10 kb (또는 그 사이의 임의의 값)의 서열 (예컨대 코딩 서열, 또한 전이유전자로서 언급됨)을 포함한다. 도너는 또한 적어도 하나의 뉴클레아제 표적 부위를 포함할 수 있다. 특정 구체예에서, 도너는 예를 들어 ZFN, TALEN, TtAgo 또는 CRISPR/Cas 뉴클레아제의 쌍에 대해, 적어도 2개의 표적 부위를 포함한다. 전형적으로, 뉴클레아제 표적 부위는 전이유전자의 절단에 대해, 전이유전자 서열 외부에, 예를 들면 전이유전자 서열에 대해 5' 및/또는 3'에 있다. 뉴클레아제 절단 부위(들)은 임의의 뉴클레아제(들)에 대한 것일 수 있다. 특정 구체예에서, 이중-가닥 도너에 함유된 뉴클레아제 절단 부위(들)은 절단된 기증자가 상동성-무관한 방법을 통해 통합되는 내인성 표적을 절단하기 위해 사용된 동일한 뉴클레아제(들)에 대한 것이다.

[0122] 도너는 그것의 발현이 통합 부위에서 내인성 프로모터, 즉 기증자가 삽입되는 내인성 유전자의 발현을 구동하는 프로모터에 의해 구동되도록 삽입될 수 있다. 그러나, 도너가 프로모터 및/또는 인핸서, 예를 들면 구성성 프로모터 또는 유도성 또는 조직 특이적 프로모터를 포함할 수 있다는 것이 드러날 것이다. 도너 분자는 모든, 일부의 내인성 유전자가 발현되거나 전혀 발현되지 않도록 내인성 유전자 안으로 삽입될 수 있다. 나아가, 비록 발현에는 필요하지 않더라도, 내인성 서열은 또한 전사 또는 번역 조절 서열, 예를 들면 프로모터, 인핸서, 인슐레이터, 말단 리보솜 유입 부위, 2A 웨პ터드를 코드화하는 서열 및/또는 폴리아데닐화 신호를 포함한다.

[0123] 본원에 기술된 도너 서열상에 운반된 전이유전자는 플라스미드, 세포 또는 다른 공급원으로부터 PCR과 같이 기술분야에 공지된 표준 기법들을 사용하여 분리될 수 있다. 사용하기 위한 도너는 수퍼코일형 원형, 풀어진 원형, 선형 등을 포함하여, 달라지는 유형의 위상을 포함할 수 있다. 다르게는, 그것은 표준 올리고뉴클레오티드 합성 기법을 사용하여 화학적으로 합성될 수 있다. 또한, 도너는 메틸화되거나 메틸화가 결핍될 수 있다. 도너는 박테리아 또는 효모 인공 염색체 (BAC 또는 YAC)의 형태일 수 있다.

[0124] 본원에 기술된 도너 폴리뉴클레오티드는 하나 이상의 비-천연 염기 및/또는 골격을 포함할 수 있다. 특히, 메틸화된 사이토신을 가진 도너 분자의 삽입은 관심 영역에서 전사 정지 상태를 이루기 위하여 본원에 기술된 방법들을 사용하여 수행될 수 있다.

[0125] 외인성 (도너) 폴리뉴클레오티드는 임의의 관심의 서열 (외인성 서열)을 포함할 수 있다. 예시적인 외인성 서열로는, 제한하는 것은 아니지만, 임의의 폴리펩티드 코딩 서열 (예컨대 cDNA), 인핸서 서열, 애피토프 태그, 마커 유전자, 절단 효소 인식 부위 및 다양한 유형의 발현 구성물을 포함한다. 마커 유전자는, 제한하는 것은 아니지만, 항생물질 내성 (예컨대 암피실린 내성, 네오마이신 내성, G418 내성, 퓨로마이신 내성)을 매개하는 단백질을 코드화하는 서열, 착색된 또는 형광 또는 발광 단백질 (예컨대 녹색 형광 단백질, 증강된 녹색 형광 단백질, 적색 형광 단백질, 루시페라제), 및 증강된 세포 성장 및/또는 유전자 증폭을 매개하는 단백질 (예컨대 다이하이드로폴레이트 환원효소)을 코드화하는 서열을 포함한다. 애피토프 태그는, 예를 들면 FLAG의 하나 이상의 복사물, His, myc, Tap, HA 또는 임의의 검출 가능한 아미노산 서열을 포함한다.

[0126] 일부 구체예에서, 도너는 추가로 그것의 세포에서의 발현이 바람직한 임의의 폴리펩티드, 이를테면 제한하는 것은 아니지만, 항원, 효소, 수용체 (세포 표면 또는 핵의), 호르몬, 림포카인, 사이토카인, 리포터 폴리펩티드, 성장 인자 및 상기 중 임의의 기능적 단편을 코드화하는 폴리뉴클레오티드를 포함한다. 코딩 서열은 예를 들면 cDNA일 수 있다.

[0127] 특정 구체예에서, 외인성 서열은 표적화된 통합을 수행하는 세포의 선택을 허용하는 마커 유전자 (상기 기술됨), 및 추가의 기능성을 코드화하는 연결된 서열을 포함할 수 있다. 마커 유전자의 비-제한적인 실례는 GFP, 약물 선택 마커(들) 등을 포함한다.

[0128] 특정 구체예에서, 전이유전자는, 예를 들면 돌연변이된 내인성 서열을 대체하기 위한 야생형 유전자를 포함할 수 있다. 예를 들어, 야생형 (또는 다른 기능적) 유전자 서열은 유전자의 내인성 복사물이 돌연변이되는 줄기세포의 계놈 안으로 삽입될 수 있다. 전이유전자는 내인성 유전자좌에서 삽입되거나, 또는 다르게는 안전한 은닉 유전자좌에 표적화될 수 있다.

[0129] 본 명세서의 교시에 따르는, 그런 발현 카세트의 구성은 분자 생물학의 기술분야에 잘 알려져 있는 방법론들을 이용한다 (예를 들면 Ausubel 또는 Maniatis 참조). 형질도입 동물을 생성하기 위해 발현 카세트를 사용하기 전에, 발현 카세트의 선택된 제어 요소와 관련된 스트레스-유도제에 대한 반응성은 발현 카세트를 적합한 세포주

(예컨대 일차 세포, 형질전환된 세포 또는 불멸화된 세포주) 안에 도입시킴으로써 시험될 수 있다.

[0130] 나아가, 비록 발현에는 필요하지 않지만, 외인성 서열은 또한 전사 또는 번역 조절 서열, 예를 들면 프로모터, 인핸서, 인슬레이터, 내부 리보솜 유입 부위, 24 웨티드를 코드화하는 서열 및/또는 폴리아데닐화 신호일 수 있다. 나아가, 관심의 유전자의 제어 요소들은 키메라 유전자 (예컨대 리포터 발현 카세트)를 생성하기 위하여 리포터 유전자에 작동 가능하게 연결될 수 있다. 예시적인 스플라이스 수용체 부위 서열은 기술분야의 숙련자들에게 알려져 있고, 단지 예를 들면 CTGACCTCTCTTCCACAG, (SEQ ID NO:1)(인간 HBB 유전자로부터 유래함) 및 TTTCTCTCACAG (SEQ ID NO:2) (인간 면역글로불린-감마 유전자로부터 유래함)를 포함한다.

[0131] 비-코딩 핵산 서열의 표적화된 삽입이 또한 이루어질 수 있다. 안티센스 RNAs, RNAi, shRNAs 및 마이크로 RNA (miRNA)를 코드화하는 서열이 또한 표적화된 삽입을 위해 사용될 수 있다.

[0132] 추가의 구체예에서, 도너 핵산은 추가의 뉴클레아제 디자인에 특이적인 표적 서열인 비-코딩 서열을 포함할 수 있다. 계속해서, 추가의 뉴클레아제는 원래의 도너 분자가 절단되고 관심의 다른 도너 분자의 삽입에 의해 변형되도록 세포에서 발현될 수 있다. 이런 방법으로, 도너 분자의 반복되는 통합은 관심의 특정 유전자좌에서 또는 안전한 은닉 유전자좌에서 형질 결합을 허용하면서 생성될 수 있다.

세포

[0134] 그러므로, 본원에는 본원에 기술된 방법에 의해 생성된 세포를 포함하여, 유전학적으로 변형된 세포, 예를 들면 비활성화된 유전자 및/또는 전이유전자를 포함하는 줄기세포가 제공된다. 전이유전자는 표적화된 방식으로 하나 이상의 뉴클레아제를 사용하여 세포의 계놈 안에 통합된다. 무작위 통합과 달리, 표적화된 통합은 전이유전자가 명시된 유전자 안에 통합되는 것을 보장한다. 전이유전자는 표적 유전자에 어느 곳에서든 통합될 수 있다. 특정 구체예에서, 전이유전자는 뉴클레아제 결합 및/또는 절단 부위에 또는 가까이에, 예를 들면 절단 및/또는 결합 부위의 부위의 상류 또는 하류의 1 내지 300 (또는 그 사이의 임의의 염기쌍 수) 염기쌍 내에서, 보다 바람직하게 절단 및/또는 결합 부위의 어느 한 쪽의 1 내지 100 염기쌍 (또는 그 사이의 임의의 염기쌍 수) 염기쌍 내에서, 보다 더 바람직하게는 절단 및/또는 결합 부위의 어느 한 쪽의 1 내지 50 염기쌍 (또는 그 사이의 임의의 염기쌍 수) 염기쌍 내에서 통합된다. 특정 구체예에서, 통합된 서열은 임의의 벡터 서열 (예컨대 바이러스 벡터 서열)을 포함하지 않는다. 특정 구체예에서, 세포는 본원에 기술된 것과 같은 뉴클레아제에 의해 만들어진 변형 (예컨대 삽입 및/또는 결실)을 그 변형이 베타-글로빈 유전자의 엑손 내에, 예를 들면 엑손 1, 2 또는 3 내에 있도록 포함한다. 특정 구체예에서, 변형은 베타 글로빈 유전자의 엑손 1의 겹상-적혈구 돌연변이를 교정한다. 특정 구체예에서, 변형은 SEQ ID NO:23 또는 24에서 또는 가까이 (예컨대 1 내지 300 염기쌍 또는 그 사이의 임의의 수의 염기쌍)에서 일어난다. 다른 구체예에서, 변형은 SEQ ID NO:23 또는 24의 1 내지 100 (또는 그 사이의 임의의 수의 염기쌍)에서 일어난다. 특정 구체예에서, 변형은 SEQ ID NO:23 및/또는 SEQ ID NO:24의 내부에서 일어나는데, 예를 들면 하나 이상의 염기쌍의 변형은 SEQ ID NO:23 또는 24 중 어느 하나에서 일어난다.

[0135] 제한하는 것은 아니지만 세포 및 세포주를 포함하여, 어떠한 세포 유형이든지 본원에 기술된 것과 같이 유전학적으로 변형되어 전이유전자를 포함하게 된다. 본원에 기술된 것과 같은 유전학적으로 변형된 세포의 다른 비-제한적인 실례는 T-세포 (예컨대 CD4+, CD3+, CD8+, 등); 수지상 세포; B-세포; 자가조직 (예컨대 부모-유도된) 세포를 포함한다. 특정 구체예에서, 세포는 줄기세포, 이를테면 이종성 다능성, 전능성 또는 다중 전능성 줄기세포 (예컨대 CD34+ 세포, 유도된 다능 줄기세포 (iPSC), 배아 줄기세포 등)이다. 특정 구체예에서, 본원에서 기술되는 세포는 환자로부터 유도된 줄기세포이다.

[0136] 본원에 기술된 세포는 장애를 가진 대상에서 장애를, 예를 들면 생체 외 치료법에 의해 치료 및/또는 방지하는데 유용하다. 뉴클레아제-변형된 세포는 표준 기법을 사용하여 팽창된 후 환자에게 재도입될 수 있다. 예컨대 Tebas *et al* (2014) *New Eng J Med* 370(10):901을 참조한다. 줄기세포의 경우에, 대상에게 주입된 후, 이들 전 구체의 기능적 단백질을 (삽입된 도너로부터) 발현하는 세포로의 생체 내 분화가 또한 일어난다. 본원에 기술된 세포를 포함하는 약학 조성을 또한 제공된다. 또한, 세포는 환자에게 투여되기 전에 저온보존될 수 있다.

전달

[0138] 뉴클레아제, 이를 뉴클레아제를 코드화하는 폴리뉴클레오티드, 도너 폴리뉴클레오티드 및 본원에 기술된 단백질 및/또는 폴리뉴클레오티드를 포함하는 조성물은 임의의 적합한 수단에 의해 전달될 수 있다. 특정 구체예에서, 뉴클레아제 및/또는 도너는 생체 내로 전달된다. 다른 구체예에서, 뉴클레아제 및/또는 도너는 환자에 대한 생체 외 전달에 유용한 변형된 세포를 제공하기 위해 분리된 세포 (예컨대 자가조직 또는 이종성 줄기세포)에 전달된다.

- [0139] 본원에 기술된 뉴클레아제를 전달하는 방법은, 예를 들면 미국 특허 번호 6,453,242; 6,503,717; 6,534,261; 6,599,692; 6,607,882; 6,689,558; 6,824,978; 6,933,113; 6,979,539; 7,013,219; 및 7,163,824에 기술되어 있고, 이것들은 전체적으로 본원에 참조로 포함된다.
- [0140] 네이키드 DNA 및/또는 RNA (예컨대 mRNA) 및 하나 이상의 성분을 코드화하는 서열을 함유하는 벡터를 포함하여, 본원에 기술된 뉴클레아제 및/또는 도너 구성물은 임의의 핵산 전달 메커니즘을 사용하여 전달될 수 있다. 제한하는 것은 아니지만, 플라스미드 벡터, DNA 미니씨클, 레트로바이러스 벡터, 렌티바이러스 벡터, 아데노바이러스 벡터, 폭스바이러스 벡터; 단순포진 벡터 및 아데노-관련 바이러스 벡터 등, 및 그것들의 조합을 포함하여 임의의 벡터 시스템이 사용될 수 있다. 또한 미국 특허 번호 6,534,261; 6,607,882; 6,824,978; 6,933,113; 6,979,539; 7,013,219; 및 7,163,824, 및 미국 특허 출원 번호 14/271,008 (이것들은 전체 내용이 본원에 참조로 포함됨)을 참조한다. 나아가, 이를 시스템 중 임의의 것이 치료에 필요한 하나 이상의 서열을 포함할 수 있다는 것이 드러날 것이다. 그러므로, 하나 이상의 뉴클레아제 및 도너 구성물이 세포 안에 도입될 때, 뉴클레아제 및/또는 도너 폴리뉴클레오티드는 동일한 전달 시스템 상에서 또는 상이한 전달 메커니즘 상에서 수행될 수 있다. 다중 시스템이 사용될 때, 각각의 전달 메커니즘은 하나 또는 다중 뉴클레아제 및/또는 도너 구성물을 코드화하는 서열 (하나 이상의 뉴클레아제를 코드화하는 mRNA 및/또는 하나 이상의 도너 구성물을 운반하는 mRNA 또는 AAV)을 포함할 수 있다.
- [0141] 종래의 바이러스 및 비-바이러스 기반 유전자 전달 방법은 뉴클레아제를 코드화하는 핵산 및 도너 구성물을 세포 (예컨대 포유류 세포) 및 표적 조직에 도입하기 위해 사용될 수 있다. 비-바이러스 벡터 전달 시스템은 DNA 플라스미드, DNA 미니씨클, 네이키드 핵산 및 리포좀, 지질 나노입자, 폴리-락테이트-글리콜산 나노입자, 폴리-아민 복합체화제 또는 폴록사미와 같은 전달 비히클과 복합체를 형성한 핵산을 포함한다. 바이러스 벡터 전달 시스템은 세포에 전달된 후 에피솜성 또는 통합된 게놈 중 어느 하나를 가지는 DNA 및 RNA 바이러스를 포함한다. 유전자 요법 과정을 개관하기 위해서는, Anderson, *Science* 256:808-813 (1992); Nabel & Felgner, *TIBTECH* 11:211-217 (1993); Mitani & Caskey, *TIBTECH* 11:162-166 (1993); Dillon, *TIBTECH* 11:167-175 (1993); Miller, *Nature* 357:455-460 (1992); Van Brunt, *Biotechnology* 6(10):1149-1154 (1988); Vigne, *Restorative Neurology 및 Neuroscience* 8:35-36 (1995); Kremer & Perricaudet, *British Medical Bulletin* 51(1):31-44 (1995); Haddada et al., in *Current Topics in Microbiology and Immunology* Doerfler 및 Bohm (eds.) (1995); 및 Yu et al., *Gene Therapy* 1:13-26 (1994)를 참조한다.
- [0142] 핵산의 비-바이러스 전달 방법은 전기천공, 리포펙션, 마이크로주입, 유전자총 (biolistics), 비로좀, 리포좀, 면역리포좀, 다가양이온 또는 지질:핵산 컨쥬게이트, 네이키드 DNA, 네이키드 RNA, 캡핑된 RNA, 인공 비리온 및 DNA의 제제-증강된 흡수를 포함한다. 예컨대 Sonitron 2000 시스템 (Rich-Mar)을 사용하는 초음파 천공법 또한 핵산의 전달을 위해 사용될 수 있다.
- [0143] 추가의 예시적인 핵산 전달 시스템은 Amaxa Biosystems (Cologne, Germany), Maxcyte, Inc. (Rockville, Maryland 및), BTX Molecular Delivery systems (Holliston, MA) 및 Copernicus Therapeutics Inc.에 의해 제공된 것들을 포함한다 (예를 들면 미국 특허 번호 6,008,336). 리포펙션은 예컨대 미국 특허 번호 5,049,386; 4,946,787; 및 4,897,355에 기술되고, 리포펙션 시약은 상업적으로 시판된다 (예컨대 TransfectamTM 및 LipofectinTM). 폴리뉴클레오티드의 효과적인 수용체-인식 리포펙션에 적합한 양이온성 및 중성 지질은 Felgner, WO 91/17424, WO 91/16024의 것들을 포함한다.
- [0144] 면역지질 복합체와 같은 표적화된 리포좀을 포함하여 지질:핵산 복합체의 제조는 기술분야의 숙련자에게 잘 알려져 있다 (예컨대 Crystal, *Science* 270:404-410 (1995); Blaese et al., *Cancer Gene Ther.* 2:291-297 (1995); Behr et al., *Bioconjugate Chem.* 5:382-389 (1994); Remy et al., *Bioconjugate Chem.* 5:647-654 (1994); Gao et al., *Gene Therapy* 2:710-722 (1995); Ahmad et al., *Cancer Res.* 52:4817-4820 (1992); 미국 특허 번호 4,186,183, 4,217,344, 4,235,871, 4,261,975, 4,485,054, 4,501,728, 4,774,085, 4,837,028, 및 4,946,787 참조).
- [0145] 추가의 전달 방법은 EnGeneIC 전달 비히클 (EDV) 안으로 전달된 핵산의 패키징의 사용을 포함한다. 이를 EDV는 항체의 한 아암이 표적 조직에 대한 특이성을 가지고 다른 아암은 EDV에 대한 특이성을 가지고 있는 이중특이적 항체를 사용하여 표적 조직에 특이적으로 전달된다. 항체는 EDV를 표적 세포 표면으로 가져간 후 EDV가 세포 안으로 세포내 이물흡수에 의해 들어간다. 일단 세포 안에서는 내용물이 방출된다 (MacDiarmid et al (2009) *Nature Biotechnology* 27(7):643 참조).

- [0146] 엔지니어링된 CRISPR/Cas 시스템을 코드화하는 핵산의 전달을 위한 RNA 또는 DNA 바이러스 기반 시스템의 사용은 신체에서 특수 세포에 바이러스를 표적화하고 바이러스 페이로드를 핵으로 건네주기 위한 고도로 진화된 과정들의 장점을 가진다. 바이러스 벡터는 대상에게 직접 투여될 수 있고 (생체 내) 또는 시험관 내에서 세포를 치료하기 위해 사용될 수 있고 변형된 세포는 대상에게 투여된다 (생체 외). CRISPR/Cas 시스템의 전달을 위한 종래의 바이러스 기반 시스템은, 제한하는 것은 아니지만, 유전자 전달을 위한 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노-관련, 백시니아 및 단순 포진 바이러스 벡터를 포함한다. 숙주 게놈에서의 통합은 레트로바이러스, 렌티바이러스 및 아데노-관련 바이러스 유전자 전달 방법으로 가능하며, 자주 삽입된 전이유전자의 장기간 발현을 초래한다. 추가로, 높은 형질도입 효율이 많은 상이한 세포 유형 및 표적 조직에서 관찰되었다.
- [0147] 레트로바이러스의 향성을 외래 외피 단백질의 통합, 표적 세포의 잠재적인 표적 집단의 팽창에 의해 변경될 수 있다. 렌티바이러스 벡터는 비-분할 세포를 형질도입시키거나 감염시키고 전형적으로 고 바이러스 역할을 유발할 수 있는 레트로바이러스 벡터이다. 레트로바이러스 유전자 전달 시스템의 선택은 표적 조직에 좌우된다. 레트로바이러스 벡터는 6 내지 10 kb 까지의 외래 서열에 대한 패키징 용량을 가지는 시스-작용 긴 말단 반복부로 구성된다. 최소 시스-작용 LTR은 벡터의 복제 및 패키징에 충분하며, 그것은 다음에 영구적인 전이유전자 발현을 제공하기 위해 표적 세포 안으로 치료제를 통합시키기 위해 사용된다. 광범위하게 사용된 레트로바이러스 벡터는 쥐과 백혈병 바이러스 (MuLV), 긴팔 원숭이 백혈병 바이러스 (GaLV), 유인원 면역결핍 바이러스 (SIV), 인간 면역결핍 바이러스 (HIV) 및 그것들의 조합을 기반으로 한 것들을 포함한다 (예컨대 Buchscher *et al.*, *J. Virol.* 66:2731-2739 (1992); Johann *et al.*, *J. Virol.* 66:1635-1640 (1992); Sommerfelt *et al.*, *Virol.* 176:58-59 (1990); Wilson *et al.*, *J. Virol.* 63:2374-2378 (1989); Miller *et al.*, *J. Virol.* 65:2220-2224 (1991); PCT/US94/05700 참조).
- [0148] 일시적인 발현이 바람직한 용도에서는, 아데노바이러스 기반 시스템이 사용될 수 있다. 아데노바이러스 기반 벡터는 많은 세포 유형에서 매우 높은 형질도입 효율을 가질 수 있고 세포 분할을 필요로 하지 않는다. 그런 벡터를 사용하면, 고역가 및 고수준의 발현이 얻어졌다. 이 벡터는 상대적으로 단순한 시스템으로 다양으로 제조될 수 있다. 아데노-관련 바이러스 ("AAV") 벡터는 또한, 예컨대 핵산 및 펩티드의 시험관 내 제조에, 및 생체 내 및 생체 외 유전자 치료 과정을 위해, 표적 핵산으로 세포를 형질도입하기 위해 사용된다 (예컨대 West *et al.*, *Virology* 160:38-47 (1987); 미국 특허 번호 4,797,368; WO 93/24641; Kotin, *Human Gene Therapy* 5:793-801 (1994); Muzyczka, *J. Clin. Invest.* 94:1351 (1994) 참조). 재조합 AAV 벡터의 구성은 많은 간행물, 이를테면 미국 특허 번호 5,173,414; Tratschin *et al.*, *Mol. Cell. Biol.* 5:3251-3260 (1985); Tratschin, *et al.*, *Mol. Cell. Biol.* 4:2072-2081 (1984); Hermonat & Muzyczka, *PNAS* 81:6466-6470 (1984); 및 Samulski *et al.*, *J. Virol.* 63:3822-3828 (1989)에 기술된다. AAV1, AAV3, AAV4, AAV5, AAV6 및 AAV8, AAV 8.2, AAV9, 및 AAV rh10 및 위타입의 AAV, 예컨대 AV2/8, AAV2/5 및 AAV2/6을 포함하여, 임의의 AAV 혈청형이 사용될 수 있다.
- [0149] 임상 실험에서 유전자 전달을 위해 적어도 6가지의 바이러스 벡터 접근법이 현재 활용되는데, 그것은 형질도입 제제를 생성하기 위하여 헬퍼 세포주 안으로 삽입된 유전자에 의해 결합 벡터를 보충하는 것을 포함하는 접근법을 활용한다.
- [0150] pLASN 및 MFG-S는 임상 실험에 사용된 레트로바이러스 벡터이다 (Dunbar *et al.*, *Blood* 85:3048-305 (1995); Kohn *et al.*, *Nat. Med.* 1:1017-102 (1995); Malech *et al.*, *PNAS* 94:22 12133-12138 (1997)). PA317/pLASN은 유전자 요법 실험에 사용된 첫 번째 치료 벡터였다 (Blaese *et al.*, *Science* 270:475-480 (1995)). 50% 이상의 형질도입 효율은 MFG-S 패키징된 벡터에 대해 관찰되었다 (Ellem *et al.*, *Immunol Immunother.* 44(1):10-20 (1997); Dranoff *et al.*, *Hum. Gene Ther.* 1:111-2 (1997)).
- [0151] 재조합 아데노-관련 바이러스 벡터 (rAAV)는 결핍 및 비병원성 파르보바이러스 아데노-관련 유형 2 바이러스를 기반으로 한 유망한 대체 유전자 전달 시스템이다. 모든 벡터는 전이유전자 발현 카세트 양 옆의 AAV 145 염기쌍 (bp) 반전 말단 반복부만을 보유한 플라스미드로부터 유도된다. 형질도입된 세포의 게놈 안으로의 통합으로 인한 효율적인 유전자 전달 및 안정한 전이유전자 전달은 이 벡터 시스템에 대한 핵심 특징이다 (Wagner *et al.*, *Lancet* 351:9117 1702-3 (1998), Kearns *et al.*, *Gene Ther.* 9:748-55 (1996)). AAV1, AAV3, AAV4, AAV5, AAV6, AAV8, AAV9 및 AAVrh10, 및 그것들의 모든 변이체를 포함하여, 다른 AAV 혈청형 또한 본 발명에 따라 사용될 수 있다.
- [0152] 복제-결핍 재조합 아데노바이러스 벡터 (Ad)는 고역가로 제조될 수 있고 많은 상이한 세포 유형을 쉽게 감염시킨다. 대부분의 아데노바이러스 벡터는 엔지니어링되어서 전이유전자가 Ad E1a, E1b, 및/또는 E3 유전자를 대체

한다; 계속해서 복제 결핍 벡터는 트란스로 결실된 유전자 기능을 공급하는 인간 293 세포에서 증식된다. Ad 벡터는 분할하지 않는, 분화된 세포, 예컨대 간, 신장 및 근육에서 발견된 것들을 포함하여, 생체 내에서 다중 유형의 조직을 형질도입할 수 있다. 종래의 Ad 벡터는 큰 운반 용량을 가진다. 임상 실험에서 Ad 벡터를 사용하는 실례는 근육 내 주사를 포함한 항-종양 면역화를 위한 폴리뉴클레오티드 치료법을 포함하였다 (Sterman et al., *Hum. Gene Ther.* 7:1083-9 (1998)). 임상 실험에서 유전자 전달을 위해 아데노바이러스 벡터를 사용하는 추가의 실례는 다음을 포함한다: Rosenecker et al., *Infection* 24:1 5-10 (1996); Sterman et al., *Hum. Gene Ther.* 9:7 1083-1089 (1998); Welsh et al., *Hum. Gene Ther.* 2:205-18 (1995); Alvarez et al., *Hum. Gene Ther.* 5:597-613 (1997); Topf et al., *Gene Ther.* 5:507-513 (1998); Sterman et al., *Hum. Gene Ther.* 7:1083-1089 (1998).

[0153]

세포 패키징은 숙주 세포를 감염시킬 수 있는 바이러스 입자를 형성하기 위해 사용된다. 그런 세포는 아데노바이러스를 패키징하는 293 세포, 및 레트로바이러스를 패키징하는 ψ 세포 또는 PA317 세포를 포함한다. 유전자 요법에 사용된 바이러스 벡터는 통상적으로 핵산 벡터를 바이러스 입자 안으로 패키징하는 생산자 세포주에 의해 생성된다. 벡터는 전형적으로 패키징 및 숙주 안으로의 후속 통합 (적용 가능하다면)에 필요한 최소 바이러스 서열을 함유하고, 다른 바이러스 서열은 발현될 단백질을 코드화하는 발현 카세트에 의해 대체된다. 바이러스 기능을 잃는 것은 미싱 바이러스 기능 (missing viral function)은 패키징 세포주에 의해 트란스로 공급된다. 예를 들어, 유전자 요법에 사용된 AAV 벡터는 전형적으로 패키징 및 숙주 계놈 안으로의 통합에 필요한 AAV 계놈으로부터의 반전된 말단 반복부 (ITR) 서열만을 가진다. 바이러스 DNA는 다른 AAV 유전자, 즉 *rep* 및 *cap*을 코드화하는 헬퍼 플라스미드를 함유하지만, ITR 서열은 없는 세포주에 패키징된다. 세포주는 또한 헬퍼로서 아데노바이러스로 감염된다. 헬퍼 바이러스는 AAV 벡터의 복제 및 헬퍼 플라스미드로부터의 AAV 유전자의 발현을 촉진한다. 헬퍼 플라스미드는 ITR 서열의 결핍으로 인해 유의미한 양으로 패키징되지 않는다. 아데노바이러스로의 오염은, 예컨대 아데노바이러스가 AAV보다 더 민감한 열 처리에 의해 감소될 수 있다.

[0154]

많은 유전자 요법 용도에서, 유전자 요법 벡터는 특정 조직 유형으로 고도의 특이성으로 전달되는 것이 바람직하다. 따라서, 바이러스 벡터는 바이러스의 외면에 바이러스 코트 단백질을 가지는 융합 단백질로서 리간드를 발현함으로써 주어진 세포 유형에 대한 특이성을 가지고 변형될 수 있다. 리간드는 관심의 세포 유형에 존재하는 것으로 알려진 수용체에 대한 친화성을 가지고 변형될 수 있다. 예를 들어, Han et al.은 *Proc. Natl. Acad. Sci. USA* 92:9747-9751 (1995)에서 Moloney 쥐과 백혈병 바이러스가 gp70에 융합된 인간 혜레굴린 (heregulin)을 발현하도록 변형될 수 있고, 재조합 바이러스는 인간 상피 성장 인자 수용체를 발현하는 특정 인간 유방암 세포를 감염시킨다고 보고하였다. 이 원리는 다른 바이러스-표적 세포상에 대해 연장될 수 있고, 여기서 표적 세포는 수용체를 발현하고 바이러스는 세포-표면 수용체에 대한 리간드를 포함하는 융합 단백질을 발현한다. 예를 들어, 필라멘트상 파지는 실제로 임의의 선택된 세포 수용체에 대한 리간드를 포함하는 융합 단백질을 발현한다. 비록 상기 설명이 주로 바이러스 벡터에 적용되긴 하지만, 동일한 원리가 비바이러스 벡터에도 적용될 수 있다. 그런 벡터는 특이적인 표적 세포에 의한 흡수를 선호하는 특이적인 흡수 서열을 함유하도록 엔지니어링될 수 있다.

[0155]

유전자 요법 벡터는 개별적인 대상에 투여함으로써, 전형적으로 하기 기술되는 것과 같이 전신적 투여 (예컨대 정맥 내, 복강 내, 근육 내, 피하 또는 두개 내 주입)에 의해 또는 국소 적용에 의해 생체 내에서 전달될 수 있다. 다르게는, 벡터는 세포에, 예컨대 개별적인 환자로부터 외식된 세포 (예컨대 럼프구, 골수 흡입, 조직 생검) 또는 보편적 도너 조혈 줄기세포에 전달되고, 이어서, 통상적으로 벡터를 통합한 세포에 대한 선택 후에 환자 안으로 세포가 재이식된다.

[0156]

뉴클레아제 및/또는 도너 구성물을 함유하는 벡터 (예컨대 레트로바이러스, 아데노바이러스, 리포좀 등)는 또한 생체 내에서 세포의 형질도입을 위한 유기체에 직접 투여될 수 있다. 다르게는, 네이키드 DNA가 투여될 수 있다. 투여는 궁극적으로 혈액 또는 조직 세포와 접촉하는 분자를 도입하기 위해 정상적으로 사용된 임의의 경로, 예를 들면 제한하는 것은 아니지만, 주사, 주입, 국소 적용 및 전기천공에 의해 이루어진다. 그런 핵산을 투여하는 적합한 방법은 활용 가능하고 기술분야의 숙련자들에게 잘 알려져 있지만, 특정 조성물을 투여하기 위해 하나 이상의 경로가 사용될 수 있고, 특정 경로는 자주 다른 경로보다 더 즉각적이고 더 효과적인 반응을 제공할 수 있다.

[0157]

본원에 기술된 폴리뉴클레오티드의 도입에 적합한 벡터는 비-통합 렌티바이러스 벡터 (IDLV)를 포함한다. 예를 들면 Ory et al. (1996) *Proc. Natl. Acad. Sci. USA* 93:11382-11388; Dull et al. (1998) *J. Virol.* 72:8463-8471; Zuffery et al. (1998) *J. Virol.* 72:9873-9880; Follenzi et al. (2000) *Nature Genetics*

25:217-222; 미국 특허 공보 번호 2009/054985를 참조한다.

[0158] 약학적으로 허용되는 담체는 부분적으로 투여되는 특정 조성물에 의해, 뿐만 아니라 조성물을 투여하기 위해 사용된 특정 방법에 의해 결정된다. 따라서, 하기에서와 같이, 활용 가능한 약학 조성물의 적합한 제형이 광범위하게 있다 (예컨대 *Remington's Pharmaceutical Sciences*, 17th ed., 1989 참조).

[0159] 뉴클레아제-코드화 서열 및 도너 구성물이 동일한 또는 상이한 시스템을 사용하여 전달될 수 있다는 것이 드러날 것이다. 예를 들어, 도너 폴리뉴클레오티드는 또는 AAV에 의해 운반될 수 있는 한편, 하나 이상의 뉴클레아제가 mRNA에 의해 운반될 수 있다. 나아가, 상이한 시스템은 동일한 또는 상이한 경로 (근육 내 주사, 꼬리 정맥 주사, 다른 정맥 내 주사, 복강 내 투여 및/또는 근육 내 주사)에 의해 투여될 수 있다. 백터는 동시에 또는 임의의 순차적인 순서로 전달될 수 있다.

[0160] 생체 외 및 생체 내 투여 둘 다에 대한 제형은 액체 또는 유화된 액체 상태의 혼탁액을 포함한다. 활성 성분은 자주 약학적으로 허용되고 활성 성분과 부합하는 부형제들과 혼합된다. 적합한 부형제로는, 예를 들면 물, 식염수, 텍스트로오스, 글리세롤, 에탄올 등, 및 그것들의 조합을 포함한다. 또한, 조성물은 미량의 보조제 물질, 예컨대 습윤제 또는 유화제, pH 완충제, 안정화제 또는 약학 조성물의 유효성을 증강시키는 다른 시약들을 함유할 수 있다.

키트

[0161] 또한 상기 방법들 중 임의의 방법을 수행하기 위한 키트가 제공된다. 키트는 전형적으로 본원에 기술된 것과 같은 하나 이상의 뉴클레아제를 코드화하는 폴리뉴클레오티드, 줄기세포 팽창에 영향을 미치는 하나 이상의 인자 및/또는 도너 폴리뉴클레오티드뿐만 아니라 뉴클레아제 및/또는 도너 폴리뉴클레오티드가 도입되는 세포 (또는 주변 배지) 안으로의 줄기세포에 영향을 미치는 인자들을 투여하기 위한 지침을 함유한다. 키트는 또한 세포, 세포의 형질전환을 위한 완충제, 세포를 위한 배양 배지, 및/또는 분석을 수행하기 위한 완충제를 함유할 수 있다. 전형적으로, 키트는 또한 지침, 키트의 다른 성분에 부착되거나 그렇지 않으면 수반하는 패키징 또는 광고 리플릿과 같은 임의의 물질을 포함하는 라벨을 함유한다.

[0163] 다음의 실시예는 뉴클레아제가 하나 이상의 ZFN 또는 하나 이상의 TALEN을 포함하는 본 발명의 예시적인 구체예들에 관련된다. 이것은 예시만의 목적에 대한 것이고 다른 뉴클레아제, 예를 들면 엔지니어링된 DNA-결합 도메인 및/또는 엔지니어링된 호밍 엔도뉴클레아제 (메가뉴클레아제) DNA-결합 도메인 및 이종성 절단 도메인의 자연적으로 발생하는 것의 융합을 가지는 호밍 엔도뉴클레아제 (메가뉴클레아제), 메가 TAL, 암축 TALEN 및 뉴클레아제 시스템, 예컨대 엔지니어링된 단일 가이드 RNA를 사용하는 TtAgo 및 CRISPR/Cas가 사용될 수 있는 것으로 드러날 것이다.

실시예

실시예 1: 징크 평거 단백질 뉴클레아제s (ZFN)의 디자인, 구성 및 일반적 특성화

[0166] 징크 평거 단백질을 디자인하고 본질적으로 Urnov *et al.* (2005) *Nature* 435(7042):646-651, Perez *et al* (2008) *Nature Biotechnology* 26(7):808-816에 기술된 것과 같이, 및 미국 특허 번호 6,534,261에 기술된 것과 같이 결합에 대해 시험한 플라스미드, AAV 또는 아데노바이러스 백터에 통합하였다. 인간 베타 글로빈 유전자좌에 특이적인 ZFN 및 TALEN에 대해, 공동-소유의 미국 특허 번호 7,888,121 및 미국 특허 공개공보 번호 20130137104 및 20130122591을 참조한다.

실시예 2: 글로빈-특이적 ZFN의 활성

[0168] 인간 글로빈 유전자좌를 표적화하는 ZFN 쌍을 사용하여 이를 ZFN이 특이적 표적 부위에서 DSB를 유도하는 능력을 시험하였다. 표시된 ZFN의 각 평거의 인식 나선 영역의 아미노산 서열을 아래의 표 1에 나타낸다. 표적 부위 (대문자로 표시된 DNA 표적 부위; 소문자로 표시된 비-접촉 뉴클레오티드)를 표 2에 나타낸다.

표 1

인간 베타 글로빈-특이적 징크 평거 단백질 인식 나선 디자인

| SBS # | 디자인 | | | | | |
|-------|-----|----|----|----|----|----|
| | F1 | F2 | F3 | F4 | F5 | F6 |
| | | | | | | |

| | | | | | | |
|--------------|---------------------------|----------------------------|------------------------|-----------------------|-----------------------|---------------------------|
| 33488 | LRHHHLTR (SEQ ID NO:12) | LRHNLRA(SEQ ID NO:13) | DQSNLRA(SEQ ID NO:14) | RNASRTR(SEQ ID NO:15) | RSDNLSE(SEQ ID NO:16) | RSQHRKT(SE Q ID NO:17) |
| 33501 | TSGSLSR(SEQ ID NO:18) | DRSDLSSR(SEQ ID NO:19) | DRSALAR(SEQ ID NO:20) | QSSNLAR(SEQ ID NO:21) | QSGHLSR(SEQ ID NO:22) | NA |
| 47773 | LRHHHLTR (SEQ ID NO:12) | LKQNLDAA (SEQ ID NO:55) | DQSNLRA(SEQ ID NO:14) | RNCARLR(SEQ ID NO:56) | RSDNLSE(SEQ ID NO:16) | RNQTRLN (SEQ ID NO:57) |
| 47817 | HHHSLKR (SEQ ID NO:58) | DRSDLSSR(SEQ ID NO:19) | TNSELDR (SEQ ID NO:59) | QSSNLAR(SEQ ID NO:21) | QSGHLSR(SEQ ID NO:22) | NA |

표 2**인간 베타 글로빈-특이적 징크 평거의 표적 부위**

| SBS # | 표적 부위 |
|-------------------------|---|
| 33488, 47773 | ggAGTCAGGTGCACCATGGTgtctgttt (SEQ ID NO:23) |
| 33501, 47817 | gtGGAGAAGTCtGCCGTTactgccctgt (SEQ ID NO:24) |

[0170] 인간 CD34+ 세포를 다음과 같이 얻었고 처리하였다. 모든 제대혈 (CD) 시편을 캘리포니아 주립대학교에 의해 승인된 지침에 따라 얻었고, IRB 리뷰로부터 미지의 의료 폐기물로 간주하였다. 세포를 수집 후 48시간 이내에 처리하였다. SCD를 가진 자원 도너로부터의 골수 (BM) 흡입물을 UCLA IRB 프로토콜 #10-001399 하에 고지에 의한 동의로 얻었다. 단핵 세포 (MNC)를 Ficoll Hypaque (Stem Cell Technologies) 밀도 원심분리를 사용하여 BM 및 CB로부터 분리하였다. 그런 다음 면역자성 칼럼 분리를 사용하여 MNC를 항-CD34 마이크로비드 (Miltenyi Biotech Inc.)와 함께 4°C에서 30분 동안 인큐베이션함으로써 CD34+ 세포에 대해 풍부하게 하였다. 그런 다음 세포를 자성 칼럼을 통해 보냈고 CD34+ 세포를 수집하여 냉동 배지 (10% 다이메틸 셀룰라이드 (Sigma Aldrich), 90% FBS)가 들어있는 저온 바이알에 넣고 액체 질소 중에 저온보존하였다. 이동된 CD34+ 세포 (mPB)를 Allcells로부터 구매하였다.

[0171] 전기천공을 위해 mRNA를 제조하기 위하여, 표 1에 나타낸 ZFN을 코드화하는 플라스미드를 SpeI (New England Biolabs)로 선형화하고 시험관 내 전사를 위한 주형으로서 사용 전에 폐놀:클로로포름에 의해 정제하였다. mMessage mMachine® T7 ULTRA 전사 키트 (Life Technologies)를 제조자의 프로토콜을 따라 사용하여 시험관 내 전사된 mRNA를 제조하였고 RNeasy® MinElute® Cleanup 키트 (Qiagen)로 세정하였다. 전기천공을 위하여, 다음의 프로토콜을 따랐다: CB CD34+ 세포를 37°C에서 해동하고, 20% 우태아 혈청 (Gemini Bio-products) 및 (1x 글루타민, 폐니실린 및 스트렙토마이신)이 첨가된 이스코브 변형 둘베코 배지 (IMDM; Life Technologies)로 세척하고 48시간 동안 글루타민, 폐니실린, 스트렙토마이신, 50 ng/ml SCF, 50 ng/ml Flt-3 및 50 ng/ml TPO (Peprotech)를 함유한 X-VIVO15 배지 (Lonza)에서 예비-자극하였다. 전기천공을 위해, 반응당 200,000 세포를 90g에서 15분 동안 회전시키고, 100 μl의 BTXpress 완충액 (Harvard Apparatus)에 재현탁한 후, 표시된 양의 ZFN mRNA 및/또는 해당되는 경우 올리고뉴클레오티드와 혼합하고, 250V에서 5 밀리초 동안 BTX ECM 830 Square Wave 전기천공기 (Harvard Apparatus)에서 펄스하였다.

[0172] 전기천공에 이어서, 세포를 10분 동안 실온에서 쉬게 한 후 배양 배지를 첨가하고 총 500 ul로 플레이트에 옮겼다. 도너 IDLV가 적절한 샘플에 대해 기술된 농도에서 최종 배양 배지에 존재하였다.

[0173] 유전자 교정 및 유전자 파괴를 다음과 같이 분석하였다: Surveyor® 뉴클레아제 분석 (Cel-1)을 사용하여 ZFN 유도된 부위-특이적 대립유전자 파괴를 측정하였다. ZFN 결합 부위를 둘러싸고 있는 410bp 영역을 200 ng의 개놈 DNA로부터 Cel1Fwd (5'-gacaggtacggctgtcatca-3' SEQ ID NO:25) 및 Cel1Rev (5'-cagcctaagggtggaaat-3' SEQ ID NO:26)를 사용하여 Accuprime Taq Hi-Fi (Life Technologies)를 사용하여 PCR 증폭시켰다. 변성, 재아닐링, 소화 및 전기영동적 및 밀도측정 분석을 기술된 대로 완료하였다 (예컨대 Joglekat *et al* (2013) Mol Ther: J of the Amer Soc Gene Ther 21:1705-17). 부위-특이적 유전자 변형을 제한 단편 길이 다형태 (RFLP)에

의해 검출하였다.

[0175] ZFN 결합 부위를 둘러싸고 있는 1.1 kb 영역을 프라이머 BgloOuterFwd (5'-atgcttagaaccgaggtagagtt-3', SEQ ID NO:27) 및 BgloOuterRev (5'-cctgagacttccacactgatg-3', SEQ ID NO:28) 및 Accuprime Taq Hi-Fi (Life Technologies)를 사용하여 PCR 증폭시켰다. PCR 생성물을 PCR 세정 키트 (Life Technologies)를 사용하여 정제하였고 10 유닛의 HhaI (New England Biolabs)를 사용하여 3.5시간 동안 37°C에서 소화시켰다. 소화 생성물을 GelGreen (Biotium)으로 사전-염색한 1.0% TBE-아가로오스 젤 상에서 분리하고 Typhoon FLA 9000 Biomolecular Imager (GE Healthcare) 상에서 영상화하였다.

[0176] 유전자 변형을 정량하기 위하여, 정량 PCR 기반 분석을 사용하였다. 두 개의 PCR 반응의 세트를 상기에서 기술한 1.1 kb PCR 생성물을 주형으로서 사용하여 수행하였다. 정제하지 않은 PCR 주형을 1:5,000으로 희석하고 그 중 1 μ l를 다음의 25 μ l 반응 각각에 사용하였다. 제1 PCR을 수행하여 변형된 게놈을 증폭시켰는데, 프라이머 HhaIFwd (5'-gaagtctgccgtactgcg-3', SEQ ID NO:29) 및 HhaIRev (5'-cccagttctattggctcc-3', SEQ ID NO:30)를 사용하였다. 제2 PCR을 프라이머 ExonIIFwd (5'-ctcggtgccttagtgatgg-3', SEQ ID NO:31) 및 ExonIIRev (5'-gactcaccctaagttctc-3', SEQ ID NO:32)를 사용하여 수행하여 인풋 주형을 표준화하였다. 이들 PCR은 둘 다 Power SYBR Green PCR Master Mix (Life Technologies)를 사용하여 정량적으로 제조하였고 ViiA7 (Life Technologies) 상에 획득하였다. 유전자 변형의 빈도를 2개의 반응과 플라스미드 표준 곡선 사이의 Ct (임계값까지의 주기) 차이를 사용하여 측정하였다. 모든 ZFN은 활성인 것으로 나타났다.

[0177] 글로빈 파라로그를 PCR-증폭하였고 Illumina MiSeq 기계 상에서 심출-서열화하여 표적 변형을 분석해내었다. PCR에 사용한 프라이머는 다음과 같다: HBB: 5'-acacgacgcttccgatctnnnngggctggcataaaagtca-3' (SEQ ID NO:33) 및 5'-gacgtgtcgttccgatcttccacatgcccagttctatt-3' (SEQ ID NO:34); HBD: 5'-acacgacgcttccgatctnnnntaaaaggcaggcagactcga-3' (SEQ ID NO:35) 및 5'-gacgtgtcgttccgatctacatgcccagttccatgc-3' (SEQ ID NO:36); HBE1: 5'-acacgacgcttcggatctnnnntgcttccgacacagctgcaa-3' (SEQ ID NO:37) 및 5'-gacgtgtcgttccgatctnnnnggaacgtctgaggttatcaat-3' (SEQ ID NO:38); HBG1 및 HBG2: 5'-acacgacgcttccgatctnnnnggaacgtctgaggttatcaat-3' (SEQ ID NO:39) 및 5'-gacgtgtcgttccgatcttccctccctgtcc-3' (SEQ ID NO:40). HBG1 및 HBG2를 공동-증폭시켰고 서열 판독을 암플리콘 내의 유전자좌-특이적 SNP를 사용하여 HBG1 또는 HBG2에 배정하였다. 전방 프라이머 내의 혼합된 염기는 서열화 중에 클러스터 디콘볼루션을 허용한다.

[0178] ZFN은 세포 유형에 따라 베타-글로빈 유전자좌에서 35 내지 65%의 대립유전자 파괴 (indels)를 유도한 한편, 베타 글로빈 파라로그에서 검출된 변형은 아래의 표 3에서 나타난 것과 같이 낮았다.

표 3

K562 또는 CD34+ 세포 (33488/33501)에서 인간 베타 글로빈 및 그것의 파라로그에서의 ZFN-관련 변형

| 유전자 | 일루미나 서열화에 의해 검출된 NHEJ (%) | |
|-------|---------------------------|-------------|
| | K562 세포 | 인간 CD34+ 세포 |
| HBB | 66.600 | 37.500 |
| HBD | 34.000 | 28.600 |
| HBE1 | 0.000 | 0.000 |
| HBG1 | 0.006 | 0.000 |
| HBG2 | 0.013 | 0.005 |
| HBGP1 | 0.024 | 0.005 |

[0180] 이동화된 말초혈 (mPB)뿐만 아니라 인간 제대혈 (CB) 둘 다로부터 분리된 CD34+에 ZFN을 코드화하는 시험관 내 전사된 mRNA를 전기친공하는 것은 Surveyor 뉴클레아제 분석에 의해 측정되는 바, 표적 유전자좌의 효과적인 절단을 초래하였다 (도 1B).

[0181] 그 결과는 ZDN이 베타-글로빈 유전자좌에서 고수준의 유전자 변형을 유도했음을 보여주었다.

실시예 4: 겸상 염기에서 ZFN 유도된 교정

[0183] 표적화된 베타-글로빈 유전자좌에서 성공적인 절단에 이어서, 본 발명자들은 이 부위에서의 겸상 염기의 교정이 상동성 도너 주형을 사용하여 가능한지를 측정하고자 하였다. 이 목적을 위해서, 두 가지 유형의 유전자 교정 주형: 짧은 DNA 올리고 및 IDLV를 디자인하였고 나란히 시험하였다. IDLV에 클론된 1.1 kb 인간 베타-글로빈 유

전자 단편 도너 주형을 겸상 돌연변이에서 교정적인 변화, 뿐만 아니라 사일런트 제한 단편 길이 다형태 (RFLP)를 포함하도록 디자인하여 상동성 재조합의 대리 (surrogate) 분석을 위한 HhaI 제한 부위를 생성하였다 (도 1C). 도너 주형은 IDLV에 의해 전달되었고 (도 2D 참조), 그것으로 최소 세포독성으로 CD34+ HSPC의 효율적인 형질도입이 허용되었고, 최소 계놈 통합으로 높은 주형 복사수가 일시적으로 생성되었다.

[0184] ZFN (상 33488/33501) + IDLV 도너로 처리된 CD34+ 세포에서 유전자 변형 수준을 HhaI RFLP 소화 (도 1D) 및 정량적 PCR-기반 (qPCR) 분석에 의해 평균 $18.0 \pm 2.2\%$ 의 대립유전자에서 (도 1E) 처음으로 측정하였다. ZFN mRNA 및 IDLV 도너 농도의 최적화를 수행하였고 (도 2 및 도 3) 세포 수 및 생존율을 유지하면서 고수준의 변형을 이루기 위하여 관심의 세포 유형에서 ZFN 시약을 적정하는 중요성을 증명하였다.

[0185] 유전자 변형/교정 주형으로서 올리고뉴클레오티드의 사용은 실험의 속도, 비용, 재생성 및 용이성의 장점들을 가질 수 있을 것이다. 베타-글로빈 유전자 (HBB)에 AvrII RFLP를 생성하기 위하여 디자인된 3' 염기쌍 서열이 도입된 mPB로부터 HSPC에서 올리고뉴클레오티드 도너 주형을 사용한 초기 실험은 15%의 유전자 변형을 산출하였다 (도 4 참조).

[0186] 올리고 도너의 사용을 개선하기 위하여, 한 가닥 또는 다른 가닥에 상용하고 ZFN 절단 부위를 중심으로 대칭적으로 증가하는 길이의 올리고뉴클레오티드의 패널을 시험하였고 (표 4 참조), 여기서 진한 글씨는 야생형 서열과 관련하여 돌연변이된 염기를 나타낸다. 's'로 표지된 올리고뉴클레오티드는 야생형 유전자 안으로 겸상 돌연변이를 도입하기 위해 디자인된 서열을 포함한다. 표에서 나타난 것과 같이, 'S'로 표지된 열에서, 야생형 서열은 이 위치에 A를 포함하는 한편 겸상 돌연변이는 T를 포함한다.

표 4

올리고 도너

| SMS 위치 | | S | O | | 1 | | 2 | | 3 | | | 4 | | 5 | SEQ ID NO |
|---------|---|---|----------|---|---|----------|---|----------|---|---|---|---|---|------------|------------|
| WT | G | A | G | G | A | G | A | A | G | T | C | T | G | C | C G T T 60 |
| SMS1 | G | A | G | G | A | A | A | A | G | T | C | T | G | C | C G T T 61 |
| SMS2 | G | A | G | G | A | G | A | A | T | C | T | G | C | C G T T 62 | |
| SMS4 | G | A | G | G | A | G | A | A | G | T | C | T | G | C | T G T T 63 |
| SMS12 | G | A | G | G | A | A | A | A | A | T | C | T | G | C | C G T T 64 |
| SMS14 | G | A | G | G | A | A | A | A | G | T | C | T | G | C | T G T T 65 |
| SMS24 | G | A | G | G | A | G | A | A | A | T | C | T | G | C | T G T T 66 |
| SMS124 | G | A | G | G | A | A | A | A | A | T | C | T | G | C | T G T T 67 |
| SMS012 | G | A | A | G | A | A | A | A | A | T | C | T | G | C | C G T T 68 |
| SMS12s | G | T | G | G | A | A | A | A | A | T | C | T | G | C | C G T T 69 |
| SMS012s | G | T | A | G | A | A | A | A | A | T | C | T | G | C | C G T T 70 |

[0188] 더 긴, 반전-가닥의 올리고뉴클레오티드의 사용은 최고 수준의 유전자 변형을 유발하였다 (도 5 참조). SCD 돌연변이 부위에서의 유전자 변형 또는 교정이 ZFN 결합 부위의 서열 또는 스페이싱을 변화시키지 않을 것이기 때문에, 사일런트 돌연변이를 도너 주형에 디자인하였고 염색체 안으로 공동-도입하여 변형된 대립유전자의 ZFN 결합 및 재-절단을 방지하였다. 반전-가닥 도너의 사용은 3' ZFN 결합 부위에서 사일런트 돌연변이의 도입과 부합한다 (도 4C).

[0189] 사일런트 돌연변이들의 다양한 조합을 도너 올리고에 도입하고, 유전자 변형의 빈도를 분석하였다. 사일런트 돌연변이 부위 (SMS) 하나 및 둘 (SMS12)의 조합은 최고 수준의 HBB 변형을 유발하였고 모든 후속 실험에서 사용하였다 (도 4D). 유전자-변형된 HBB 대립유전자와 변형-후-재절단 대립유전자 사이의 역관계는 SMS 도너가 실제로 ZFN 재절단을 차단하는 역할을 한다는 것을 시사한다 (도 4E). HSPC 푸울 및 과열-형성 유닛에서 HBB 전사물, 적혈구 콜로니 (BFU-E)의 조사는 임의의 SMS 대립유전자로부터의 mRNA의 스플라이싱의 변경에 대한 증가를 발견하지 못하였다.

[0190] 높은 ZFN mRNA (상 47773/47817) 농도에서 SMS012를 함유하는 도너 올리고뉴클레오티드의 사용은 SMS12 도너의 사용보다 많은 HBB-변형 대립유전자를 유발하였다 (도 17B).

[0191] ZFN 및 SMS124 도너로의 처리 후에 HBB에서의 문자적 결과의 포괄적인 분석은 상동성-지시된 및 NHEJ-기반 유전

자 변형의 조합 (1.5%) 또는 올리고뉴클레오티드의 NHEJ-기반 포획 (0.4%)으로부터의 것들과 같이 단지 미량 수준의 예상치 못한 DNA 수복 사건을 나타냈다 (아래의 표 5 참조).

표 5

ZFN 및 올리고뉴클레오티드 도너로의 HSPC의 처리로부터의 분자적 결과

| 결과 | 빈도 (%) |
|----------------------|--------|
| 변화 없음 | 46.1 |
| NHEJ-매개된 결실 | 30.5 |
| 올리고뉴클레오티드-주형의 유전자 변형 | 18.8 |
| NHEJ-매개된 삽입 | 2.8 |
| 하프 유전자 변형, 하프 NHEJ | 1.5 |
| 도너 올리고뉴클레오티드의 포획 | 0.4 |

[0193] 마지막으로, 부분적인 상동성-지시된 유전자 변형 사건 (SMS12 대 SMS24)의 과정은 올리고뉴클레오티드-주형의 유전자 변형 중에 새로운 DNA 합성이 절제된 DAB의 좌측 3' 단일-가닥 단부를 사용하여 일어난다는 개념을 강력하게 지지한다 (도 6 참조). ZFN, 올리고뉴클레오티드 및 HSPC 도너의 최적의 조합은 대립유전자의 30 내지 40%의 유전자 변형을 허용하였다 (도 4F 참조).

[0194] 그러므로, ZFN은 상동성 도너 주형과 나란히 전달될 때 일차 HSPC에서 고수준의 부위-특이적 유전자 변형뿐 아니라 고수준의 표적화된 DSB 둘 다를 유도할 수 있다.

실시예 5: ZFN 특이성

[0196] 베타-글로빈이 다른 글로빈 유전자 (엘타-, 앱실론-, A 감마-, G 감마-, 및 위-베타-글로빈)에 대해 높은 상동성을 가지기 때문에, ZFN을 이들 영역에서의 절단을 피하도록 디자인하였다 (도 6A). 33488/33501 쌍을 사용한 상동성 글로빈 유전자에서의 절단률을 Surveyor 뉴클레아제 분석을 사용하여 및 대용량 DNA 서열화에 의해 분석하였다. 이를 영역의 각각의 분석은 mPB, CB, 및 SCD BM CD34+ 세포에서 단지 고도로 상동성인 엘타-글로빈 유전자에서만 표적-외 변형을 나타냈다 (도 6B, 표 3). 본 발명자들은 표적-외 부위를 확인하기 위하여 보다 포괄적이고, 편향되지 않은 계놈-전체 접근법으로 표적-외 ZFN 절단의 이런 직접 절단을 고려하였다. 이 분석은 비-상동성 IDLV가 이중-가닥 파괴 부위에서 포획될 성향을 기반으로 하고 IDLV 포획에 대한 허용성에 대해 선택된 K562 세포에서 통합의 하류 응집된 통합 부위 분석 (CLIS)을 사용한다. 사용한 과정을 아래에 간략하게 기술한다:

[0197] 베타-글로빈 유전자 클러스터 부위를 다음의 프라이머들: HBD: 5'-ggttcattttattctaca-3' (SEQ ID NO:41) 및 5'- gtaatctgaggtaggaaac-3' (SEQ ID NO:42); HBBP1: 5'-cacccttgaccaatagattc-3' (SEQ ID NO:52) 및 5'- gagactgtggatagtata-3' (SEQ ID NO:43); HBE1: 5'-cattatcacaaactttagtgtcc-3' (SEQ ID NO:44) 및 5'- agtctatgaaatgacaccatat-3' (SEQ ID NO:45); HBG1: 5'- gcaaaggctataaaaaattagc-3' (SEQ ID NO:46) 및 5'- gagatcatccaggtgcattt-3' (SEQ ID NO:47); HBG2: 5'-ggcaaaggctataaaaaattaagca-3' (SEQ ID NO:48) 및 5'- gagatcatccaggtgcattt-3' (SEQ ID NO:49)을 사용하여 PCR-증폭시켰고 상기 기술한 것과 같이 Surveyor® 뉴클레아제에 의해 분석하였다.

[0198] K562 세포를 15 μg/mL의 ZFN mRNA로 전기천공하고 2E+08 T U/mL (MOI=250)의, MND 프로모터 (Challita, P. M. et al. (1995) *J of Viro* 69, 748-55)에 의해 구동되고 베타-글로빈 유전자좌에 대해 임의의 상동성이 없는 GFP 전이유전자를 함유하는 IDLV로 형질도입하여 이중 가닥 파괴에서의 통합이 NHEJ-매개된 포획으로 인한 것임을 확실하게 하였다 (도 7A 참조). ZFN+IDLV의 4개의 개별적인 생물학적 복제물을, 잠재적인 계놈-전체 표적-외 부위에 대한 높은 민감성을 유지하도록 완료하는 한편 두 개의 대조표준은 바탕 통합의 수준을 측정하는 것을 완료하였다. 대조 세포들에는 GFP IDLV를 주었지만, ZFN의 작용 없이 IDLV를 포획하는 자연적으로 발생하는 DSB의 부위를 검출하는 ZFN mRNA는 없었다. K562 세포를 10% 유태아 혈청 (Gemini Bio-products) 및 1% 폐니실린/스트렙토마이신/L-글루타민 (Gemini Bio-products)이 첨가된 RPMI1640 (Cellgro)에서 60일 동안 배양하였다. 이 기간 동안에, GFP 양성에 대한 유동 세포분석계에 의한 분석을 정기적으로 수행하여 임의의 비-통합 IDLV가 샘플에서 회색된 것을 확실하게 하였다. 60일째에, 샘플을 형광-활성화 세포 분류 (FACS)에 의해 분류하여 GFP+ 세포를 분리하였고, 그것을 5일 동안 팽창시키고, 계놈 DNA를 추출한 후, 샘플을 비제한적 선형 증폭 Mediated-PCR35에 의해 벡터 통합 부위의 서열화에 대해 준비하였다. 응집된 통합 부위 (CLIS) 분석을 앞서 개발한 바와

같이 (Gabriel, R., et al(2011) *Nat Biotech.* 29(9):816-23) 500 염기쌍의 CLIS 원도우로 수행하여 ZFN 쌍 서열에 대한 각 부위의 상동성 수준을 측정하였다. 상동성 백분율을 20 뉴클레오티드까지의 스페이서 길이를 가지는 ZFN 결합 (동종이량체 및 이종이량체)의 모든 통합에 대한 부위에서 매칭 염기쌍의 수로서 정의하였다. 각 CLIS에 대한 가장 높은 백분율 상동성을 아래의 표 6에 열거한다. 이 표에 대해, 확인한 모든 CLIS를 그것들의 통합 부위의 200 염기쌍 내에서 ZFN 표적 부위에 대한 최고 백분율의 상동성과 함께 열거한다. ZFN 결합 (동종이량체 및 이종이량체)의 모든 가능한 상호작용을 평가하였고 20 nt까지의 스페이서 길이를 분석하였다. 절단의 두 ZFN 부위를 진하게 강조한다. CLIS에 의해 확인한 다른 모든 파괴되기 쉬운 부위들은 이탤릭체로 표시한다. Chr, 염색체; CLIS, 응집된 통합 부위; L, '좌측' ZFN 단량체 (SBS#33488) ; R, '우측' ZFN 단량체 (SBS#33501); 한편 ZFN 이량체의 수는 ZFN 표적 부위에 대한 상동성의 가장 큰 백분율을 제공하는 스페이서 크기에 상응한다. 20,000개의 무작위 게놈 부위를 이 방법을 사용하여 분석하여 게놈을 가로지르는 ZFN 상동성의 바탕 수준을 측정하였다.

표 6

CLIS 및 ZFN 표적 부위에 대한 상동성 백분율

| Chr | 위치 | 가장 가까운 RefSeq 유전자 (500 kb 이내) | 동일성 (%) | CLIS의 수 | ZFN 이량체 | 샘플 |
|-----|-----------|-------------------------------|---------|---------|---------|----------------|
| 11 | 5248238 | HBB | 100% | 18 | L_5_R | ZFN+GFP |
| 11 | 5255649 | HBD | 91% | 10 | L_5_R | ZFN+GFP |
| 7 | 12238685 | TMEM106b | 52.94% | 3 | L_12_R | ZFN+GFP |
| 19 | 58024409 | ZNF773 | 62.50% | 6 | L_18_L | ZFN+GFP; GFP단독 |
| 1 | 121484829 | EMBP1 | 55.88% | 6 | R_14_L | ZFN+GFP; GFP단독 |
| 10 | 117392565 | ATRNL1 | 55.88% | 3 | L_15_R | ZFN+GFP; GFP단독 |
| 19 | 27733086 | - | 56.25% | 3 | L_17_L | GFP 단독 |
| 12 | 112392669 | TMEM116 | 52.94% | 2 | L_4_R | GFP 단독 |
| 12 | 693139 | NINJ2 | 35.29% | 2 | L_1_R | GFP 단독 |
| 16 | 46428309 | ANKRD26P1 | 53.12% | 2 | L_3_L | GFP 단독 |
| 5 | 144123495 | KCTD16 | 53.12% | 2 | L_4_L | GFP 단독 |
| 3 | 183525263 | YEATS2 | 53.13% | 2 | L_15_L | GFP 단독 |
| 5 | 53018665 | NDUFS4 | 58.82% | 2 | R_9_L | GFP 단독 |
| 3 | 19674417 | EFHB | 56.25% | 2 | L_9_L | GFP 단독 |
| 9 | 127395708 | NR6A1 | 55.88% | 2 | L_6_R | GFP 단독 |
| 7 | 137218714 | DGKI | 52.94% | 2 | L_12_R | GFP 단독 |

[0199] 예측한 바대로, CLIS 분석은 베타-글로빈 및 델타-글로빈에서의 트래핑을 나타냈다 (도 7, 표 6). ZFN 결합 부위에 유의미한 상동성을 가지는 다른 잠재적인 표적-외 부위를 CLIS 분석에서 발견하지 못하였다.

[0200] 이를 결과는 게놈-전체 규모에서 의도된 표적 부위에 대한 ZFN의 이런 쌍의 고수준의 특이성을 증명한다.

0202] 실시예 6: 변형된 HSPC의 시험관 내 분화

[0203] 유전자-변형된 HSPC가 정상적인, 광역 스펙트럼의 적혈구성 및 골수성 분화를 할 수 있는지를 확실히 하기 위해, 처리된 세포들을 단일 세포로서 및 별크 배양에서 모두 분석하였다. 단일-세포를 HSPC 분화를 촉진하는 사이토카인을 함유하는 메틸셀룰로오스 배지에서 콜로니-형성 가능성에 대해 모니터링하였다. 간단하게 설명하면, 전기천공 후 1일 후에, 35 mm 세포 배양 접시당 100 및 300의 생존 세포를 MethoCult™ 최적 메틸셀룰로오스-기반 배지 (Stem Cell Technologies)에 이중으로 플레이팅하였다. 올리고뉴클레오티드 도너를 사용하는 실험을 위해, 세포를 전기천공후 5일 후에 플레이팅하였다. 5% CO₂, 37°C 및 습기 분위기에 2주 동안 배양한 후, 상이한 유형의 콜로니들을 특성화하고 형태를 기반으로 계수하였다. ZFN 단독으로 (쌍 33488/33501) 또는 ZFN과 올리고뉴클레오티드 도너로 처리된 HSPC는 유사한 수 및 패턴의 적혈구계 및 골수성 클론을 생성하였다.

[0204] SCD BM 샘플에 대한 시험관 내 적혈구성 분화 기법을 Giarratana et al. ((2005) *Nat Biotech* 23:69-74)로부터 채택하였고 Romero and Urbinati et al ((2013) *J Clin Invest* 123(8):3317-3330)에 기술되어 있다. 전기천공

후 1일 후에, 세포를 적혈구 배양에 넣었다. 하류 분석을 위해 대다수의 세포를 얻기 위하여, 세포를 8일과는 반대로 프로토콜의 12일에 공동-배양에 넣어 증가된 적혈구 팽창을 허용하였다. 세포를 실험 22일째에 수확하고, 펠릿화한 후, HPLC 분석을 수행하였다 (하기 참조). HBB 스플라이싱의 분석을 또한 RT-PCR에 의해 5'-acatggcttgcacacaac-3' (액손 1, SEQ ID NO:50) 및 5'-gaaattggacagcaagaaagc-3' (액손 3, SEQ ID NO:51)을 사용하여 수행하였고 액손 건너뜀 또는 인트론 보유는 관찰하지 못하였다.

[0205] 개별적인 적혈구계 콜로니들의 유전형 분석은 예상된 빈도로 의도된 편집의 존재를 확인하였다. 또한, HSPC를 벌크 배양에서 적혈구로 분화하도록 유도하였다 (Giarratana, M. et al. (2011) *Blood* 118, 5071-9). 세포를 적혈구 분화 전과정을 통해 샘플링하였고 편집된 세포의 빈도를 대용량 DNA 서열화에 의해 측정하였다. 메틸셀룰로오스에서 성장한 적혈구 콜로니에 대한 본 발명자들의 단일-세포 분석과 일관되게, 변형된 세포의 빈도의 변화는 푸울의 유전자 변형의 출발 수준과 관계없이 벌크 배양에서 적혈구로 분화하는 동안 본질적으로 발견되지 않았다 (도 9A).

[0206] 글로빈 사슬의 분석을 위해, HPLC를 수행하였다 (도 9B). SCD BM 샘플에 대해, 적혈수 세포의 말기 단계에 의해 생성된 혼모글로빈 (Hb) 종의 HPLC를 Wilber, A. et al. ((2011) *Blood* 117:2817-26)에 기술된 것을 약간 변형시켜 수행하였다. 적혈구계 세포 (1-2E+07)를 적혈구 배양이 끝날 때 (22일)에 수확하고, 펠릿화한 후, -80°C에서 용해할 때까지 냉동 보관하였다. 해동시, 세포를 25 uL의 용혈물 시약 (Helena Laboratories)에 재-현탁하고 밤새 냉장보관한 후 20,800 x g에서 10분 동안 4°C에서 원심분리하여 세포 파편을 제거하였다. 깨끗해진 상층액을 양이온-교환 칼럼 (Ultra2 Variant Resolution Analyzer; Primus Diagnostics) 및 인간 혼모글로빈에 대해 보정된 샘플을 사용한 HPLC에 의한 Hb 생성의 특성화에 사용하였다. Hb 종에 상응하는 피크를 HPLC 기기에 수반된 소프트웨어로 보유 시간을 기준으로 확인하였다. 각 샘플에 대해 생성된 HbA의 상대 백분율을 아세틸화된 태아 혼모글로빈을 포함한 일차 혼모글로빈 피크의 각각에 대한 곡선 아래의 면적의 총 합계를 기준으로 계산하였다, HbFAC; 태아 혼모글로빈, HbF; 야생형 혼모글로빈, HbA; 및 겸상 혼모글로빈, HbS.

[0207] mPB 샘플에 대하여, 적혈구계 세포를 적혈구 배양 18일째에 수확하였다. 세포를 펠릿화하고 물에서 10분 동안 RT에서 용해시켰다. 20,000 g에서 5분 동안 원심분리하여 세포 파편을 제거한 후에, 세포 용해물을 -80°C에서 냉동 보관하였다. 해동시, 세포 용해물을 이동상 A에 1:10으로 희석하고 약한 양이온-교환 칼럼 (PolyCAT ATM, PolyLCINC.)을 사용하여 HPLC (Infinity 1260, Agilent)에 의해 특성화하였다. FASC 참조 물질 (Trinity Biotech)을 사용하여 통상적인 혼모글로빈 (HbF, HbA, HbS, HbC)의 용출 시간을 규정하였다. 분석 및 피크 통합을 OpenLAB CDS Chemstation 소프트웨어를 사용하여 수행하였다. 각 샘플에 대해 생성된 HbA의 상대 백분율을 아세틸화된 태아 혼모글로빈을 포함한 혼모글로빈 피크의 각각에 대한 곡선 아래의 면적의 총 합계를 기준으로 계산하였다, HbFAC; 태아 혼모글로빈, HbF; 야생형 혼모글로빈, HbA; 및 겸상 혼모글로빈, HbS.

[0208] 병렬 실험으로, 본 발명자들은 14%의 HBB 대립유전자를 겸상 형태로 전환시키고, HPLC에 의해 분석하여 적혈구 분화 후 18일 후 존재하는 글로빈 사슬을 보였다. 이를 세포에 의해 생성된 혼모글로빈의 대략 18%가 겸상 혼모글로빈이었고, 이것은 단백질 수준에서의 유전자 변형 사건의 영향을 증명한다 (도 9B).

실시예 7: 변형된 세포의 마우스로의 이식

[0210] ZFN-변형된 세포가 그것의 조혈 재증식 능력을 유지하는 지의 여부를 측정하기 위하여, ZFN- 및 도너-처리된 또는 mock-처리된 CB-유도된 HSPC를 면역-결핍 NSG 마우스에 이종이식하였다.

[0211] 간단하게 설명하면, 다수의 건강한 개체로부터의 신선한 CB CD34+ 세포를 2일 동안 예비-자극한 후에 ZFN mRNA로 전기천공하고 도너 IDLV (2E+07TU/mL; MOI=50)로 형질도입시켰다. 글루타민, 폐니실린, 스트렙토마이신, 50 ng/ml SCF, 50 ng/ml Flt-3 및 50 ng/ml TPO (Peprotech)를 함유한 X-VIVO15 배지 (Lonza)를 포함하는 예비-자극 배지를 사용하였다. 회복 후 1일 후에, 0.1% BSA (Sigma)가 첨가된 PBS (Corning) 중의 1E+06 생존 세포를 6- 내지 8-주령의 NSG 마우스 (The Jackson Laboratory)에 250 cGy 총 신체 방사선 조사 후에 꼬리-정맥 주사에 의해 이식하였다. 나란히 배양하였지만 mRNA, 전기천공 또는 IDLV (mock-처리된)에 노출되지 않은 CD34+ 세포의 대조 샘플을 사용하였다. 다중 분석을 위해 작은 부분표본을 시험관 내에서 배양하였다. mPB 샘플을 전기천공 후 1일에 냉동시켰고, 나중에 해동하였으며 상기와 같이 이식 전 1일에 회복시켰다.

[0212] 인간 세포의 이식을 유동 세포분석법에 의해 이식 후 5주 후에 V450-컨쥬케이트된 항-인간 CD45 대 APC-컨쥬케이트된 항-쥐과 CD45를 사용하여 평가하였다. 항체들과 인큐베이션 후에, 적혈구를 BD FACS-용해 용액 (BD Biosciences)으로 용해시켰다. 이식의 백분율을 %huCD45+/(%huCD45+ + %muCD45+)로 규정하였다. 말초혈 및 BM에서의 이식 및 혈통 분포를 이식 후 8주 및 16주 후에 V450-컨쥬케이트된 항-인간 CD45, V500-컨쥬케이트된 항

-쥐과 CD45, FITC-컨쥬게이트된 항-인간 CD33, PerCP-컨쥬게이트된 항-인간 CD3, PE-컨쥬게이트된 항-인간 CD56, PE-Cy7-컨쥬게이트된 항-인간 CD19, 및 APC-컨쥬게이트된 항-인간 CD34 (모든 항체는 BD Biosciences)를 사용하여 유동 세포분석에 의해 평가하였다.

[0213] 결과 변수들에 대한 평균 및 표준 편차를 포함하는 기술 통계학을 기록하고 그래프로 나타냈다. 유전자 변형 수준, 이식 및 혈통과 관련된 정량적 결과에 대해, 쌍-방식의 비교를 쌍을 이루지 않은 t-시험에 의해 또는 2개 이상의 그룹이 있다면 일원변량분석의 틀 안에서 수행하였다. 용량-반응 분석을 수행하여 ZFN 및 도너 비율 최적화를 평가하였다. 구체적으로, 본 발명자들은 선형 혼합 모델 접근법을 사용했지만 무작위 효과와 같은 연속 적이고 고정된 효과 및 실험으로 용량을 처리하였다. 시간 경과에 따른 SCD 환자 BM 세포의 배수 팽창을 반복 변량분석에 의해 평가하였다. 본 발명자들의 통계학적 연구에서, 가설 시험은 2번이었고, p-값에 대해 중요 한 계값 0.5를 사용하였다. 모든 통계학적 분석을 SAS 버전 9.3 (SAS Institute Inc. 2012)을 사용하여 수행하였다.

[0214] 시험관 내에서의 세포의 병렬 배양은 유전자 변형률이 qPCR RFLP 분석에 의해 5% 내지 20%의 범위이고, 평균 값이 $14.5 \pm 6.4\%$ 인 것을 보였다 (도 10). 대용량 서열화로는 $10.5 \pm 4.0\%$ 의 대립유전자가 겹상 위치에서 $32.0 \pm 9.9\%$ 에서 볼 수 있는 삽입 또는 결실 (indel)로 교환된 염기를 함유한 것으로 나타났다. 또한, 이들 세포의 조혈 잠재력을 시험관 내 콜로니-형성 분석으로 평가하였다; 세포는 모든 분석 가능한 혈통에서 미처리 mock 샘플과 관련하여 광역 콜로니-형성 능력을 유지하였다.

[0215] 주사 후 5 및 8주 후에, 이식된 마우스들을 인간 및 쥐과 둘 다의 총 CD45+ 세포 중 인간 CD45+ 세포의 백분율로서 측정된, 인간 HSPC의 이식에 대해 평가하였다. ZFN 및 IDLV-처리된 HSPC의 이식 수준은 미처리 대조표준의 그것과 비교할만하였는데, 5주째에 평균 $14.8 \pm 11.4\%$ 이었고 8주째에 45%로 증가하였다 (도 10C). 마우스의 말초 혈의 혈통 분석 (T 세포에 대해 CD3, 골수성 세포에 대해 CD33, HSPC에 대해 CD34, B 세포에 대해 CD19 및 NK 세포에 대해 CD56)은 이 모델에서 예상한 대로였다 (도 11). ZFN 및 올리고뉴클레오티드 도너 둘 다로 처리된 세포를 받은 마우스에서 인간 세포의 분석은 ZFN 및 IDLV-처리된 HSPC의 그것과 유사한 HSPC 분화 스펙트럼을 나타냈다 (도 10E).

[0216] 마우스들을 이식 후 16주에 안락사시켜서 인간 세포 이식 및 유전자 변형 수준의 평가를 허용하였다. ZFN 및 IDLV-처리된 세포를 받은 마우스에 대해, 말초혈의 평가는 고수준의 이식 및 예상한 혈통 분포를 나타냈다. 골수 (BM) 구획의 분석은 유사하였다 (도 12). 마찬가지로, ZFN 및 올리고뉴클레오티드 둘 다로 처리된 세포를 받은 마우스에서, 이식 수준 및 혈통 분포는 말초혈 및 BM에서 ZFN 및 IDLV-처리된 세포를 받은 마우스의 집단에서 볼 수 있는 것과 유사하였다.

[0217] 마우스에 존재하는 인간 세포가 실제로 ZFN 시약에 의해 변형된 것인지를 측정하기 위하여, 게놈 DNA를 각 마우스의 BM 및 비장 조직으로부터 분리하고 인간 베타-글로빈 유전자를 탐색하였다. ZFN 및 IDLV-처리된 세포를 받은 마우스의 조직에서 qPCR에 의한 HhaI RFLP의 분석은 이들 마우스에서 유전자 변형이 일어났음을 나타냈지만, 인풋 세포에 대해 관찰된 10.5% 유전자 변형 수준보다 적은 빈도였다 (BM 및 비장에 대해 각각 $0.14 \pm 0.32\%$ 및 $0.16 \pm 0.18\%$) (도 13). 서열 분석은 BM 및 비장에서 인간 세포의 겹상 돌연변이에서의 유전자 변형을 확인해주었는데, 각각 평균 $0.21 \pm 0.39\%$ 및 $0.27 \pm 0.31\%$ 으로 비교할만하였다 (도 13B). 절단 부위에서 NHEJ에 의해 유발된 삽입 및 결실 (indel)의 분석은 HDR에 의한 것보다 높은 수준의 변화를 나타냈는데, indel을 함유한 BM에서 모든 서열 판독의 $4.8 \pm 7.8\%$ 였고 비장에서는 $3.8 \pm 3.7\%$ 였다 (도 14).

[0218] ZFN 및 올리고뉴클레오티드-처리된 세포를 받은, 투여 전에 벌크 집단에서 겹상 염기에서 17.3%의 유전자 변형 및 19.8% indel을 가진 마우스의 BM 및 비장의 게놈 분석 또한 유전자 변형의 유지를 증명하였다. 이들 조직으로부터의 DNA에 대용량 서열화를 수행하였고, 그 결과 BM 및 비장에서 각각 $0.85 \pm 0.81\%$ 및 $2.11 \pm 1.19\%$ 의 표적화된 유전자 변형을 나타냈다 (도 13C). 이들 조직에서 indel의 평가는 BM에서 $3.34 \pm 2.65\%$ 및 비장에서 $5.86 \pm 2.30\%$ 의 수준을 나타냈다 (도 14).

[0219] 이들 결과는 ZFN에 의한 부위-특이적 DNA 절단 및 상동성-지시된 유전자 변형이 진행된 CD34+ 세포가 이식할 수 있고 다중-혈통 분화를 진행할 수 있음을 증명한다.

0220] 실시예 8: 겹상 골수에서 유전자 교정

[0221] 겹상-적혈구 질환 환자들은 G-CSF를 사용한 줄기세포 이동에 대한 후보가 아니기 때문에, 본 발명자들은 이들 환자의 BM 흡입물로부터 CD34+ HSPC를 얻었다. 부위-특이적 유전자 교정을 ZFN mRNA 및 IDLV 도너를 사용하여 수행하였다. 세포를 적혈구 팽창 배지에 넣고 계속해서 수립된 방법을 사용하여 분화시켰다 (Romero et al, 상

기 동일, Giarratana *et al*, 상기 동일). 또한, 세포의 일부를 콜로니-형성 가능성에 대해 평가하였다. ZFN+IDLV (쌍 33488/33501)로 처리된 세포는 mock, 비-전기천공된 대조 샘플에 비교하여 적당하게 낮은 (35%) 콜로니-형성 능력을 보였다 (도 15).

[0222] 초기 적혈구 팽창에 이어서 (그러나 적출 전에) 세포를 계놈 분석을 위해 수확하였다. 이들 샘플의 qPCR에 의한 RFLP 분석은 평균 $20.1 \pm 8.8\%$ 의 유전자 변형 수준을 나타냈다 (도 16A). 이들 결과는 심층 서열화에 의해 확인되었는데, 판독의 $18.4 \pm 6.7\%$ 에서 SCD 돌연변이의 교정을 보여준다 (도 16B). 추가로, 서열화는 Δ SCD 돌연변이의 교정을 함유한 대부분의 대립유전자가 또한 HhaI 위치에서의 염기 변화를 함유하였음을 확인해주었고, 대부분의 HDR-유도된 사건이 그런 두 개의 염기 사이에 적어도 22 bp의 거리를 포함하는 것을 나타낸다. 적혈구 배양이 끝난 후에, 샘플을 고압 액체 크로마토그래피 (HPLC)에 의한 글로빈 사량체의 분석을 위해 수집하였다 (도 16C).

[0223] ZFN 및 IDLV-처리된 샘플로부터 유도된 적혈구계 세포에서 HbA 피크의 존재는 유전자 교정이 베타S 대립유전자의 베타A 대립유전자로의 기능적 전환을 유도하였음을 증명한다. 그런 피크는 mock-처리된 세포로부터 유도된 적혈구에서는 볼 수 없었다. HbF 피크는 ZFN 처리된 샘플에서 HbS 피크의 감소로 인한 상대적인 증가를 보여주었다. ZFN 및 IDLV-처리된 샘플에서 HbA의 상대적 유도는 평균 $5.3 \pm 0.02\%$ 이었고, 단백질 교정 수준은 10.0%까지였다 (도 16D).

[0224] 이를 결과는 겸상-적혈구 질환 환자들의 골수로부터 유래된 HSPC의 표현형을 교정하는 IDLV 도너와 조합된 ZFN의 능력을 증명한다.

[0225] 여기서 보여준 결과들은 인간 조혈 줄기/선구 세포에서 시험관 내 징크-핑거 뉴클레아제를 사용하여 겸상-적혈구 질환 돌연변이의 고수준의 유전자 교정을 증명한다. 본 연구에서, 본 발명자들은 ZFN 쌍이 베타-글로빈 유전자좌를 절단하도록 디자인하였다. 상동성 도너 주형 (울리고뉴클레오티드로서 또는 인테그라제-결핍 렌티바이러스 벡터를 통해 전달됨)과 조합하여, 이를 ZFN은 선구 세포에서 고수준으로 상동성-지시된 수복을 유도할 수 있다.

[0226] ZFN 절단 부위의 분석은 대부분의 뉴클레아제 활성이 베타-글로빈 유전자좌 표적 부위에서 일어나고, 이때 상동성 델타-글로빈 유전자에서는 절단이 더 작은 부분으로 일어나는 것을 나타냈다. 성인 적혈구 세포에서 저수준의 델타-글로빈 발현 (모든 글로빈의 <3.5%)으로 인해, 세포의 하위세트에서 이 유전자에서의 ZFN 활성을 해로운 것으로 보이지 않는다. IDLV 단부-포획을 사용하여 K562 세포에서 ZFN 표적-외 절단 부위의 편향되지 않은, 계놈-전체 평가는 ZFN의 높은 특이성을 증명하였고, 이때 자연적으로 파괴되기 쉬운 부위로의 바탕 통합만을 볼 수 있었다.

[0227] ZFN 및 도너-처리된 세포를 마우스에 이식했을 때, 세포의 이식 및 혈통 분포는 mock-처리된 및 ZFN과 도너-처리된 샘플 사이에 동등하였다. 그러나, 이식 전의 시험관 내 샘플에서 평균 10 내지 20%의 유전자 변형 수준에도 불구하고, 이식 후 16주 후의 마우스의 비장 및 BM에서 인간 세포에서의 유전자 교정 수준은 현저하게 더 낮았다. 이를 발견은 기술분야에서 최근에 발표된 작업과 일관되며 (Genovese *et al* (2014) *Nature* 510:235-240) 더 많은 성숙한 선구 세포가 효과적으로 교정되는 한편으로, 지속적인 유익 (예컨대 임상적 유익)이 선행의, 보다 원시적인 조혈 줄기세포의 변형에 의해 제공되는 것을 함축할 수 있다.

[0228] 부위-특이적 유전자 파괴의 효율 (CCR5 파괴와 관련하여 Holt *et al* (2010) *Nat biotech* :839-47 참조)과 비교하여 보다 원시적인 HSPC에서의 상동성-지시된 유전자 교정의 효율은 교정적인 염기를 함유하는 도너 주형을 사용하는 데에서의 차이로 인한 것일 수 있고, 그것은 공-전달되며 세포의 DNA 손상 수복 경로는 NHEJ 보다는 HDR을 사용하여 DSB를 해결해야 한다. NHEJ가 조용한, 원시적인 HSC에서 선호되기 때문에 (Mohrin *et al* (2010) *Cell stem cell* 7:174-85 참조), 수복 경로 선택에서의 편향은 원시적인 HSC에서 유전자 교정을 제한해야 할 것이다. 그러므로, 한 이론에 의해 구속되지 않으면서, ZFN이 성숙한 및 원시적인 세포 집단에서 유사한 효율로 작용하는 것이 가능하고, 그러나 각 세포 유형에서 활성 수복 경로는 달라서 HDR은 보다 성숙한 세포에서 보다 활성이고 원시적인 HSC에서 그렇게 적어진다. 특히, 마우스에 이식된 세포는 유전자 변형에 대한 것보다 더 큰 정도로 indel의 인풋 수준을 유지하였다 (IDLV 및 울리고 실험에 대한 유전자 변형에서 각각 43.9 및 11.7배 변화와 비교하여 indel의 7.4 및 4.3배 변화).

[0229] 겸상-적혈구 질환 환자들로부터의 BM CD34+ 세포에서의 실험은 임상적 번역에 대한 전망을 제공한다. 표본적인 겸상 돌연변이 (적어도 적혈구계 선구 세포들에서)의 교정 수준은 이들 실험에 대해 평균 18%였고, 분화시, 이들 세포는 교정된, 야생형 헤모글로빈 (HbA)을 생성하였다. SCD에 대한 동종이계의 조혈 줄기세포 이식으로부터

의 데이터를 기반으로, 10 내지 30%의 도너 키메라 현상은 정상적인, 도너-유도된 적혈구의 선택적 장점의 결과로서 유의미한 임상적 개선을 초래할 수 있다 (Andreani *et al* (2011) *Chimerism* 2(1):21-22 및 Walters *et al* (2001) *Am Soc Blood and Marrow Transpl* 7:665-73 참조). 또한, SCD 돌연변이에 대한 이형접합체는 정상적으로 질환의 증상을 경험하지 않는다. 그러므로, 각 HSC에서 단지 하나의 대립유전자를 교정하는 것이 SCD와 관련된 증상을 중 적어도 일부를 완화시키는 것을 충분히 증명할 수 있다.

[0230] 혈모글로빈증, 특히 SCD에 대한 렌티바이러스 기반 유전자 요법에서의 최근의 진보 (Romero *et al*, 상기 동일, 및 Chandrakasan and Malik (2014) *Hematol/oncol Clin of North Amer* 28:199-216 참조)에도 불구하고, 치료용 전이유전자의 장기-지속성 및 적절하게 조절된 발현에 대한 요구로 인해 잠재적인 문제가 남아있다. HSC에서 표본적인 A 내지 T 질환-유발 겹상 형질도입의 표적화된 뉴클레아제를 사용하는 부위-특이적 교정은 그것의 내인성 프로모터 및 유전자좌 제어 영역 하에 베타-글로빈의 발현을 유지하는 독특한 능력을 제공한다. 나아가, 게놈 교정 시약은 영구적인 교정을 초래하기 위해 세포의 1회, 일시적인, 생체 외 처리만을 필요로 한다.

[0231] 이들 데이터를 함께 취하면 SCD에 대한 잠재적인 치료로서 HSPC에서의 게놈 편집의 지속적인 발달을 지지한다.

[0232] 본원에서 언급된 모든 특허, 특히 출원 및 공보는 그것들의 전체 내용이 참조로 본원에 포함된다.

[0233] 발명이 비록 이해를 분명히 할 목적으로 예시 및 실시예에 의하여 일부 상세하게 제공되었지만, 기술분야의 숙련자들에게는 다양한 변화 및 변형이 발명의 사상 또는 범주로부터 벗어나지 않으면서 실시될 수 있다는 것이 명백해질 것이다. 따라서, 전술한 설명 및 실시예들은 제한하는 것으로 해석되지 않아야 한다.

도면

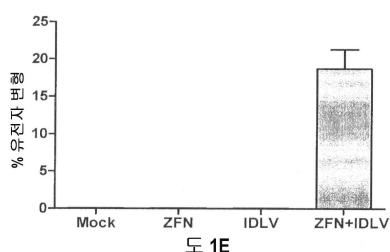
도면1a

도 1A

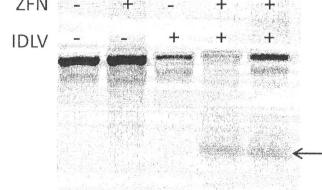
5'-TCAAACAGACACCATGGTGCATCTGACTCCTGAGGAGAAGTCTGCCGTTACTGCCCTGTGGGGC



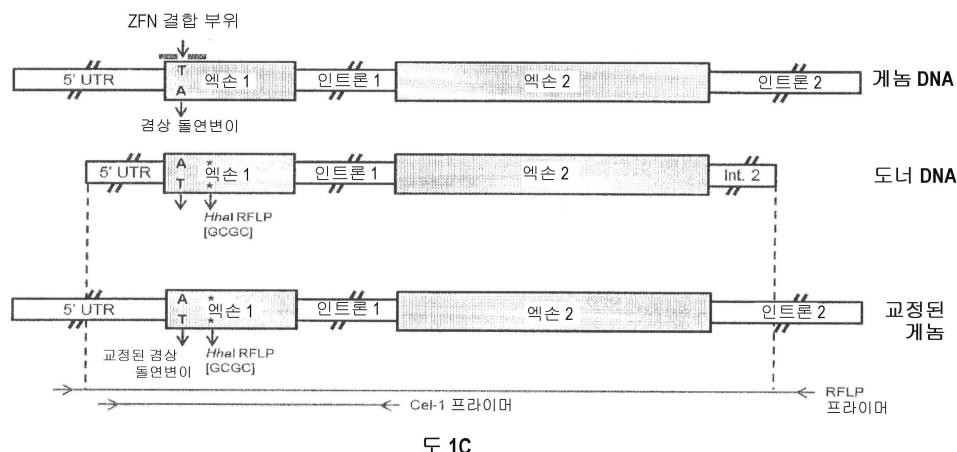
도 1B



도 1D

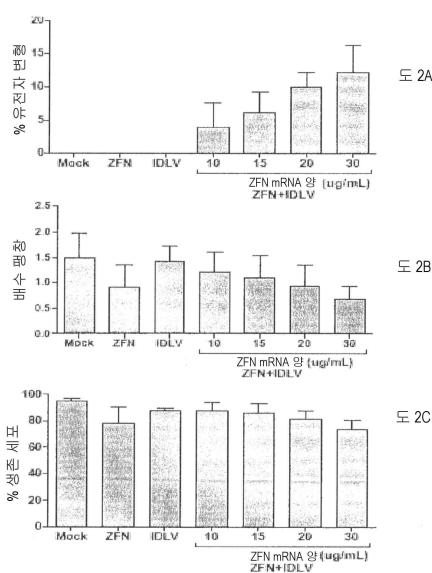


도면1b

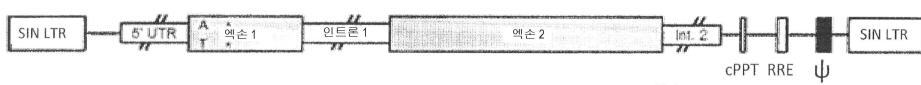


도 1C

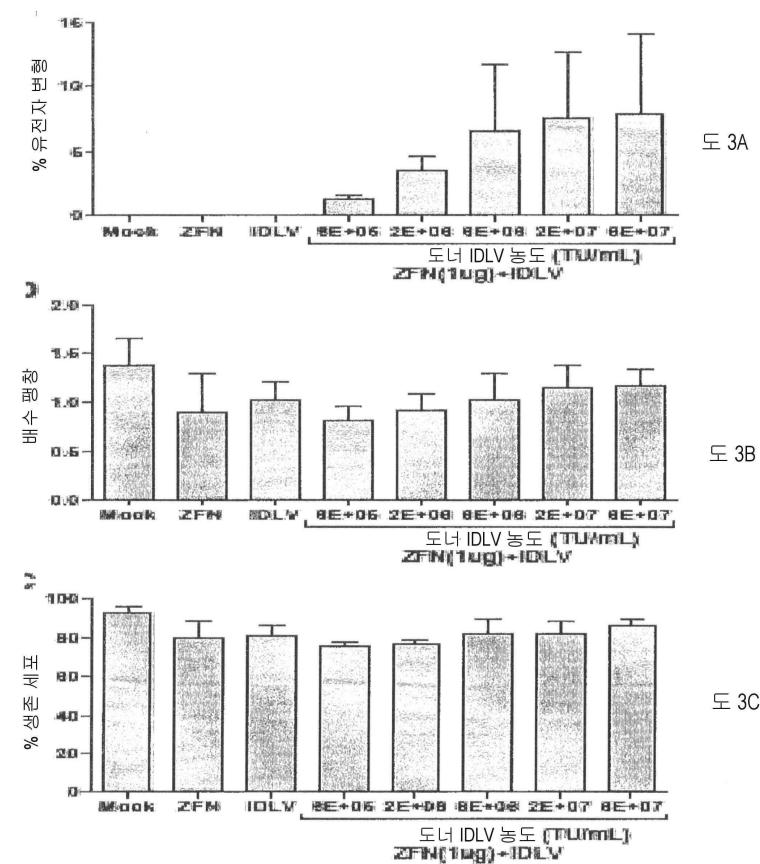
도면2



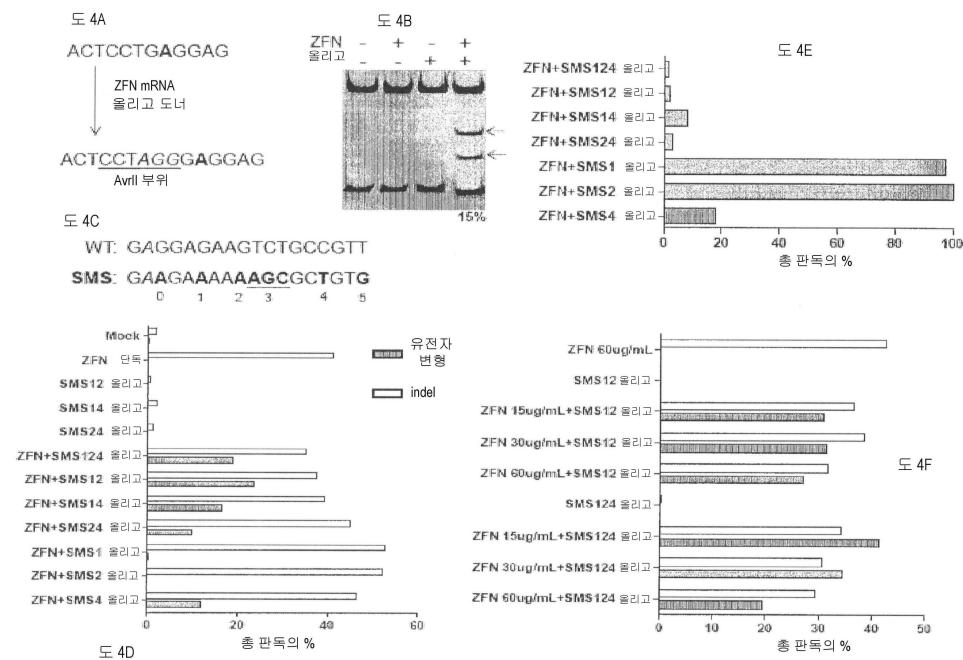
도 2D



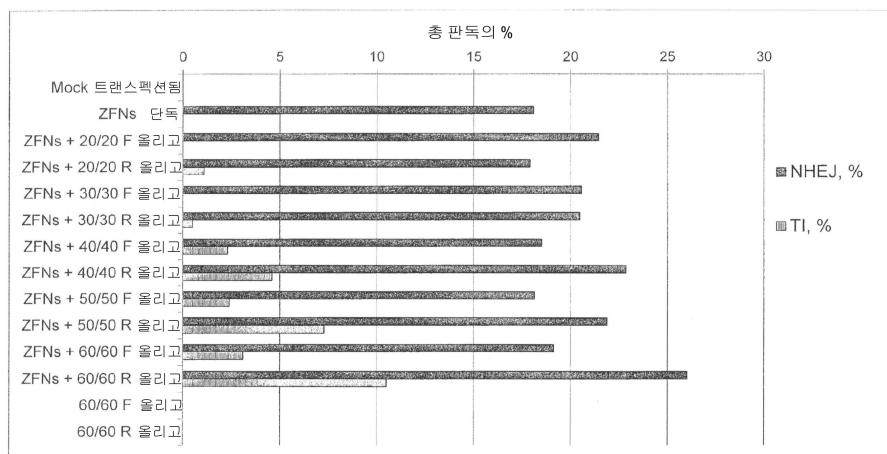
도면3



도면4

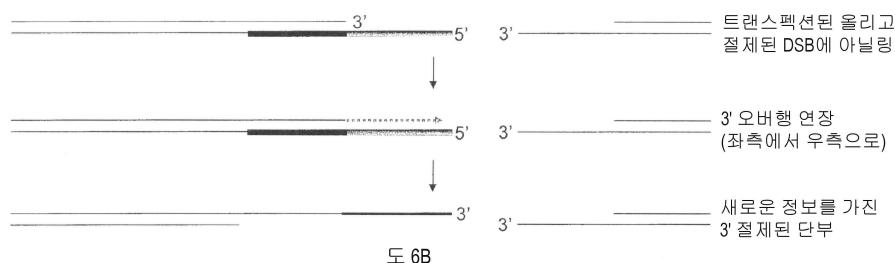


도면5



도면6

| 좌측에서 우측으로 | | | 우측에서 좌측으로 | | | 도 6A |
|-----------|-------|--------|-----------|-------|------|------|
| SMS1 | SMS12 | SMS124 | SMS124 | SMS24 | SMS4 | |
| 0.03 | 6.0 | 35.2 | 35.2 | 0.03 | 0.2 | |



도면7a



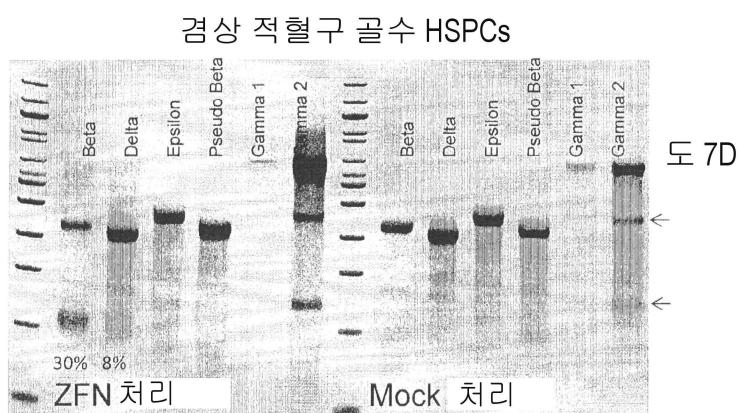
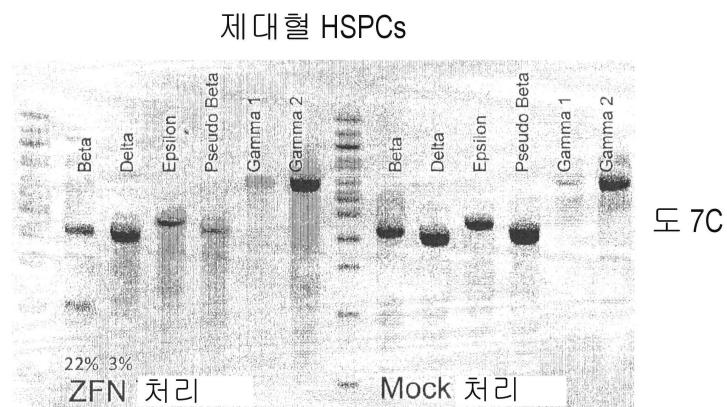
도 7B

```

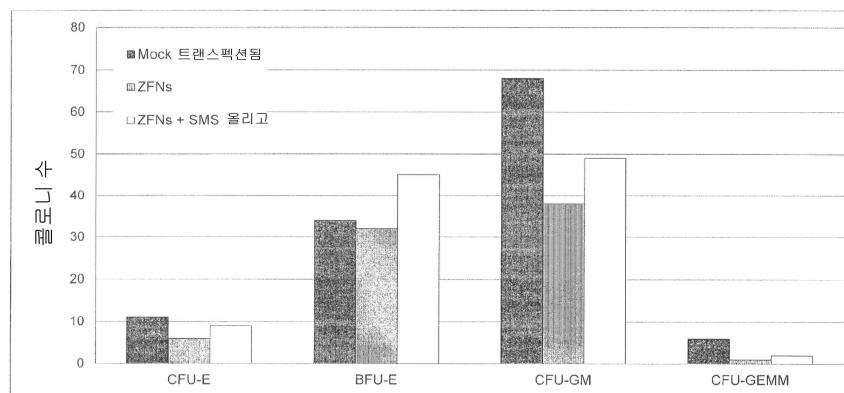
HBB      C A C C A T G G T G C A T C T G A C T C C C T G A G G A G A A G T C T G C C G T T A
HBD      C A C C A T G G T G C A T C T G A C T C C C T G A G G A G A A G T C T G C C G T T A
HBE1     C A T C A T G G T G C A T T T T A C T G C T G A G G A G A A G G C T G C C G T C A
HBPP1    C A C T G T A G T G C A T T T C A C T G C T G A C A A G A A G G C T G C T G C C A
HBG1     C G C C A T G G G T C A T T T C A C A G A G G G A G G A C A A G G C T A C T A T C A
HBG2     C G C C A T G G G T C A T T T C A C A G A G G G A G G A C A A G G C T A C T A T C A

```

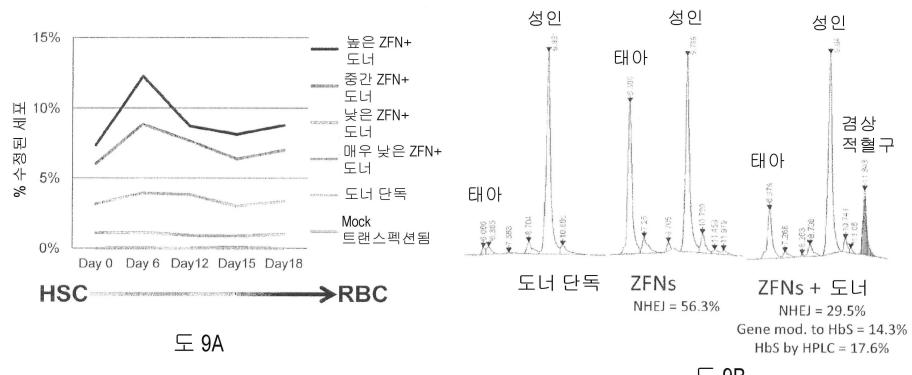
도면7b



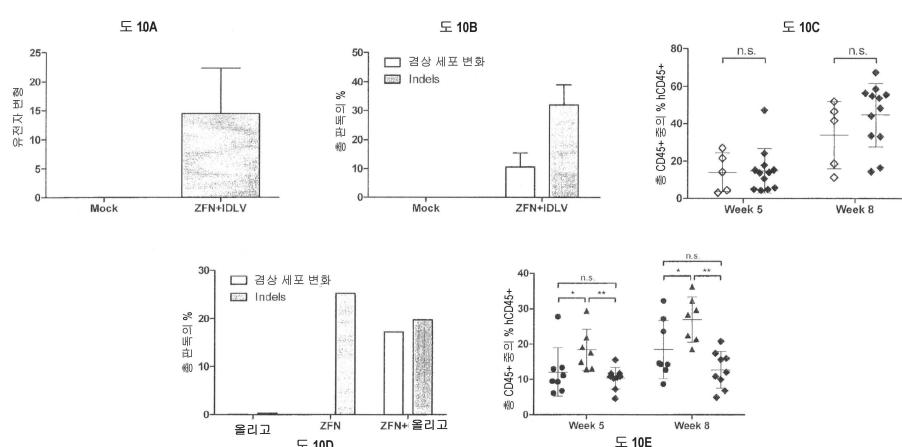
도면8



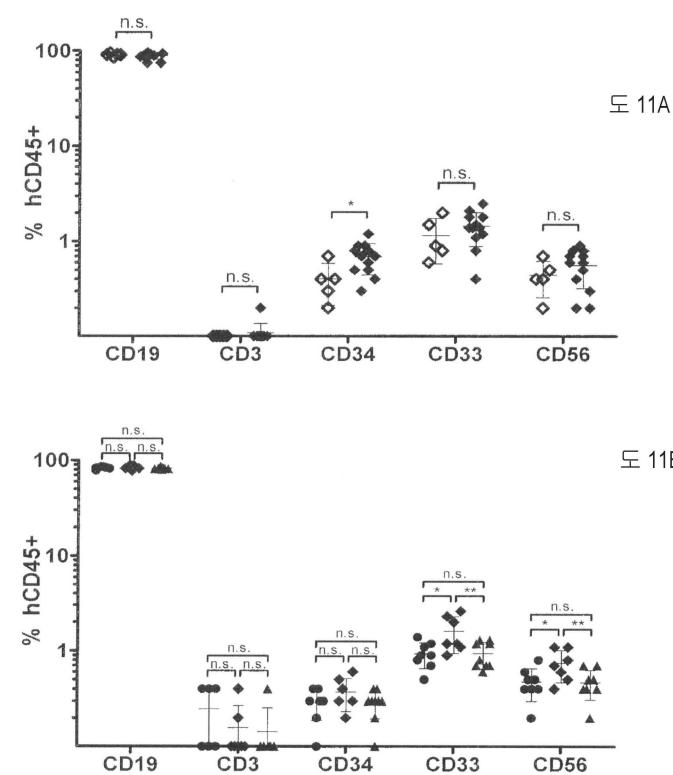
도면9



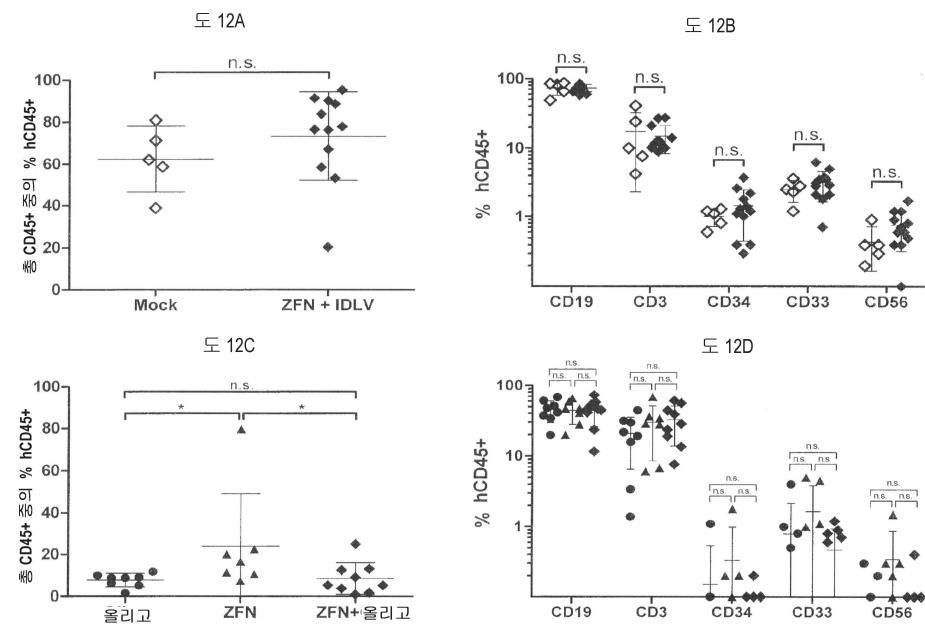
도면10



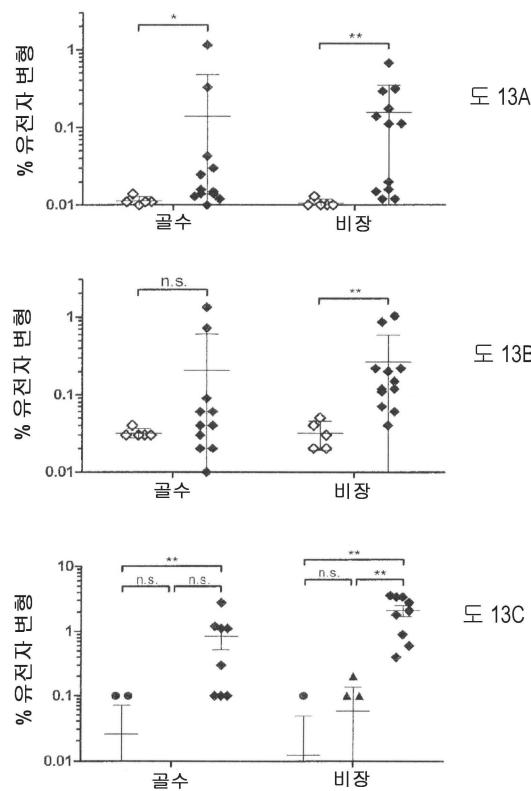
도면11



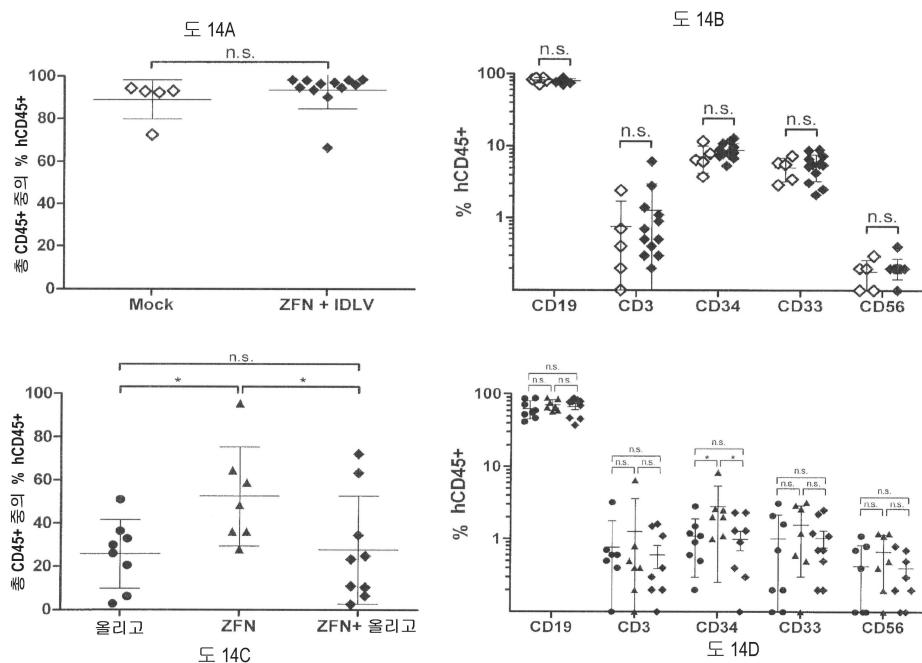
도면12



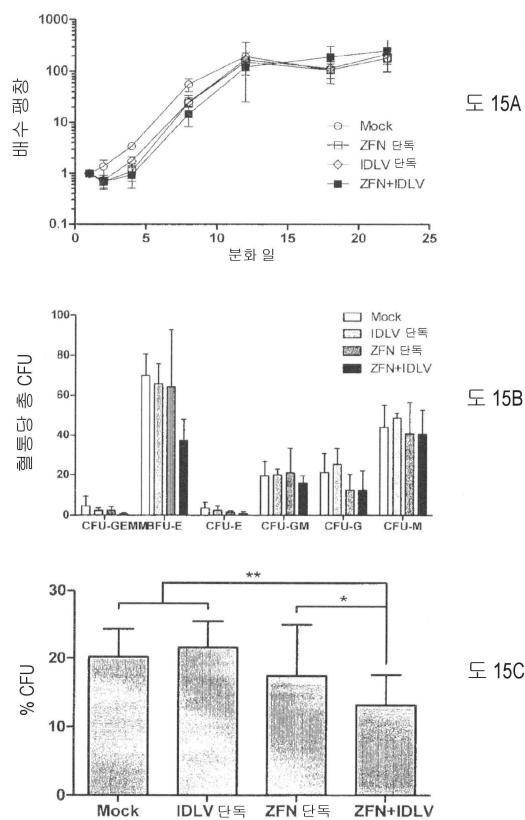
도면13



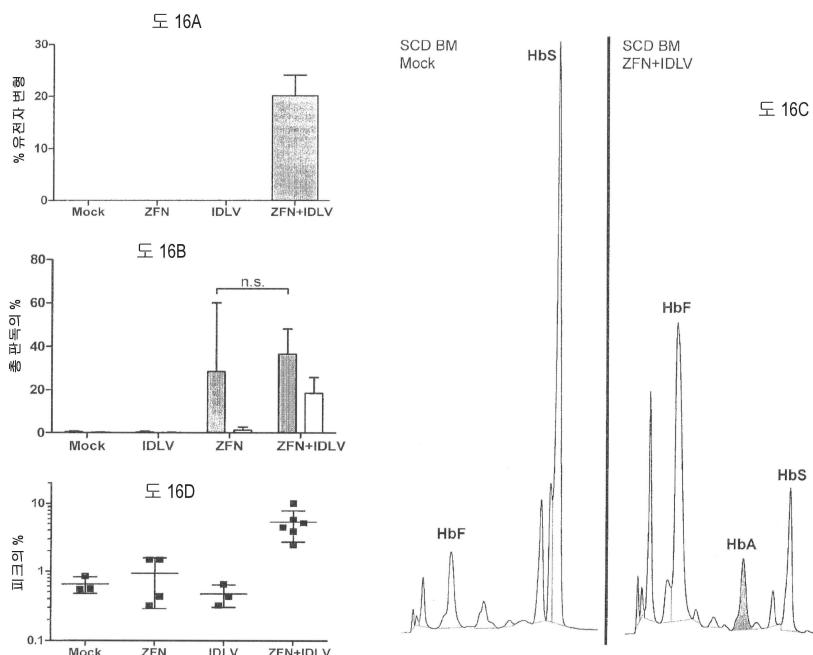
도면14



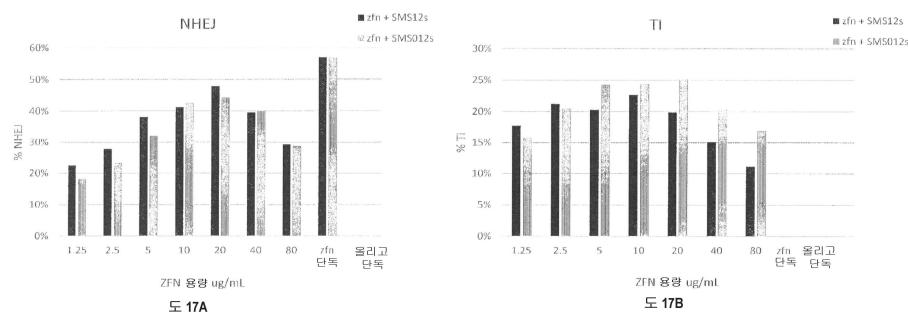
도면15



도면16



도면17



서 열 목록

SEQUENCE LISTING

<110> SANGAMO BIOSCIENCES, INC.

<120> METHODS AND COMPOSITIONS FOR NUCLEASE-MEDIATED GENOME

ENGINEERING AND CORRECTION IN HEMATOPOIETIC STEM CELLS

<130> 8328-0127.40

<140> PCT/US2015/050411

<141> 2015-09-16

<150> 62/051,159

<151> 2014-09-16

<160> 71

<170> PatentIn version 3.5

<210> 1
<211> 25
<212> DNA
<213> Homo sapiens
<400> 1
ctgacacctt ctcttcctcc cacag 25
<210> 2
<
211> 13
<212> DNA
<213> Homo sapiens
<400> 2
tttctctcca cag 13
<210> 3
<211> 64
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 3
tcaaacagac accatgggtgc atctgactcc tgaggagaag tctgccgtta ctgccctgtg 60
gggc 64
<210> 4
<211> 12
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 4
actcctgagg ag 12
<210> 5
<211> 15
<212> DNA
<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 5

actccctaggg aggag 15

<210> 6

<211> 41

<212> DNA

<213> Homo sapiens

<400> 6

caccatggtg catctgactc ctgaggagaa gtcgtccgtt a 41

<210> 7

<211> 41

<212> DNA

<213> Homo sapiens

<400> 7

caccatggtg catctgactc ctgaggagaa gactgctgtc a 41

<210> 8

<211> 41

<212> DNA

<213> Homo sapiens

<400> 8

catcatggtg cattttactg ctgaggagaa ggctgccgtc a 41

<210> 9

<211> 41

<212> DNA

<213> Homo sapiens

<400> 9

cactgttgtc catttcactg ctgacaagaa ggctgctgcc a 41

<210> 10

<211> 41

<212> DNA

<213> Homo sapiens

<400> 10

cgccatgggt cattcacag aggaggacaa ggctactata a

41

<210> 11

<211> 41

<212> DNA

<213> Homo sapiens

<400> 11

cgccatgggt cattcacag aggaggacaa ggctactata a

41

<210> 12

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 12

Leu Arg His His Leu Thr Arg

1 5

<210> 13

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 13

Leu Arg His Asn Leu Arg Ala

1 5

<210> 14

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 14

Asp Gln Ser Asn Leu Arg Ala

1 5
<210> 15
<211> 7
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic peptide
<400> 15
Arg Asn Ala Ser Arg Thr Arg
1 5
<210> 16
<211> 7
<212> PRT
<213> Artificial Sequence
<220>
<223> Description of Artificial Sequence: Synthetic peptide
<400> 16
Arg Ser Asp Asn Leu Ser Glu
1 5
<210> 17
<211> 7
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic peptide
<400> 17
Arg Ser Gln His Arg Lys Thr
1 5
<210> 18
<211> 7
<212> PRT
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic peptide

<400> 18

Thr Ser Gly Ser Leu Ser Arg

1 5

<210> 19

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 19

Asp Arg Ser Asp Leu Ser Arg

1 5

<210> 20

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 20

Asp Arg Ser Ala Leu Ala Arg

1 5

<210> 21

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

peptide

<400> 21

Gln Ser Ser Asn Leu Ala Arg

1 5

<210> 22

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
peptide

<400> 22

Gln Ser Gly His Leu Ser Arg

1 5

<210> 23

<211> 28

<212> DNA

<213> Homo sapiens

<400> 23

ggagttaggt gcaccatggc gtctgttt 28

<210> 24

<211> 28

<212> DNA

<213> Homo sapiens

<400> 24

gtggagaagt ctgcgttac tgccctgt 28

<210> 25

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 25

gacaggttacg gctgtcatca 20

<210> 26

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 26

| | |
|--|----|
| cagcctaagg gtggaaaaat | 20 |
| | |
| <210> 27 | |
| <211> 24 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> Description of Artificial Sequence: Synthetic | |
| primer | |
| <400> 27 | |
| atgcttagaa ccgaggtaga gttt | 24 |
| | |
| <210> 28 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> Description of Artificial Sequence: Synthetic | |
| primer | |
| <400> 28 | |
| cctgagactt ccacactgat g | 21 |
| | |
| <210> 29 | |
| <211> 19 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220> | |
| ><223> Description of Artificial Sequence: Synthetic | |
| primer | |
| <400> 29 | |
| gaagtctgcc gttactgcg | 19 |
| | |
| <210> 30 | |
| <211> 20 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> Description of Artificial Sequence: Synthetic | |
| primer | |
| <400> 30 | |
| cccagttct attggctcc | 20 |

<210> 31

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 31

ctcggcgct ttagtgatgg 20

<210> 32

<211> 19

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 32

gactcacccct gaagttctc 19

<210> 33

<211> 44

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<220><221> modified_base

<222> (21)..(24)

<223> a, c, t, g, unknown or other

<400> 33

acacgacgct ctccgatct nnngggctg ggcataaaag tcag 44

<210> 34

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 34

gacgtgtgct ctccgatct tccacatgcc cagtttctat t 41

<210> 35

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<220><221> modified_base

<222> (21)..(24)

<223> a, c, t, g, unknown or other

<400> 35

acacgacgct ctccgatct nnntaaaag gcagggcaga gt cga 45

<210> 36

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 36

gacgtgtgct ctccgatct acatgccag tttccatttg c 41

<210> 37

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<220><221> modified_base

<222> (21)..(24)

<223> a, c, t, g, unknown or other

<400> 37

acacgacgct ctccgatct nnntgctt ccgacacagc tgcaa 45

<210> 38

<211> 41

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 38

gacgtgtgct ctccgatct tcaccctca ttcccatgca t

41

<210> 39

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<220><221> modified_base

<222> (21)..(24)

<223> a, c, t, g, unknown or other

<400> 39

acacgacgct ctccgatct nnnggaacg tctgaggta tcaat

45

<210> 40

<211> 38

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 40

gacgtgtgct ctccgatct tcctccctc cttgtcc

38

<210> 41

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 41

| | |
|--|----|
| ggttcattt tcatttcac a | 21 |
| | |
| <210> 42 | |
| <211> 21 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> Description of Artificial Sequence: Synthetic | |
| primer | |
| <400> 42 | |
| gtaatctgag ggttagaaaa c | 21 |
| | |
| <210> 43 | |
| <211> 20 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> Description of Artificial Sequence: Synthetic | |
| primer | |
| <400> 43 | |
| gagactgtgg gatagtata | 20 |
| | |
| <210> 44 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220> | |
| ><223> Description of Artificial Sequence: Synthetic | |
| primer | |
| <400> 44 | |
| cattatcaca aacttagtgt cc | 22 |
| | |
| <210> 45 | |
| <211> 22 | |
| <212> DNA | |
| <213> Artificial Sequence | |
| <220><223> Description of Artificial Sequence: Synthetic | |
| primer | |
| <400> 45 | |
| agtctatcaa atgacaccat at | 22 |

<210> 46

<211> 24

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 46

gcaaaggcta taaaaaaaat tagc

24

<210> 47

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 47

gagatcatcc aggtgctttg

20

<210> 48

<211> 27

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 48

ggcaaaggct ataaaaaaaaa ttaagca

27

<210> 49

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
primer

<400> 49

gagatcatcc aggtgcttta

20

<210> 50

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 50

| | |
|----------------------|----|
| acatggctt ctgacacaac | 20 |
|----------------------|----|

<210> 51

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 51

| | |
|-------------------------|----|
| gaaattggac agcaagaaag c | 21 |
|-------------------------|----|

<210> 52

<211> 20

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

primer

<400> 52

| | |
|------------------------|----|
| cacccttgcac caatagattc | 20 |
|------------------------|----|

<210> 53

<211> 18

<212> DNA

<213> Homo sapiens

<400> 53

| | |
|---------------------|----|
| gaggagaagt ctgccgtt | 18 |
|---------------------|----|

<210> 54

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 54

gaagaaaaaa gcgcgttg 18

<210> 55

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 55

Leu Lys Gln Asn Leu Asp Ala

1 5

<210> 56

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 56

Arg Asn Cys Ala Arg Leu Arg

1 5

<210> 57

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

peptide

<400> 57

Arg Asn Gln Thr Arg Leu Asn

1 5

<210> 58

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic peptide

<400> 58

His His His Ser Leu Lys Arg

1 5

<210> 59

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic peptide

<400> 59

Thr Asn Ser Glu Leu Asp Arg

1 5

<210> 60

<211> 18

<212> DNA

<213> Homo sapiens

<400> 60

gaggagaagt ctgccgtt 18

<210> 61

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 61

gaggaaaagt ctgccgtt 18

<210> 62

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 62

| | |
|--------------------|----|
| gaggagaat ctgccgtt | 18 |
|--------------------|----|

<210> 63

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 63

| | |
|---------------------|----|
| gaggagaagt ctgctgtt | 18 |
|---------------------|----|

<210> 64

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 64

| | |
|---------------------|----|
| gagggaaaat ctgccgtt | 18 |
|---------------------|----|

<210> 65

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 65

| | |
|----------------------|----|
| gagggaaaagt ctgctgtt | 18 |
|----------------------|----|

<210> 66

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide
<400> 66

| | |
|----------------------|----|
| gaggagaaaat ctgctgtt | 18 |
|----------------------|----|

<210> 67

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 67

| | |
|----------------------|----|
| gaggagaaaat ctgctgtt | 18 |
|----------------------|----|

<210> 68

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 68

| | |
|---------------------|----|
| gaagaaaaat ctgccgtt | 18 |
|---------------------|----|

<210> 69

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 69

| | |
|---------------------|----|
| gtggaaaaat ctgccgtt | 18 |
|---------------------|----|

<210> 70

<211> 18

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 70

gtagaaaaat ctgccgtt

18

<210> 71

<211> 9

<212> PRT

<213> Unknown

<220><223> Description of Unknown: 'LAGLIDADG' family

meganuclease motif peptide

<400> 71

Leu Ala Gly Leu Ile Asp Ala Asp Gly

1

5