



등록특허 10-2572759



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년08월29일
(11) 등록번호 10-2572759
(24) 등록일자 2023년08월25일

- (51) 국제특허분류(Int. Cl.)
C12N 9/22 (2006.01) *A61K 35/28* (2015.01)
A61P 7/06 (2006.01) *C12N 15/85* (2006.01)
C12N 15/90 (2006.01)
- (52) CPC특허분류
C12N 9/22 (2013.01)
A61K 35/28 (2013.01)
- (21) 출원번호 10-2019-7005015
- (22) 출원일자(국제) 2017년08월24일
심사청구일자 2020년08월24일
- (85) 번역문제출일자 2019년02월20일
- (65) 공개번호 10-2019-0039530
- (43) 공개일자 2019년04월12일
- (86) 국제출원번호 PCT/US2017/048397
- (87) 국제공개번호 WO 2018/039440
국제공개일자 2018년03월01일
- (30) 우선권주장
62/378,978 2016년08월24일 미국(US)
(뒷면에 계속)
- (56) 선행기술조사문헌
Blood, 2015, Vol. 126, No. 23:204
WO2014036219 A2

(73) 특허권자
상가모 테라퓨틱스, 인코포레이티드
미국 캘리포니아 94005 브리즈번 마리나 불러바드
7000

(72) 발명자
밀러, 제프리 씨.
미국 94804 캘리포니아주 리치몬드 스위트
에이100 커넬 불러바드 501 포인트 리치몬드 테크
센터 상가모 테라퓨틱스, 인코포레이티드 내
레바, 에드워드 제이.
미국 94804 캘리포니아주 리치몬드 스위트
에이100 커넬 불러바드 501 포인트 리치몬드 테크
센터 상가모 테라퓨틱스, 인코포레이티드 내

(74) 대리인
양영준, 김영

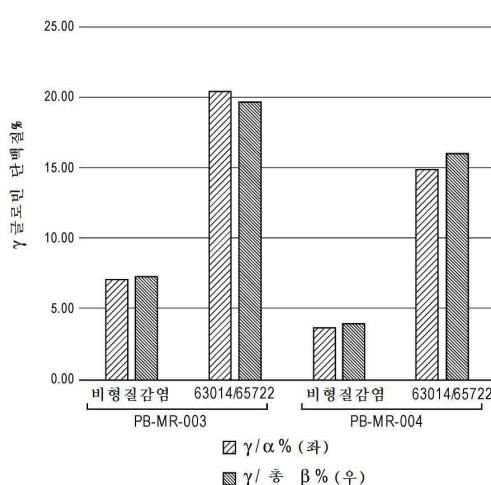
전체 청구항 수 : 총 25 항

심사관 : 한지혜

(54) 발명의 명칭 조작된 뉴클레아제를 이용하는 유전자 발현의 조절

(57) 요약

본 개시내용은 게놈 조작 분야, 특히 조혈 세포 게놈의 표적화된 변형에 관한 것이다.

대 표 도 - 도6

(52) CPC특허분류

A61K 38/00 (2013.01)
A61P 7/06 (2018.01)
C12N 15/85 (2013.01)
C12N 15/90 (2013.01)
C07K 2319/09 (2013.01)
C07K 2319/81 (2013.01)

(30) 우선권주장

62/443,981 2017년01월09일 미국(US)
62/545,778 2017년08월15일 미국(US)

명세서

청구범위

청구항 1

아연 평거 뉴클레아제 (ZFN)이며, 서열번호 29에 나타낸 바와 같은 아미노산 서열을 포함하는 좌측 ZFN 및 서열 번호 31에 나타낸 바와 같은 아미노산 서열을 포함하는 우측 ZFN을 포함하는 아연 평거 뉴클레아제.

청구항 2

제1항에 따른 하나 이상의 아연 평거 뉴클레아제를 암호화하는 폴리뉴클레오타이드.

청구항 3

제2항에 있어서, 상기 폴리뉴클레오타이드는 mRNA인, 폴리뉴클레오타이드.

청구항 4

제3항에 있어서, 서열번호 28 또는 서열번호 30을 포함하는 폴리뉴클레오타이드.

청구항 5

제1항의 아연 평거 뉴클레아제를 포함하는 단리된 세포.

청구항 6

제5항에 있어서, 상기 단리된 세포는 줄기 세포 또는 전구체 세포인, 단리된 세포.

청구항 7

제6항에 있어서, 상기 단리된 세포는 인간 세포인, 단리된 세포.

청구항 8

제5항 내지 제7항 중 어느 한 항에 있어서, 상기 단리된 세포의 게놈은 상기 아연 평거 뉴클레아제에 의해 변형되는 것인, 단리된 세포.

청구항 9

제8항에 있어서, 상기 게놈 변형은 삼입, 결절 및 이들의 조합으로 이루어진 군으로부터 선택되는 것인, 단리된 세포.

청구항 10

제8항에 있어서, 상기 게놈 변형은 BCL11A 인핸서 서열의 +58 영역 내인, 단리된 세포.

청구항 11

제5항의 단리된 세포를 증식, 유래 또는 분화시킴으로써 생산된 세포.

청구항 12

제8항에 있어서, 상기 단리된 세포는 상기 게놈 변형이 없는 세포에 비해 감마 또는 베타 글로빈의 증가된 발현을 나타내는 것인, 단리된 세포.

청구항 13

제1항에 따른 아연 평거 뉴클레아제 또는 제2항 내지 제4항 중 어느 한 항의 폴리뉴클레오타이드 또는 제5항 내지 제7항 중 어느 한 항의 단리된 세포를 포함하는, 이상혈색소증을 갖는 대상체에서 글로빈 생산을 증가시키는

데 사용하기 위한 약제학적 조성물.

청구항 14

세포에서 내인성 BCL11A 인핸서 서열을 변형시키는 시험관내 방법으로서, 상기 내인성 BCL11A 인핸서 서열이 변형되도록 상기 세포에 제1항에 따른 아연 평거 뉴클레아제 또는 제2항 내지 제4항 중 어느 한 항의 폴리뉴클레오타이드를 투여하는 단계를 포함하는, 세포에서 내인성 BCL11A 인핸서 서열을 변형시키는 시험관내 방법.

청구항 15

제14항에 있어서, 상기 세포 내로 외인성 서열을 도입하는 단계를 더 포함하며, 여기서 도입은 상기 외인성 서열이 상기 내인성 BCL11A 인핸서 서열 내로 삽입되도록 수행하는 것인, 세포에서 내인성 BCL11A 인핸서 서열을 변형시키는 시험관내 방법.

청구항 16

제14항에 있어서, 상기 변형은 결실을 포함하는 것인, 세포에서 내인성 BCL11A 인핸서 서열을 변형시키는 시험관내 방법.

청구항 17

제13항에 있어서, 상기 이상혈색소증은 베타-탈라세미아 또는 겸상 적혈구병인, 약제학적 조성물.

청구항 18

내인성 BCL11A 인핸서 서열 내에 게놈 변형을 포함하는 유전자 변형된 세포를 생산하는 시험관내 방법으로서,

- 세포를 제2항 내지 제4항 중 어느 한 항의 폴리뉴클레오타이드와 접촉시키는 단계;
- 상기 세포에 상기 폴리뉴클레오타이드로부터 융합 단백질을 발현시키는 데 도움이 되는 조건을 실시하는 단계; 및
- 상기 유전자 변형된 세포를 생산하는 데 충분한 상기 발현된 융합 단백질을 이용하여 상기 내인성 BCL11A 인핸서 서열을 변형시키는 단계

를 포함하는, 내인성 BCL11A 인핸서 서열 내에 게놈 변형을 포함하는 유전자 변형된 세포를 생산하는 시험관내 방법.

청구항 19

제18항에 있어서, 적어도 1종의 사이토카인을 이용하여 상기 세포를 자극하는 단계를 더 포함하는, 내인성 BCL11A 인핸서 서열 내에 게놈 변형을 포함하는 유전자 변형된 세포를 생산하는 시험관내 방법.

청구항 20

- (a) 제1항의 아연 평거 뉴클레아제 또는 (b) 제2항 내지 제4항 중 어느 한 항의 폴리뉴클레오타이드 중 하나 이상을 포함하는 키트.

청구항 21

제11항의 세포를 포함하는 키트.

청구항 22

제11항의 세포를 포함하는, 이상혈색소증을 갖는 대상체에서 글로빈 생산을 증가시키는데 사용하기 위한 약제학적 조성물.

청구항 23

제22항에 있어서, 상기 이상혈색소증은 베타-탈라세미아 또는 겸상 적혈구병인, 약제학적 조성물.

청구항 24

제2항 내지 제4항 중 어느 한 항의 폴리뉴클레오타이드를 포함하는 단리된 세포.

청구항 25

제11항의 세포로부터 생산된 세포주.

발명의 설명

기술 분야

[0001]

관련 출원에 대한 상호 참조

[0002]

본 출원은 2016년 8월 24일자로 출원된 미국 가출원 특허 제62/378,978호; 2017년 1월 9일자로 출원된 미국 가출원 특허 제62/443,981호; 및 2017년 8월 15일자로 출원된 미국 가출원 특허 제62/545,778호의 유익을 주장하며, 이 기초출원의 개시내용은 그들의 전문이 본 명세서에 참고로 포함된다.

[0003]

기술분야

[0004]

본 개시내용은 게놈 조작 분야, 특히 조혈 세포 게놈의 표적화된 변형에 관한 것이다.

배경기술

[0005]

게놈 서열분석 노력은 인간 게놈이 20,000 내지 25,000개의 유전자(그러나 2000개 보다 적은 전사 조절자)를 함유한다는 것을 나타낸다는 것을 고려할 때, 모든 그의 다양한 일시적, 발생적 그리고 조직 특이적 징후에서 유전자 발현을 제어하도록 상호작용하여야 한다는 것이 분명하게 된다. 유전자의 발현은 DNA 요소와 상호작용하는 일반적 그리고 특이적인 전사 조절자의 고도로 복잡한 혼합물에 의해 제어된다. 이들 DNA 요소는 국소 DNA 요소, 예컨대 코어 프로모터와 그의 관련된 전사 인자 결합 부위뿐만 아니라 원위 요소, 예컨대 인핸서, 사일런서, 절연체 및 좌위 제어 영역(locus control region: LCR)을 포함한다(문헌[Matson *et al.* (2006) *Ann Rev Genome Hum Genet* 7: 29-50] 참조).

[0006]

인핸서 요소는 SV40 바이러스 게놈에서 처음 등장되었고, 이어서, 인간 면역글로불린 중쇄 좌위에서 발견되었다. 현재 다수의 유전자의 발현에서 조절 역할을 하는 것으로 알려진 인핸서는 유전자 발현의 일시적 그리고 공간적 패턴에 주로 영향을 미치는 것으로 나타난다. 또한 인핸서는 표적화된 유전자의 코어 프로모터로부터 면 거리에서 발현을 조절하는 기능을 할 수 있으며, 프로모터에 대해 임의의 특정 서열 배향에 의존하지 않는다는 것이 발견되었다. 인핸서는 그들이 인트론 서열에 또는 심지어 유전자의 3' 말단을 지나서 위치될 수 있는 코어 프로모터 영역의 상류 또는 하류의 수백 킬로베이스에 위치될 수 있다.

[0007]

게놈 DNA의 표적화된 절단을 위한 다양한 방법 및 조성물이 기재되었다. 이러한 표적화된 절단 사건은, 예를 들어, 표적화된 돌연변이유발을 유도하기 위해 사용될 수 있고, 세포 DNA 서열의 표적화된 결실을 유도하며, 그리고 사전결정된 염색체 좌위에서 표적화된 재조합을 용이하게 한다. 예를 들어, 미국 특허 제9,255,250호; 제9,200,266호; 제9,045,763호; 제9,005,973호; 제9,150,847호; 제8,956,828호; 제8,945,868호; 제8,703,489호; 제8,586,526호; 제6,534,261호; 제6,599,692호; 제6,503,717호; 제6,689,558호; 제7,067,317호; 제7,262,054호; 제7,888,121호; 제7,972,854호; 제7,914,796호; 제7,951,925호; 제8,110,379호; 제8,409,861호; 미국 특허 공개 제2003/0232410호; 제2005/0208489호; 제2005/0026157호; 제2005/0064474호; 제2006/0063231호; 제2008/0159996호; 제2010/00218264호; 제2012/0017290호; 제2011/0265198호; 제2013/0137104호; 제2013/0122591호; 제2013/0177983호; 제2013/0196373호; 제2015/0056705호 및 제2015/0335708호를 참조하며, 이의 개시내용은 그들의 전문이 참고로 포함된다.

[0008]

오류 유발 과정, 예컨대 비상동성 말단 결합(non-homologous end joining: NHEJ)에 의한 파손의 수선, 공여자의 비상동성 지향 말단 또는 수선 주형을 이용하는 수선(상동성 지향 수선 또는 HDR)이 관심 대상(표적화된 통합)의 유전자 또는 삽입의 넉아웃을 초래할 수 있도록 이들 방법은 종종 이중 가닥 파손(double strand break: DSB) 또는 표적 DNA 서열에서의 틈을 유도하도록 조작된 절단 시스템의 사용을 수반한다. 예를 들어, 미국 특허 제9,045,763호; 제9,200,266호; 제9,005,973호; 및 제8,703,489호 참조. 이들 기법은 또한 게놈 영역의 특정 결실의 도입, 또는 특정 점돌연변이 또는 국소화된 변경(또한 유전자 보정으로서 알려짐)의 도입을 포함하는, 공여자 올리고뉴클레오타이드의 사용을 통해 게놈 서열에서 부위 특이적 변화를 도입하기 위해 사용될 수 있다. 절단은 특정 뉴클레아제, 예컨대 조작된 아연 핑거 뉴클레아제(ZFN), 전사-활성체 유사 효과기 뉴클레아제(TALEN)의 사용을 통해, 또는 특정 절단을 가이드하기 위해 조작된 crRNA/tracr RNA('단일 가이드 RNA')를 이용

하는 CRISPR/Cas 시스템을 사용하여 일어날 수 있다. 추가로, 표적화된 뉴클레아제는 게놈 편집 및 유전자 요법에서 사용하기 위한 잠재력을 또한 가질 수 있는 아거노트(Argonaute) 시스템(예를 들어, 'TtAgo'로서 알려진 써무스 써모필루스(*T. thermophilus*)로부터 유래, 문헌[Swarts *et al.* (2014) *Nature* 507(7491): 258-261] 참조)에 기반하여 개발 중에 있다.

[0009] 적혈구 세포(Red blood cell: RBC) 또는 적혈구는 혈액의 주된 세포 성분이며, 인간에서 세포의 1/4을 차지한다. 성숙 RBC는 핵 및 다수의 다른 세포소기관을 결여하며, 폐로부터 조직까지 산소를 전달하는 기능을 할 뿐만 아니라 이산화탄소를 조직 밖으로 그리고 제거를 위해 폐로 다시 운반하는 금속단백질인 헤모글로빈으로 차 있다. 이 단백질은 RBC 건조 중량의 대략 97%를 구성하며, 혈액의 산소 운반 능력을 약 70배까지 증가시킨다. 헤모글로빈은 2개의 알파(α)-유사 글로빈 쇄 및 2개의 베타(β)-유사 글로빈 쇄 및 4개의 헵 기를 포함하는 이형사량체이다. 성인에서, α₂β₂ 사량체는 헤모글로빈 A(HbA) 또는 성인 헤모글로빈으로서 지칭된다. 전형적으로, 알파 및 베타 글로빈 쇄는 대략 1:1 비로 합성되며, 이 비는 헤모글로빈 및 RBC 안정화에 관해 중요한 것으로 여겨진다. 발생 중인 태아에서, 산소가 모체의 혈액계를 통해 아기의 혈액계로 전달될 수 있도록 헤모글로빈 A보다 산소에 대해 더 큰 결합 친화도를 갖는 상이한 형태의 헤모글로빈인 태아 헤모글로빈(HbF)이 생산된다. 서열이 매우 유사한 태아 글로빈을 암호화하는 2종의 유전자가 있으며, 이는 베타 글로빈 유전자 좌위에서의 그들의 배열 순서에 기반하여 HBG1(또한 G감마로서 지칭됨) 및 HBG2(A감마)로 지칭된다. 성인 헤모글로빈과 같이, 태아 헤모글로빈 단백질은 2개의 α 글로빈 쇄를 함유하지만, 성인 β-글로빈 쇄 대신에, 이는 2개의 태아 감마(γ)-글로빈 쇄(즉, 태아 헤모글로빈은 α₂γ₂임)를 가진다. 대략 30주의 임신에서, 태아에서의 감마 글로빈의 합성은 하락하기 시작하는 반면, 베타 글로빈의 생산은 증가한다. 대략 10개월령까지, 일부 HbF는 성인기까지 지속되지만(총 헤모글로빈의 대략 1 내지 3%), 신생아의 헤모글로빈은 거의 모두 α₂β₂이다. 감마-내지 베타-글로빈의 생산으로부터 스위치의 조절은 상당히 복잡하며, 베타 글로빈 전사의 동시 상향 조절과 함께 감마 글로빈 전사의 하향 조절을 주로 수반한다.

[0010] 헤모글로빈 쇄를 암호화하는 서열에서의 유전자 결합은 겸상 적혈구 빈혈 및 알파 및 베타 탈라세미아를 포함하는 이상혈색소증으로서 알려진 질환의 그룹을 초래할 수 있다. 이상혈색소증을 갖는 대다수의 환자에서, 감마 글로빈을 암호화하는 유전자가 존재하는 채로 남아있지만, 발현은 상기 기재한 바와 같이 출산 주위에 생기는 정상 유전자 퇴행에 기인하여 상당히 낮다.

[0011] 미국에서 사하라 이남 아프리카 혈통 사람의 대부분에서 5000명의 사람 중 1명이 겸상 적혈구병(SCD)을 갖는 것으로 추정된다(Roseff (2009) *Immunohematology* 25(2):67) 말라리아에 대한 보호에 기인하는 겸상적혈구 성숙의 이형접합 보인자에 대한 이점이 되는 것으로 나타났으며, 따라서 이 형질은 사하라 이남 아프리카에서 집단의 28%까지 겸상 세포 형질을 갖는 것으로 추정되도록, 시간에 따라 양성으로 선택될 수 있다(Elguero *et al.* (2015) *PNAS USA* 112 (22): 7051). 아미노산 #6에서 글루탐산을 발린으로 치환한 결과로서 β글로빈 유전자에서의 돌연변이(DNA 수준에서 GAG의 GTG로의 돌연변이)에 의해 겸상세포병이 야기되며, 얻어진 헤모글로빈은 "헤모글로빈 S" 또는 "HbS"로서 지칭된다. 더 낮은 산소 조건 하에서, HbS의 데옥시 형태에서 입체배좌 이동은 E와 F 나선 사이의 단백질 상의 소수성 패치를 노출시킨다. 헤모글로빈 내 베타쇄의 위치 6에서 발린의 소수성 잔기는 소수성 패치와 결합되어, HbS 분자가 응집하고 섬유성 침전물을 형성하도록 야기할 수 있다. 이들 응집물을 결국 RBC의 이상 또는 '겸상세포생성'을 야기하여, 세포 유연성의 상실을 초래한다. 겸상세포생성 RBC는 더 이상 모세혈관상에 끼어들 수가 없으며, 겸상 세포 환자에서 혈관막힘위기를 초래할 수 있다. 추가로, 겸상세포 RBC는 정상 RBC보다 더 약하며, 용혈되어, 결국 환자에서 빈혈을 야기하는 경향이 있다.

[0012] 겸상 세포 환자의 치료 및 관리는 급성 에피소드 동안 항생제 치료, 통증 관리 및 수혈을 수반하는 평생 동안의 문제이다. 일 접근은 부분적으로 감마 글로빈의 생산을 증가시킴으로써 그의 효과를 별휘하는 하이드록시유레아의 사용이다. 만성 하이드록시유레아 요법의 장기간 부작용은 여전히 알려져 있지 않지만, 그러나, 치료는 낮은 환자 순응도를 야기하는 원치않는 부작용을 제공하며, 환자에 따라 변할 수 있는 효능을 가진다(Brandow and Panepinto (2011) *Am J Hematol* 86(9):804-806). 겸상 세포 치료 효능의 증가에도 불구하고, 환자의 수명 예상은 여전히 단지 50대 중반 내지 후반이며, 질환의 관련된 사망률은 환자의 삶의 질에 엄청난 영향을 가진다.

[0013] 탈라세미아는 또한 헤모글로빈에 관한 질환이며, 전형적으로 글로빈 쇄의 감소된 발현을 수반한다. 이는 유전자의 조절 영역에서의 돌연변이를 통해 또는 감소된 발현 또는 감소된 수준 또는 기능성 글로빈 단백질을 초래하는 글로빈 암호 서열에서의 돌연변이로부터 일어날 수 있다. 알파 글로빈 좌위에서 돌연변이에 의해 야기되는 알파 탈라세미아는 서아프리카 및 남아시아 혈통의 사람과 주로 관련되며, 말라리아 내성을 부여할 수 있다. 베타 글로빈 좌위에서의 돌연변이에 의해 야기되는 베타 탈라세미아는 전형적으로 그리스 및 터키와 이탈리아의 해안 지역의 지중해 혈통 사람과 주로 관련된다. 소탈라세미아에서, β 글로빈 대립유전자 중 하나만이 돌연변

이를 보유한다. 개체는 소구성 빈혈을 앓고 있으며, 검출은 보통 정상 평균 적혈구 용적보다 더 낮게 수반된다 (<80fL). 소탈라세미아가 있는 대상체의 대립유전자는 β^+/β 또는 $\beta 0/\beta$ 이다(여기서 ' β^+ '는 소량의 β 쇄 형성이 일어나도록 허용하는 대립유전자를 지칭하며, ' β '는 야생형 β 글로빈 대립유전자를 지칭하고, ' $\beta 0$ '는 베타-글로빈 발현의 완전한 부재와 관련된 β 글로빈 돌연변이를 지칭한다). 탈라세미아 중간 대상체는 종종 정상적인 삶을 영위할 수 있지만, 그들의 빈혈의 중증도에 따라서 특히 질병 또는 임신 시 가끔 수혈이 필요할 수 있다. 이들 환자의 대립유전자는 β^+/β^+ 또는 $\beta 0/\beta^+$ 일 수 있다. 대탈라세미아는 대립유전자가 둘 다 탈라세미아 돌연변이를 가질 때 일어난다. 이는 심하게 작은 적혈구이며 저색소성 빈혈이다. 치료되지 않는다면, 이는 빈혈, 비장증대 및 중증의 골 기형을 야기하며, 20세 전에 사망으로 진행된다. 치료는 주기적 수혈; 비장증대에 대한 비장절제술 및 수혈로 야기된 철 과부하의 퀼레이트화로 이루어진다. 적절한 공여자가 동정될 수 있지만, 이 절차가 상당한 위험을 가질 수 있다면 골수 이식이 또한 중증의 탈라세미아의 치료를 위해 사용된다.

[0014]

SCD와 베타 탈라세미아 둘 다의 치료를 위해 제안된 일 접근은 비정상 성인 헤모글로빈을 기능적으로 대체하는 HbF를 갖기 위한 목적으로 감마 글로빈의 발현을 증가시키는 것이다. 상기 언급한 바와 같이, 하이드록시유레아를 이용하는 SCD 환자의 치료는 감마 글로빈 발현을 증가시킴에 있어서 그의 효과에 기인하여 부분적으로 성공적이 되는 것으로 생각된다. 감마 글로빈 발현에 영향을 미치는 것으로 발견된 화합물의 제1 그룹은 세포독성 약물이다. 약학적 조작에 의해 감마-글로빈의 드노보 합성을 야기하는 능력은 실험 동물에서 5-아자사이티딘을 이용하여 처음 나타났다(DeSimone (1982) *Proc Nat'l Acad Sci USA* 79(14):4428-31). 후속 연구는 β -탈라세미아 및 겹상 적혈구병을 갖는 환자에서 HbF를 증가시키는 5-아자사이티딘의 능력을 확인하였다(Ley, et al., (1982) *N. Engl. J. Medicine*, 307: 1469-1475, 및 Ley, et al., (1983) *Blood* 62: 370-380). 추가로, 단쇄지방산(예를 들어, 뷰티레이트 및 유도체)은 HbF를 증가시키기 위한 실험 시스템에서 나타났다(Constantoulakis et al., (1988) *Blood* 72(6):1961-1967). 또한, '태아 헤모글로빈의 유전적 지속성'(HPFH)으로서 알려진 병태를 갖는 인간 집단의 세그먼트가 있으며, 여기서, 상승된 양의 HbF가 성인기에 지속된다(HPFH 이형접합체에서 10 내지 40%(문헌[Thein et al. (2009) *Hum. Mol. Genet* 18 (R2): R216-R223] 참조). 이는 희귀한 병태이지만, 임의의 관련된 베타 글로빈 이상의 부재 하에서, 개체의 헤모글로빈의 100%가 HbF일 때조차 임의의 상당한 임상 징후와 관련되지 않는다. 베타 탈라세미아를 갖는 개체가 또한 일치되는 HPFH를 가질 때, HbF의 발현은 질환의 중증도를 줄일 수 있다(Potoka and Gladwin (2015) *Am J Physiol Lung Cell Mol Physiol*. 308(4): L314-L324). 추가로, 겹상 적혈구병의 천연 과정의 중증도는 환자에 따라 상당히 다를 수 있으며, 이 가변성은 더 경증의 질환을 갖는 일부 개체가 더 고수준의 HbF를 발현시킨다는 사실에 대해 부분적으로 추적될 수 있다.

[0015]

HbF의 발현을 증가시키기 위한 일 접근은 유전자의 동정을 수반하는데, 이의 산물은 감마 글로빈 발현의 조절에서 어떤 역할을 한다. 하나의 이러한 유전자는 램프구 발생에서 그의 역할 때문에 처음 동정된 BCL11A이다. BCL11A는 감마 글로빈 발현의 발생 단계-특이적 조절에 수반되는 것으로 생각되는 아연 평거 단백질을 암호화한다. BCL11A는 성인 적혈구 전구체 세포에서 발현되며, 그의 발현의 하향 조절은 감마 글로빈 발현의 증가를 야기한다. 추가로, BCL11A mRNA의 스플라이싱은 발생적으로 조절된다. 배아 세포에서, BCL11A-S 및 BCL11A-XS로서 알려진 더 짧은 BCL11A mRNA 변이체가 주로 발현되는 반면, 성인 세포에서, 더 긴 BCL11A-L 및 BCL11A-XL mRNA 변이체가 우세하게 발현된다. 문헌[Sankaran et al. (2008) *Science* 322 p. 1839] 참조. BCL11A 단백질은 베타 글로빈 좌위와 상호작용하여 그의 임체배좌를 변경시키고, 그에 따라 상이한 발생 단계에서 그의 발현을 변경시키는 것으로 나타난다. BCL11A 유전자에 표적화된 저해 RNA의 사용이 제안되었지만(예를 들어, 미국 특허 공개 제2011/0182867호) 이 기술은 수명 동안 다크 치료를 필요로 하는 몇몇 잠재적 결함(즉, 완전한 넉다운이 달성되지 않을 수도 있으며, 이러한 RNA의 전달은 문제가 될 수 있고, RNA가 지속적으로 제공되어야 한다)을 가진다.

[0016]

BCL11A 인핸서 서열의 표적화는 HbF를 증가시키기 위한 메커니즘을 제공한다. 예를 들어, 미국 특허 제2015/0132269호 및 국제 특허 출원 공개 WO 2016/183298 참조. 게놈 와이드 관련 연구는 증가된 HbF 수준과 관련된 BCL11A 유전자 좌위에서 유전자 변이의 세트를 동정하였다. 이들 변이는 단계-특이적, 계통-제한적 인핸서 영역으로서 작용하는 BCL11A의 비암호 영역에서 발견되는 작은 뉴클레오타이드 다형체(small nucleotide polymorphism: SNP)의 수집물이다. 추가적인 연구는 이 BCL11A 인핸서가 BCL11A 발현을 위해 적혈구 세포에서 필요하지만, B 세포에서의 그의 발현에 대해서는 필요하지 않다는 것을 나타내었다(문헌[Bauer et al. (2013) *Science* 343:253-257] 참조). 인핸서 영역은 BCL11A 유전자의 인트론 2 내에서 발견되었으며, 인트론 2에서 DNaseI 과민성의 3개 영역(종종 조절 잠재력과 관련된 염색질 상태를 나타냄)을 동정하였다. 이들 3개 영역은 BCL11A의 전사 개시 부위로부터의 거리(킬로베이스)에 따라 "+62", "+58" 및 "+55"로서 동정되었다. 이들 인핸서 영역은 길이가 대략 350(+55); 550(+58); 및 350(+62)개의 뉴클레오타이드이다(Bauer 2013, 이하 참조).

[0017] 인간의 치료적 처치에서 사용하기 위한 뉴클레아제를 개발할 때, 뉴클레아제가 최고의 안전성 특징을 갖는 것은 필수적이다. 구체적으로, 뉴클레아제는 매우 저수준의 표적을 벗어난 절단을 가져야 한다. 사용자-특정된 표적 이외의 위치에서 상당한 수의 이중 가닥 절단은 표적을 벗어난 유전자의 억제를 야기할 수 있으며, 드문 예에서, 염색체 전위의 발생을 야기할 수 있다(문헌[Hoban and Bauer (2016) *Blood*, 127(21):2525-2535 및 Tsang *et al.* (2017) *Nature Methods*, in press] 참조). 특이성의 개선은 조작된 뉴클레아제와 게놈 DNA 사이의 비특이적 상호작용을 제거함으로써 달성될 수 있다(미국 가출원 특허 제62/378,978호 및 제62/443,981호).

[0018] 따라서, 예를 들어 이상혈색소증, 예컨대 겹상 적혈구병 및 베타 탈라세미아를 치료하기 위해 BCL11A 유전자 발현의 변경을 위한 추가적인 고도로 특이적인 방법 및 조성물에 대한 필요가 남아있다.

발명의 내용

[0019] 본 발명은 유전자 요법 및 게놈 조작에서 사용하기 위한 고도로 특이적인 조성물 및 방법을 기재한다. 구체적으로, 기재된 방법 및 조성물은 BCL11A 유전자, 예를 들어, 1종 이상의 추가적인 유전자의 조절자로서 작용하는 유전자를 (예를 들어, 그의 발현을 완전히 또는 부분적으로 없앰으로써) 비활성화시키는 것에 관한 것이다. 특히, 본 발명은 특정 세포 계통(예를 들어, 적혈구)에서 그의 활성을 감소시키거나 또는 낙아웃시키기 위해 BCL11A 유전자에서 인핸서 기능을 방해하는 방법 및 조성물을 기재한다. 추가적으로, 본 발명은 BCL11A 인핸서 기능을 방해하는 방법 및 조성물을 제공하되, 인핸서 서열은 BCL11A 유전자의 암호 서열 내에 위치되지 않으며, 제공되는 시약은 고도로 특이적인 활성을 제공하였다. 이들 상황에서 BCL11A 유전자의 결과로 인한 하향 조절은 감마 글로빈의 증가된 발현을 초래하며, 표적을 벗어난 절단 사건의 수는 감소된다.

[0020] 일부 양상에서, 본 발명은 4, 5 또는 6개의 평거를 포함하는 아연 평거 단백질(ZFP)을 포함하는 비천연 유래 아연 평거 단백질을 포함하며, 각각의 평거는 DNA 표적 하위 부위(subsite)를 인식하는 인식 나선 영역을 포함하되, 인식 나선 영역은 표 1의 단일 행에 나타낸 순서로 서열을 포함한다. 각각의 아연 평거 내에서, 7개의 아미노산 인식 나선 영역은 (아연 배위 잔기를 포함하는, 대략 30개의 잔기의) 아연 평거 골격 내에서 -1 내지 +6으로 넘버링된다. 특정 실시형태에서, 본 명세서에 기재된 아연 평거 단백질의 성분 아연 평거 중 1, 2, 3개 이상은 위치 -5, 위치 -14 또는 위치 -5와 -14 둘 다에서(인식 나선 영역에 대해 사용되는 -1 내지 +6 넘버링의 계속되는 넘버링) 아미노산에 대한 돌연변이를 포함하지만, 이들로 제한되지 않는 인식 나선 영역 밖의 하나 이상의 잔기에 대한 돌연변이를 추가로 포함한다. 예를 들어, 미국 가출원 특허 제62/378,978호 및 제62/443,981호에 기재된 Qm4 및 Qm14 돌연변이 참조. 아연 평거 단백질의 성분 아연 평거는, 예를 들어 미국 특허 제8,772,453호에 기재된 임의의 링커에 의해 연결될 수 있다. 특정 실시형태에서, ZFP는 다음과 같이 표기되는 단백질에 대해 표 1로 나타내는 인식 나선을 포함한다: 63014(서열번호 1에 나타내는 표적 부위에 결합함) 및 65722(서열번호 2에 나타낸 표적 부위에 결합함).

[0021] 특정 실시형태에서, 본 명세서에 기재된 바와 같은 아연 평거 단백질은 기능성 도메인(예를 들어, 전사 활성화 도메인, 전사 억제 도메인, 절단 도메인(아연 평거 뉴클레아제를 형성함) 등)에 융합된다. 임의의 링커는 미국 특허 제 9,394,531호 및 제9,567,609호에 기재된 바와 같은 링커를 포함하지만, 이들로 제한되지 않는 절단 도메인 및 아연 평거 단백질에 작동 가능하게 연결하기 위해 사용될 수 있다. 더 나아가, FokI 절단 도메인이 사용될 때, 야생형에 대해 넘버링된 이량체화 도메인에 대한 ELD 또는 KKR 돌연변이, FokI 도메인의 잔기 525에 대한 돌연변이(K에서 S로), 및 이량체화 도메인에 대한 ELD 또는 KKR 돌연변이와 FokI 도메인의 잔기 525에 대한 돌연변이(K에서 S로)의 조합을 포함하지만, 이들로 제한되지 않는, 촉매적 도메인, 이량체화 도메인에서의 추가적인 돌연변이, 인산염 접촉 잔기에 대한(이량체화 또는 촉매적 도메인에서는 아님) 추가적인 돌연변이 및 촉매적 도메인, 이량체화 도메인 중 임의의 하나에서의 돌연변이와 인산염 접촉 잔기에 대한 돌연변이 조합이 존재할 수 있다. 미국 특허 제7,888,121호; 제7,914,796호; 제8,034,598호; 제8,623,618호 및 미국 특허 공개 제2011/0201055호 및 미국 가출원 특허 제62/378,978호 및 제62/443,981호 참조.

[0022] 특정 실시형태에서, 아연 평거 뉴클레아제(ZFN)는 쌍의 ZFN에 대해 표적 부위 중 하나에서 또는 근처에서 또는 둘 다에서 절단하도록 이량체화 쌍에서 사용될 수 있으며, 예를 들어, 표 1의 "좌측 상대"(예를 들어, 63014)는 BCL11A 인핸서 서열을 절단하기 위해 표 1의 "우측 상대"(예를 들어, 65722)와 이량체를 형성할 수 있다. 특정 실시형태에서, ZFN의 쌍은 다음의 아미노산 서열을 포함한다:

63014:

MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGIHGVPAAMAERPFQCRI
 CMQNFSQSNLRAHIRHTHTGEKPFACDICGRKFARNFSLTMHTKIHTGSQKPF
 QCRCIMQNFSSTGNLTNHIRHTHTGEKPFACDICGRKFATSGSLTRHTKIHTHPR
 APIPKPFQCRCIMQNFSQSNLRAHIRHTHTGEKPFACDICGRKFAAQCCLFHHT
 KIH- 링 커 -

ELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGK
 HLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMERYVEENQ
 TRDKHLPNEWWKVYPSSVTEFKFLVSGHFKGNYKAQLTRLNHTNCNGA
 VLSVEELLIGGEMIKAGTLTLEEVRKFNNGEINFRS (서열 번호 3); 및
 MDYKDHDGDYKDHDIDYKDDDDKMAPKKRKVGIHGVPAAMAERPFQCRI
 CMQKFARNDHRTTHKIHTGEKPFQCRCIMQNFSQKAHLIRHIRTHTGEKPF
 CDICGRKFAQKGTLEHTKIHTGSQKPFQCRICMQNFSRGRDLSRHIRTHTGE
 KPFACDICGRKFARRDNLHSHTKIH- 링 커 -

ELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGK
 HLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRVYVKENQ
 TRNKHINPNEWWKVYPSSVTEFKFLVSGHFSGNYKAQLTRLNRKTNCNGA
 VLSVEELLIGGEMIKAGTLTLEEVRKFNNGEINF (서열 번호 4),

[0023]

여기서, 링 커 서열은, 예를 들어 미국 특허 제9,394,531호 및 제9,567,609호에 기재된 바와 같은 당업계에 공지된 임의의 링 커 서열일 수 있다. 특정 실시형태에서, 63014에 대한 링 커는 L7c5 링 커(LRGSQLVKS(서열번호 5))를 포함하거나 또는 이것으로 이루어지고, 65722에서 사용되는 링 커는 L0 링 커(LRGSQLVKS(서열번호 6), 미국 특허 제9,567,609호)를 포함하거나 또는 이것으로 이루어진다. 상기 나타낸 서열 링 커에 대해 C-말단의 FokI 절단 도메인 서열은 또한 아연 평거 단백질에 작동 가능하게 연결된 대안의 FokI 도메인을 포함할 수 있다. 특정 실시형태에서, FokI 절단 도메인은 촉매 도메인, 이량체화 도메인, 인산염 접촉 잔기에 대한 대안의 또는 추가적인 돌연변이, 및 촉매 도메인, 이량체화 도메인과 인산염 접촉 잔기 중 임의의 하나에 대한 돌연변이의 조합을 포함할 수 있다.

[0024]

다른 양상에서, 본 발명은 계놈 조작의 목적을 위해 인간 줄기 세포 또는 전구체 세포(HSC/PC)에 대한 적어도 하나의 뉴클레아제(예를 들어, BCL11A 인핸서 서열에 결합하는 뉴클레아제)의 전달을 포함한다. 특정 실시형태에서, 뉴클레아제는 4, 5 또는 6개의 평거를 포함하는 아연 평거 단백질(ZFP)을 포함하며, 각각의 평거는 표적 하위 부위를 인식하는 인식 나선 영역을 포함하되, 인식 나선 영역은 표 1의 단일행에 나타낸 순서로 서열을 포함한다. 다른 실시형태에서, ZFN 뉴클레아제는 63014/65722로 표기되는 뉴클레아제의 쌍을 포함한다. 본 명세서에서 기재되는 뉴클레아제(들)는 링 커(예를 들어, DNA-결합 도메인과 절단 도메인 사이), 예를 들어 LRGSQLVKS(서열번호 5) 또는 LRGSQLVKS(서열번호 6))를 포함하지만, 이들로 제한되지 않는 제9,567,609호에 나타내는 바와 같은 링 커를 추가로 포함할 수 있다.

[0025]

일부 실시형태에서, 뉴클레아제는 웨타이드로서 전달되는 반면, 나머지에서 이는 적어도 하나의 뉴클레아제를 암호화하는 핵산으로서 전달된다. 일부 실시형태에서, 하나 초파의 뉴클레아제가 사용된다. 일부 바람직한 실시형태에서, 뉴클레아제를 암호화하는 핵산은 mRNA이고, 일부 실시형태에서, mRNA는 보호된다. 일부 양상에서, mRNA는 화학적으로 변형될 수 있다(예를 들어, 문헌[Kormann *et al.* (2011) *Nature Biotechnology* 29(2):154-157] 참조). 다른 양상에서, mRNA는 ARCA 캡을 포함할 수 있다(미국 특허 제7,074,596호 및 제8,153,773호). 추가 실시형태에서, mRNA는 비변형 뉴클레오타이드와 변형 뉴클레오타이드의 혼합물을 포함할 수 있다(미국 특허 공개 제2012/0195936호 참조). 바람직한 실시형태에서, 뉴클레아제(들)를 암호화하는 핵산은 전기천공법을 통해 HSC/PC에 전달된다. 일부 실시형태에서, 뉴클레아제는 전사 인자의 결합 부위에서 또는 근처에서 절단한다. 일부 양상에서, 전사 인자는 GATA-1이다.

[0027]

다른 양상에서, 본 발명은, 예를 들어 세포의 야생형 서열에 비교하여 내인성 BCL11A 인핸서 서열이 본 명세서

에 기재된 바와 같은 뉴클레아제(예를 들어, 표 1에 나타냄)에 의해 유전자 변형되는 세포 또는 세포주를 포함한다. BCL11A 인핸서에 대한 유전자 변형은 글로빈(베타 및 감마) 유전자 발현의 변형을 초래한다. 본 명세서에 기재된 바와 같은 뉴클레아제 변형 세포 또는 세포주는 구조, 기능 및 구조와 기능 둘 다의 조합이 야생형과 구별 가능하다. 유전자 변형된 세포 또는 세포주는 변형에 대해 이형접합적 또는 동형접합적일 수 있다. 변형은 삽입(예를 들어, 이식유전자 삽입) 결실 및 삽입과 결실의 조합을 포함할 수 있으며; 이러한 삽입, 결실 및 삽입과 결실의 조합은 통상적으로 "삽입결실(indel)"로서 지칭된다. 일부 바람직한 실시형태에서, 삽입결실은 전사 인자 결합 부위의 파괴를 초래한다. 특정 실시형태에서, 변형은 뉴클레아제(들) 결합 부위(들), 절단 부위(들), 및 결합 부위와 절단 부위의 조합에서이거나 또는 근처에서이고, 예를 들어, 절단 부위(들)의 상류 또는 하류의 1 내지 300개(또는 그 사이의 임의의 값) 이내, 더 바람직하게는 표 1에 나타내는 결합 부위(들), 절단 부위(들), 및 결합 부위와 절단 부위의 조합의 측면 중 하나의 1 내지 100개의 염기쌍(또는 그 사이의 임의의 값) 내에서, 훨씬 더 바람직하게는 결합 부위(들), 절단 부위(들), 및 결합 부위와 절단 부위의 조합의 측면 중 하나에 대한 1 내지 50개 염기쌍(또는 그 사이의 임의의 값) 이내이다. 특정 실시형태에서, BCL11A 인핸서 서열의 유전자 변형은 표 1에 나타내는 서열(표적 부위) 내이고/이거나 이들 사이에 있다. 변형은 또한 절단 부위 내 하나 이상의 뉴클레오타이드에 대한 변형을 포함할 수 있다. 변형은 또한 결합 부위 내 하나 이상의 뉴클레오타이드에 대한 변형을 포함할 수 있다. 변형은 절단 부위에서, 그리고 결합 부위 중 하나 이상에서 하나 이상의 뉴클레오타이드에 대한 변형을 추가로 포함할 수 있다. 특정 실시형태에서, 뉴클레아제 표적 부위(들) 중 하나 이상은 변형되지 않는다. 다른 실시형태에서, 뉴클레아제(들)에 대한 표적 부위 중 적어도 하나는 변형된다. 특정 실시형태에서, 변형은 BCL11A 인핸서의 "+58" 영역에 또는 근처에, 예를 들어, 임의의 서열번호 1 및 서열번호 2에 나타낸 뉴클레아제 결합 부위에 또는 근처에 있다. 임의의 세포 또는 세포주는 본 명세서에 기재된 바와 같은 뉴클레아제, 예를 들어 줄기 세포(조혈 줄기 세포, 예컨대 CD34+ 조혈 줄기 세포) 또는 적혈구(RBC) 전구체 세포에 의해 변형될 수 있다.

[0028] 또한 본 명세서에 기재된 바와 같은 뉴클레아제에 의한 변형 후에 얻어지는 세포 또는 세포주, 예를 들어 본 명세서에 기재되는 바와 같은 뉴클레아제-변형 세포 또는 세포주로부터 유래된 세포 또는 세포주가 기재된다. 본 명세서에 기재된 바와 같은 변형된 줄기 세포로부터 유래된 부분적으로 또는 완전히 분화된 세포가 또한 제공된다(예를 들어, RBC 또는 RBC 전구체 세포). 뉴클레아제-변형 세포로부터 유래된 세포는 시험관내에서 (배양물) 증식되고, 분화되며, 증식과 분화 둘 다의 조합일 수 있거나 또는, 예를 들어 뉴클레아제-변형 줄기 세포의 시험관외 투여 후에 살아있는 대상체 내에서 분화될 수 있다. 본 명세서에 개시된 임의의 유전자 변형된 세포 또는 세포주는 감마 글로빈의 증가된 발현을 나타낼 수 있다. 본 명세서에 기재된 바와 같은 유전자 변형된 세포를 포함하는 약제학적 조성물과 같은 조성물이 또한 제공된다.

[0029] 다른 양상에서, 본 발명은 유전자 변형된 세포를 제공하기 위해 표적 세포에 공여자 핵산의 전달을 포함하며, 이때 공여자는 세포 내로 통합된다. 공여자는 표 1의 뉴클레아제(들)를 암호화하는 핵산 전에, 후에 또는 함께 전달될 수 있다. 공여자 핵산은 세포의 게놈, 예를 들어, 내인성 좌위 내로 통합될 외인성 서열(이식유전자)을 포함할 수 있다. 일부 실시형태에서, 공여자는 표적화된 절단 부위와의 상동성 영역에 측접되는 전장 유전자 또는 이의 단편을 포함할 수 있다. 일부 실시형태에서, 공여자는 상동성 영역을 결여하며, 상동성 독립적 메커니즘(즉, NHEJ)을 통해 표적 좌위 내로 통합된다. 공여자는 임의의 핵산 서열, 예를 들어 뉴클레아제-유도 이중가닥 파손의 상동성-지형 수선을 위한 기질로서 사용될 때, 내인성 염색체 좌위(예를 들어, BCL11A 인핸서 영역)에서 생성될 공여자-구체화된 결실 또는 대안적으로(또는 추가로), 생성될 내인성 좌위의 신규한 대립유전자 형태(예를 들어, 전사 인자 결합 부위가 없는 점 돌연변이)를 야기하는 핵산을 포함할 수 있다. 일부 양상에서, 공여자 핵산은 올리고뉴클레오타이드이되, 통합은 유전자 보정 사건 또는 표적화된 결실을 야기한다.

[0030] 다른 양상에서, 뉴클레아제, 공여자 및 뉴클레아제와 공여자 둘 다의 조합은 바이러스, 비바이러스 및 바이러스와 비바이러스 유전자 전달 방법의 조합에 의해 전달된다. 바람직한 실시형태에서, 공여자는 아데노-연관 바이러스(adeno-associated virus: AAV)를 통해 세포에 전달된다. 일부 예에서, AAV는 캡시드 혈청형과의 비교에서 이종성 혈청형을 갖는 LTR을 포함한다.

[0031] 일부 양상에서, 인핸서의 DNaseI 과민감 영역(예를 들어, BCL11A 인핸서의 +58 영역) 내의 영역을 포함하는 결실은 표 1에 나타내는 바와 같은 하나 이상의 뉴클레아제를 이용하여 이루어진다. 이들 결실은 약 1개의 뉴클레오타이드 내지 약 551개의 뉴클레오타이드를 포함할 수 있다. 따라서, 결실은 1, 5, 10, 15, 20, 25, 30, 40, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550개의 뉴클레오타이드 또는 그 사이의 임의의 값을 포함할 수 있다. 일부 실시형태에서, 결실은 하나 이상의 전사 인자에 대한 결합 영역을 포함한다. 일부 바람직한 실시형태에서, 결실은 GATA-1 결합 부위, 또는 다른 인자와 조합한 GATA-1에 대한 결합 부위를 포함한다.

- [0032] 일부 실시형태에서, 표 1의 DNA 결합 도메인은 기능성 도메인에 융합된다. 일부 양상은 유전자 발현을 조절할 수 있는 도메인과 DNA 결합 도메인의 융합을 포함한다. 일부 실시형태에서, 융합 단백질은 유전자 발현 조절 도메인에 융합된 표 1의 DNA 결합 도메인을 포함하며, 여기서 조절자는 유전자 발현을 억제한다.
- [0033] 일부 실시형태에서, HSC/PC 세포는 본 발명의 뉴클레아제, DNA 결합 단백질, 및 뉴클레아제와 DNA 결합 단백질의 조합(즉, 표 1에 나타낸 바와 같은 ZFP)과 접촉된다. 일부 실시형태에서, 뉴클레아제, DNA 결합 단백질, 및 뉴클레아제와 DNA 결합 단백질의 조합은 핵산으로서 전달되고, 다른 실시형태에서, 그들은 단백질로서 전달된다. 일부 실시형태에서, 핵산은 DNA 결합 단백질, 및 뉴클레아제와 DNA 결합 단백질의 조합을 암호화하는 mRNA이며, 그리고 추가 실시형태에서, mRNA는 보호될 수 있다. 일부 실시형태에서, mRNA는 화학적으로 변형될 수 있고, ARCA 캡, 비변형과 변형 뉴클레오타이드의 혼합물, 및 ARCA 캡 및 비변형과 변형 뉴클레오타이드의 혼합물의 조합을 포함할 수 있다. 부분적으로 또는 완전히 분화된 세포를 포함하는, 이들 세포로부터 유래된 세포 또는 세포주가 또한 제공된다.
- [0034] 일부 양상에서, HSC/PC는 대상체로부터 HSC/PC의 분리반출법, 또는 채취된 골수로부터의 정제 후에 생체외에서 본 발명의 뉴클레아제, DNA 결합 단백질, 및 뉴클레아제와 DNA 결합 단백질의 조합과 접촉된다. 일부 실시형태에서, 본 명세서에 기재된 뉴클레아제는 BCL11A 인핸서 영역 내에서 변형을 야기하여, 예를 들어 야생형 세포, 다른 변형된(예를 들어, 뉴클레아제-변형된) 세포, 및 야생형과 다른 변형된 세포의 조합과 별개인 구조적으로, 기능적으로, 그리고 구조적과 기능적을 조합한 유전자 변형된 세포를 얻는다. 추가 실시형태에서, BCL11A 인핸서 영역 변형을 함유하는 HSC/PC는 대상체에게 다시 도입된다. 일부 예에서, BCL11A 인핸서 영역 변형을 함유하는 HSC/PC는 도입 전에 확장된다. 다른 양상에서, 유전자 변형된 HSC/PC는 골수 이식에서 대상체에게 제공되며, HSC/PC는 생체내에서 생착, 분화 및 성숙된다. 일부 실시형태에서, HSC/PC는 G-CSF-유도 동원, 플레릭사포르-유도 동원, 및 G-CSF-와 플레릭사포르-유도 동원의 조합 후에 대상체로부터 단리되고, 나머지에서, 세포는 인간 골수 또는 인간 탯줄로부터 단리된다. 일부 양상에서, 대상체는 변형된 HSC/PC를 포함하는 이식편의 도입 전에 약한 골수소멸성 절차로 치료되는 반면, 다른 양상에서, 대상체는 격렬한 골수소멸성 조건화 섭생으로 치료된다. 일부 실시형태에서, 본 발명의 방법 및 조성물은 이상혈색소증을 치료하거나 또는 예방하기 위해 사용된다. 일부 양상에서, 이상혈색소증은 탈라세미아이다. 일부 양상에서, 이상혈색소증은 베타 탈라세미아인 반면, 다른 양상에서, 이상혈색소증은 경상 적혈구병이다.
- [0035] 일부 실시형태에서, HSC/PC는 공여자 분자와 추가로 접촉된다. 일부 실시형태에서, 공여자 분자는 바이러스 벡터에 의해 전달된다. 공여자 분자는 기능성 폴리펩타이드(예를 들어, cDNA 또는 이의 단편)를 암호화하는 하나 이상의 서열을 프로모터와 함께 또는 프로모터 없이 포함할 수 있다. 추가적인 서열(암호 또는 비암호 서열)은, 공여자 분자가 2A 펩타이드, SA 부위, IRES 등을 암호화하는 서열을 포함하지만, 이들로 제한되지 않는 비활성화를 위해 사용될 때에 포함될 수 있다.
- [0036] 일부 양상에서, 본 발명의 방법 및 조성물은 생체내에서 HSC/PC와 접촉하는 방법을 포함한다. 뉴클레아제, DNA 결합 단백질, 또는 뉴클레아제와 DNA 결합 단백질의 조합은 당업계에 공지된 방법에 의해 인시추로 HSC/PC에 전달된다. 일부 실시형태에서, 본 발명의 뉴클레아제 및/또는 DNA 결합 단백질은 이것이 필요한 대상체에게 투여되는 바이러스 입자를 포함하는 반면, 다른 실시형태에서, 뉴클레아제, DNA 결합 단백질, 또는 뉴클레아제와 DNA 결합 단백질의 조합은 나노입자(예를 들어, 리포좀)를 포함한다. 일부 실시형태에서, 바이러스 입자, 나노입자, 또는 바이러스 입자와 나노입자의 조합은 기관(예를 들어, 골수)에 전달되어, HSC/PC가 남아있다.
- [0037] 다른 양상에서, 본 명세서에서 상동성-독립적 메커니즘을 통해 세포 계놈 내로 공여자 핵산을 통합하는 방법이 기재된다. 상기 방법은 공여자 핵산이 DSB의 부위에 통합되도록, 세포 계놈에서 이중 가닥 파손(DSB)을 생성하는 단계 및 본 명세서에 기재된 바와 같이 뉴클레아제를 이용하여 공여자 분자를 절단하는 단계를 포함한다. 특정 실시형태에서, 공여자 핵산은 비상동성 의존적 방법(예를 들어, NHEJ)을 통해 통합된다. 상기 언급한 바와 같이, 생체내 절단 시, 공여자 서열은 DSB의 위치에서 세포 계놈 내로 표적화된 방식으로 통합될 수 있다. 공여자 서열은 DSB를 생성하기 위해 사용되는 뉴클레아제 중 하나 이상에 대해 동일한 표적 부위 중 하나 이상을 포함할 수 있다. 따라서, 공여자 서열은 통합이 요망되는 내인성 유전자를 절단하기 위해 사용되는 동일한 뉴클레아제 중 하나 이상에 의해 절단될 수 있다. 특정 실시형태에서, 공여자 서열은 DSB를 유도하기 위해 사용되는 뉴클레아제와 상이한 뉴클레아제 표적 부위를 포함한다. 표적 세포의 계놈에서 DSB는 임의의 메커니즘에 의해 생성될 수 있다. 특정 실시형태에서, DSB는 관심 대상의 영역, 및 절단 도메인 또는 절단 절반-도메인 내의 서열에 결합하도록 조작된 아연 평거 결합 도메인을 포함하는 융합 단백질인 하나 이상의 아연-평거 뉴클레아제(ZFN)에 의해 생성된다.

- [0038] 일 양상에서, 공여자는 관심 대상의 유전자에 결합하거나, 발현을 조절하거나 또는 결합과 발현의 조절 둘 다인 관심 대상의 조절 단백질(예를 들어, ZFP TF, TALE TF 또는 CRISPR/Cas TF)을 암호화할 수 있다. 일 실시형태에서, 조절 단백질은 DNA 서열에 결합하고, 다른 조절 인자의 결합을 방지한다. 다른 실시형태에서, 조절 단백질의 결합은 표적 DNA의 발현을 조절할 수 있다(즉, 유도하거나 또는 억제한다).
- [0039] 일부 실시형태에서, 유전자이식 HSC/PC 세포, 유전자이식 동물 또는 유전자이식 HSC/PC 세포와 동물의 조합은 인간 유전자를 암호화하는 이식유전자를 포함한다. 일부 예에서, 유전자이식 동물은 내인성 좌위에서 네아웃, 및 그의 인간 상대로 내인성 유전자의 대체를 포함함으로써, 생체내 시스템의 발생을 허용하며, 여기서 인간 단백질은 단리에서 연구될 수 있다. 이러한 유전자이식 모델은 관심 대상의 인간 단백질과 상호작용하거나 또는 변형할 수 있는 소분자 또는 거대 생체분자 또는 다른 독립체를 동정하기 위한 선별 목적을 위해 사용될 수 있다. 일부 양상에서, 이식유전자는 줄기 세포(예를 들어, 배아 줄기 세포, 유도만능 줄기 세포, 조혈 줄기 세포 등) 또는 본 명세서에 기재된 임의의 방법에 의해 얻어지는 동물 배아 내로 선택된 좌위(예를 들어, 세이프-하버(safe-harbor))에 통합되며, 이어서, 살아있는 동물이 태어나도록 배아는 이식된다. 이어서, 동물은 성적 성숙에 대해 상승되고, 새끼를 생산하도록 허용되며, 새끼 중 적어도 일부는 편집된 내인성 유전자 서열 또는 통합된 이식유전자를 포함한다.
- [0040] 다른 양상에서, 본 명세서에서 세포에서 유전자 발현(예를 들어, BCL11A, 글로빈 유전자, 및 BCL11A와 글로빈 유전자의 조합)을 변경시키는 방법이 제공되며, 상기 방법은 1종 이상의 단백질이 발현되고 유전자 발현이 변경되는 조건 하에서 본 명세서에 기재된 바와 같은(표 1에 나타냄) 1종 이상의 뉴클레아제를 세포 내로 도입하는 단계를 포함한다. 특정 실시형태에서, 글로빈 유전자(예를 들어, 감마 글로빈 또는 베타 글로빈)의 발현은 변경된다(예를 들어, 증가된다). 본 명세서에 기재된 임의의 방법은 세포의 계놈 내로 공여자 서열(예를 들어, 외인성 또는 내인성 프로모터의 제어 하에서 이식유전자 또는 이의 단편)을 통합하는 단계, 예를 들어 BCL11A 유전자 내 뉴클레아제 절단 부위에서 또는 근처에서 공여자를 통합하는 단계를 추가로 포함할 수 있다. 공여자 서열은 바이러스 벡터를 이용하여, 올리고뉴클레오파이드로서, 플라스미드 상에서 그리고 바이러스 벡터, 올리고뉴클레오파이드로서 또는 플라스미드 상으로부터 선택되는 하나 이상의 방법의 조합으로 세포에 도입된다. 유전자 발현이 변경된 세포는, 예를 들어, 적혈구(RBC) 전구체 세포, 조혈 줄기 세포(예를 들어, CD34+ 세포) 및 RBC 전구체 세포와 조혈 줄기 세포의 조합일 수 있다.
- [0041] 다른 실시형태에서, 본 명세서에서 내인성 BCL11A 인핸서 서열 내의 계놈 변형(BCL11A 인핸서 서열의 뉴클레오파이드 서열에 대한 변형)을 포함하는 유전자 변형된 세포를 생산하는 방법이 제공되며, 상기 방법은: a) 세포를 4, 5 또는 6개의 아연 평거 도메인을 포함하는 아연 평거 뉴클레아제를 암호화하는 폴리뉴클레오파이드(예를 들어, DNA 또는 mRNA)와 접촉시키는 단계로서, 각각의 아연 평거 도메인은 표 1의 단일행에 나타낸 순서로 인식나선 영역을 포함하는, 상기 접촉시키는 단계; b) 상기 세포를 폴리뉴클레오파이드로부터 아연 평거 단백질을 발현시키는 데 도움이 되는 조건으로 처리하는 단계; 및 c) 내인성 BCL11A 인핸서 서열을 변형시키는 단계가 제공된다. 특정 실시형태에서, 세포는(예를 들어, 단계 (a) 전에) 적어도 1종의 사이토카인으로 자극된다. 폴리뉴클레오파이드는 형질감염을 통해, 비바이러스 벡터를 이용하여, 바이러스 벡터를 이용하여, 화학적 수단에 의해 또는 전기장에 대한 노출(예를 들어, 전기천공법)을 포함하지만, 이들로 제한되지 않는 임의의 적합한 방법을 이용하여 세포와 접촉될 수 있다.
- [0042] 본 명세서에 기재된 방법에 의해 생산된 세포로부터 유래된 세포를 포함하는 본 명세서에 기재된 계놈 변형 중 한 또는 조합을 포함하는 세포가 또한 제공된다.
- [0043] 또한 글로빈 유전자 발현에서의 증가가 필요한 환자를 치료하는 방법이 제공되며, 상기 방법은 환자에게 약제학적 제제를 투여하는 단계를 포함하되, 약제학적 제제는 환자에서 글로빈 유전자 발현을 증가시키기 위한 충분한 양으로 본 명세서에 기재된 바와 같은 유전자 변형된 세포, 단백질, 폴리뉴클레오파이드 및 유전자 변형된 세포, 단백질과 폴리뉴클레오파이드로부터 선택되는 하나 이상의 조합을 포함한다. 특정 실시형태에서, 환자는 탈라세미아 또는 겸상 적혈구병을 갖는 것으로 알려져 있거나, 갖는 것으로 의심되거나 또는 이들이 발생할 위험에 있다.
- [0044] 본 발명의 핵산, 단백질, 유전자 변형된 세포 및 핵산, 단백질과 유전자 변형된 세포로부터 선택되는 하나 이상의 조합을 포함하는 키트가 또한 제공된다. 키트는 뉴클레아제를 암호화하는 핵산(예를 들어, 적합한 발현 벡터에 함유된 유전자를 암호화하는 RNA 분자 또는 ZFN, TALEN 또는 CRISPR/Cas 시스템), 뉴클레아제 단백질의 분취액, 공여자 분자, 줄기 세포 자기 재생의 적합한 변형제("줄기세포능(stemness)"), 세포, 완충제, 설명서(예를 들어, 본 발명의 방법을 수행하기 위함) 등(이들 키트 부품의 다양한 조합을 포함)을 포함할 수 있다. 본 발명

은 뉴클레아제에 의해 이루어진 적어도 하나의 게놈 변형(예를 들어, 표 1의 단일행에 나타낸 바와 같음)을 포함하는 유전자 변형된 세포(예를 들어, 줄기 세포, 예컨대 조혈(CD34+) 줄기 세포 또는 RBC 전구체 세포)를 포함하지만, 이들로 제한되지 않되, 게놈 변형은 내인성 BCL11A 인핸서 서열 내이고, 추가로 게놈 변형은 삽입, 결실 및 이들의 조합으로 이루어진 군으로부터 선택되고, 임의의 서열번호 1과 서열번호 2에서, 근처에서 또는 사이에서 변형을 포함한다. 특정 실시형태에서, 세포는 본 명세서에 기재된 바와 같은 줄기 세포로부터 유래된 유전자 변형된 분화 세포(예를 들어, 조혈 줄기 세포 또는 RBC 전구체 세포로부터 유래된 RBC)이다.

[0045] 뉴클레아제는 적어도 하나의 아연 평거 뉴클레아제(ZFN)(예를 들어, 표 1에 나타내는 바와 같음), 적어도 하나의 TALEN, 및 적어도 하나의 ZFN과 적어도 하나의 TALEN의 조합을 포함할 수 있다. 뉴클레아제(들)는 단백질 형태로, 뉴클레아제(들)를 암호화하는 폴리뉴클레오타이드로서, 또는 단백질 형태와 뉴클레아제(들)를 암호화하는 폴리뉴클레오타이드의 조합으로서 세포 내로 도입될 수 있다. 특정 실시형태에서, 게놈 변형은 이식유전자를 암호화하는 공여자 폴리뉴클레오타이드의 통합을 포함하는 삽입을 포함한다. 또한 본 명세서에 기재된 바와 같은 유전자 변형된 세포 중 하나 이상을 포함하는 약제학적 조성물이 제공된다.

[0046] 또한 인식 나선 영역을 포함하는 4, 5 또는 6개의 아연 평거 도메인을 포함하는 아연 평거 단백질을 포함하는 DNA-결합 단백질이 제공되되, 아연 평거 단백질은 표 1의 단일행에 나타낸 순서로 인식 나선 영역을 포함한다. 또한 표 1에 나타낸 표적 부위의 일부(예를 들어, 적어도 4, 5, 6개 이상의) 염기쌍을 포함하는 서열에 결합하는 복수의 반복부를 포함하는 TALE 단백질이 제공된다. 본 명세서에 기재된 바와 같은 아연 평거 단백질 또는 TALE 단백질 및 야생형 또는 조작된 절단 도메인 또는 절단 절반-도메인을 포함하는 융합 단백질은 본 명세서에 기재된 단백질(ZFP, TALE, ZFN, TALEN)을 암호화하는 폴리뉴클레오타이드로서 제공된다. 본 명세서에 기재된 바와 같은 하나 이상의 폴리뉴클레오타이드, 단백질, 및 폴리뉴클레오타이드와 단백질의 조합을 포함하는 세포(예를 들어, 단리된 줄기 세포, 예컨대 조혈(CD34+) 줄기 세포)가 또한 제공된다. 또한 본 명세서에 기재된 바와 같은 하나 이상의 단백질, 폴리뉴클레오타이드, 세포 또는 이들의 조합을 포함하는 키트가 제공된다.

[0047] 세포(예를 들어, RBC 전구체 세포, 조혈 줄기 세포 및 RBC 전구체 세포와 조혈 줄기 세포의 조합)에서 글로빈 유전자 발현을 변경시키는 방법이 또한 기재되며, 상기 방법은 하나 이상의 단백질이 발현되고 글로빈 유전자(예를 들어, 감마 글로빈, 베타 글로빈, 및 감마 글로빈과 베타 글로빈의 조합)의 발현이 변경되는(예를 들어, 증가되는) 조건 하에서, 본 명세서에 기재된 하나 이상의 뉴클레아제를 암호화하는 하나 이상의 폴리뉴클레오타이드를 암호화하는 하나 이상의 폴리뉴클레오타이드를 도입하는 단계를 포함한다. 특정 실시형태에서, 상기 방법은, 예를 들어, 바이러스 벡터를 이용하여, 올리고뉴클레오타이드로서 또는 플라스미드 상에서 세포의 게놈 내로 공여자 서열을 통합하는 단계를 추가로 포함한다. 공여자 서열은 내인성 또는 외인성 프로모터의 제어 하에 이식유전자를 포함할 수 있다.

[0048] 또한 내인성 BCL11A 인핸서 서열(예를 들어, 표 1에 나타낸 바와 같은 표적 부위) 내에서 게놈 변형을 포함하는 유전자 변형된 세포를 생산하는 방법이 제공되며, 상기 방법은 (a) 세포를 4, 5 또는 6개의 아연 평거 도메인을 포함하는 아연 평거 뉴클레아제를 포함하는 융합 단백질을 암호화하는 폴리뉴클레오타이드와 접촉시키는 단계로서, 각각의 아연 평거 도메인은 표 1의 단일 행에 나타낸 순서로 인식 나선 영역을 포함하는, 상기 접촉시키는 단계; (b) 상기 세포를 폴리뉴클레오타이드로부터의 융합 단백질을 발현시키는 데 도움이 되는 조건으로 처리하는 단계; 및 (c) 유전자 변형된 세포를 생산하는 데 충분한 발현된 융합 단백질을 이용하여 내인성 BCL11A 인핸서 서열을 변형시키는 단계를 포함한다. 특정 실시형태에서, 상기 방법은 적어도 1종의 사이토카인을 이용하여 세포를 자극하는 단계를 추가로 포함한다. 폴리뉴클레오타이드(들)는, 예를 들어, 비-바이러스 전달 시스템, 바이러스 전달 시스템, 전달 비히클, 및 비-바이러스 전달 시스템, 바이러스 전달 시스템 및 전달 비히클로부터 선택된 조합을 이용하여 세포 내부로 전달될 수 있고, 세포에 전기장 처리하거나 또는 전달 메커니즘으로서 세포막을 사용하는 것을 포함할 수 있다(소위 '스퀴즈 기법(Squeeze Technology)', 예를 들어, 문헌[Sharei *et al.* (2015) *PLOS ONE* doi: 10.1371/journal/pone.0118803] 참조).

[0049] 글로빈 유전자 발현의 증가가 필요한 환자(예를 들어, 이상혈색소증, 예컨대 탈라세미아(예를 들어, β -탈라세미아) 또는 겸상 적혈구병을 갖는 것으로 알려져 있거나, 갖는 것으로 의심되거나 또는 이들이 발생할 위험에 있는 환자)를 치료하는 방법이 또한 제공되며, 상기 방법은 환자에서 글로빈 유전자 발현을 증가시키기에 충분한 양으로 본 명세서에 기재된 약제학적 조성물(예를 들어, 단백질, 폴리뉴클레오타이드, 세포 또는 백질, 폴리뉴클레오타이드 및 세포로부터 선택되는 조합)을 환자에게 투여하는 단계를 포함한다.

[0050] 이들 및 다른 양상은 전체로서 본 개시내용에 비추어 당업자에게 용이하게 명확할 것이다.

도면의 간단한 설명

[0051]

도 1A 및 도 1B는 올리뉴클레오타이드 이중가닥 통합 부위 분석의 개요를 도시하는 개략도를 도시한 도면. 도 1은 얻어진 절단 사전의 분획에 포함된 보충적 올리고뉴클레오타이드 이중가닥 DNA의 존재 하에 세포가 ZFN으로 처리되는 분석의 제1 단계를 나타낸다. 도 1B는 세포가 7일 동안 배양되고, 계놈 DNA가 단리되며, 통합 부위에 측집하는 계놈의 세그먼트는 통합된 올리고뉴클레오타이드 이중가닥에 프라이머를 이용하여 어댑터-매개 PCR을 통해 증폭되는 후속 단계를 나타낸다. 이어서, 앰플리콘은 후보 절단 부위를 나타내도록 서열분석된다.

도 2는 본래의 51857/51949 또는 모 ZFN 쌍을 이용하는 올리고뉴클레오타이드 이중가닥 통합 부위 분석에 의해 동정된 455개의 잠재적인 표적을 벗어난 좌위를 도시한다. 상부 63개의 좌위는 회색으로 강조하며, 후속 삽입결실 분석에서 분석하였다. 각각의 좌위는 절단 위치를 나타내는 염색체 및 염기 수뿐만 아니라 해당 좌위에서 검출된 구성요소 수의 표시에 의해 동정된다.

도 3a 내지 도 3c는 아연 평거 스파스풀드 내의 예시적인 인산염 접촉 잔기를 도시한 도면. 도 3a는 DNA 문자와의 아연 평거 상호작용을 도시하며, 야생형 알기닌 측쇄의 위치 및 DNA 문자의 인산염 골격과 상호작용하는 방법을 나타낸다. 도 3b(서열번호 7 내지 17)은 이 알기닌이 각각의 아연 평거의 1차 서열에 위치되고(볼드체 화살표를 통해 표시된 'R' 잔기의 수직 행) 또한 대응하는 인산염 접촉(개개로 박스 표시된 'R' 잔기)를 제거하기 위해 ZFP 골격에서 글루타민으로 치환된 해당 알기닌을 강조한 경우를 나타내는 예시적인 ZFN 서열을 도시한다. 서열은 또한 인식 나선 영역(여기서 잔기 -1 내지 +3, +5 및 +6은 박스표시 및 음영표시됨)뿐만 아니라 C-말단의 아연 평거 도메인과 절단 도메인 사이의 링커 부분(절단 도메인은 나타내지 않음)을 나타낸다. 도 3c는 특이성 최적화 동안 세린으로 치환될 수 있는 FokI 절단 도메인에 존재하는 라이신 측쇄와 접촉되는 다른 잠재적 골격의 공간적 위치를 추가로 도시한다.

도 4는 CD34 세포의 처리를 통한 변형 수준을 다양한 시험 좌위에서 본래의 51857/51949 쌍 또는 최적화된 63014/65722 쌍 중 하나와 비교하는 그래프를 도시한 도면. 하부에 걸쳐 또는 x-축 상에서, 잠재적 절단 표적으로서 동정된 좌위를 나타내며, 수직 또는 y-축 상에 각각의 부위에 대한 삽입결실%를 나타낸다. y-축은 로그 규모로 나타낸다는 것을 주목한다. 진한 회색 막대는 51857/51949 쌍에 의해 절단되는 좌위, 및 검출된 절단의 양을 나타내는 반면, 밝은 회색 막대는 최적화된 쌍에 의해 절단되는 해당 좌위이고, 여기서, 측정된 거의 모든 절단은 표적화된 BCL11a 표적 서열에 있다.

도 5A 및 도 5B는 BCL11A 특이적 ZFN 및 적혈구 분화에 의한 치료 후에 hCD34+ 세포에서 이루어진 글로빈 mRNA의 상대적 비를 검출하는 그래프를 도시한 도면. 2명의 건강한 인간 공여자(PB-MR-003 또는 PB-MR-004)로부터 유래된 CD34+ 세포를 ZFN 쌍으로 처리하거나 또는 처리하지 않았고, 이어서, α , β 및 γ 글로빈 발현을 분석하였다. ZFN 처리 후에 발견된 γ 글로빈 mRNA의 양을 결정하는 최고의 방법은 γ 글로빈 대 β 글로빈 비(도 5A), 또는 γ 글로빈 대 α 글로빈 mRNA 비 중 하나로서 발현의 변화를 표현하는 것이다(도 5B).

도 6은 처리된 CD34+ 세포에서 생산된 γ -글로빈 단백질의 상대적 양을 도시한 그래프를 도시한 도면. 상기와 같이, 건강한 인간 공여자로부터 유래된 두 CD34+ 세포 로트를 사용하였다(PB-MR-003 및 PB-MR-004). 이 실험에서, 약 15% 내지 20% 수준까지 태아 글로빈 단백질 백분율의 대략 3 내지 4배 상승이 공여자 로트 둘 다에서 BCL11A 인핸서의 63014/65722-매개 붕괴 시 HSPC의 적혈구 자손에서 관찰되었다.

도 7은 상기 기재한 바와 같은 63014/65722 처리 공여자 로트("+ZFN")가 생착된 마우스에서 상대적 인간 키메라 현상(chimerism)을 나타내는 그래프를 도시한 도면. 인간 키메라 현상은 FACS를 이용하여 그들의 표면 상에서 hCD45 마커를 보유하는 세포의 검출을 통해 측정된다. 이식 후 8 또는 12주에 수집한 말초 혈액 중의 인간 hCD45+ 세포의 백분율을 나타낸다. 데이터는 비형질감염 대조군("(-)") 및 ZFN 형질감염 HSPC("+ZFN")의 생착 후 비교할 만한 인간 키메라 현상을 이용하여 이 연구에서 양호한 생착 수준을 나타내었다. 속이 빈 원 및 삼각형은 개개 동물을 나타낸다.

도 8은 인간 세포가 그들의 세포 표면 상에서 hCD45의 존재에 의해 동정된 생착 마우스의 골수에서 검출된 키메라 현상의 백분율을 도시한 도면. 생착 후 12주에 샘플을 분석하였다.

도 9a 내지 도 9d는 계통 특이적 세포 표면 마커를 인식하는 항체를 이용하여 제12주에 얻은 생착 마우스에서 골수 세포의 FACS 분석에 의해 시험되는 다양한 조혈 세포 계통의 재구성을 도시하는 그래프를 도시한 도면. 데이터는 BCL11A-특이적 ZFN mRNA 처리된 CD34+ 세포 자손("14/'22")과 비형질감염 세포의 그것("(-)") 사이에서 주사 후 제12주에 골수 내 모든 분석 인간 조혈 계통의 비슷한 제시를 나타내었다. 공여자 둘 다("003"과 "004")로부터 유래된 세포에 대해 림프구, 골수, 적혈구 및 HSPC로부터의 데이터를 나타낸다(각각 도 9a 내지

도 9d). 데이터는 Bc111A ZFN mRNA 처리된 CD34+ 세포 자손과 비형질감염 세포의 그것 사이에서 주사 후 제12주에 골수 내 모든 분석 인간 조혈 계통의 비슷한 제시를 나타내었다.

도 10은 심층 서열분석에 의해 분석되는 생착 마우스의 말초 혈액으로부터 단리된 DNA 내 BCL11A 표적에서의 유전자 변형 수준을 도시하는 그래프를 도시한 도면. 유입 세포에 대한 데이터(ZFN 형질감염 후 2일, ("+")), 그리고, 이어서 생착 후 8 또는 12주에 혈액 세포에 대한 데이터를 나타내고, 유전자 변형의 양호한 보유를 입증하였다. 비형질감염 세포를 처리선에서 "(-)"로 나타낸다.

도 11은 ZFN-처리 세포("+)의 생착 후 골수 세포 샘플에 대한 BCL11A 표적에서 유전자 변형의 양을 도시하는 그래프를 도시한 도면. 비처리 세포를 처리선에서 "(-)"로 나타낸다. 비슷한 변형이 BCL11A 의존적 계통(CD19 마커를 발현시키는 B 세포; CD45 및 고수준의 CD38을 발현시키는 원시 전구체) 및 BCL11A 독립적(골수) 계통 둘다에서 관찰된다. 유입 유전자 변형 수준은 PB-MR-004 공여자 샘플에서보다 PB-MR-003 공여자 샘플에서 더 높았지만, PB-MR-004 유래 세포는 일관되게 더 높은 변형 수준, 즉, PB-MR-003으로부터 유래된 것보다 마우스에서 변형의 더 양호한 보유를 나타내었다.

도 12는 14일 동안 시험관내 분화 후 12주 골수 세포로부터 유래된 적혈구 세포 내 유전자 변형의 양을 도시하는 그래프를 도시한 도면. 데이터는 상기 기재한 두 상이한 공여자에 의해 본래 생착된 마우스로부터 유래되며, ZFN 처리("+ ZFN")에 의해 매개된 BCL11A 변형은 적혈구 분화 동안 현저하게 변경되지 않는다는 것을 입증하였다. ZFN으로 처리하지 않은 세포를 "(-) ZFN"으로 나타낸다.

도 13A 내지 도 13B는 mRNA를 암호화하는 γ -글로빈의 상대적 양을 도시하는 그래프이며, 여기서 γ -글로빈 mRNA의 농도는 γ 글로빈/ β 글로빈 mRNA의 비로서(도 9a) 또는 γ 글로빈/ α 글로빈 mRNA의 비로서 도시된다(도 9b). 비형질감염 ("(-) ZFN") 샘플과 ZFN 처리된 샘플("+ ZFN") 둘다에서, γ -글로빈 대 β 글로빈 또는 γ -글로빈 대 α -글로빈 mRNA 비는 동일한 그룹으로부터의 개개 마우스의 적혈구 자손 간에 크게 상이하였다. 그러나, 이런 가변성 두 상이한 인간 공여자의 사용에 의해 도입된 가변성에도 불구하고, 63014/65722 처리 샘플 평균은 그들의 각각의 비형질감염 상대에 비해 γ 글로빈 mRNA 수준의 대략 1.5 내지 2배 증가를 나타낸다.

도 14는 생착 마우스로부터의 골수 유래 세포에서 γ -글로빈 단백질의 양(γ -글로빈/ α -글로빈 또는 γ -글로빈/총 β -유사 단백질의 비로서 표현)의 차이를 도시하며, 여기서 골수 세포는 시험관내 차이 프로토콜로 제출되는 것을 도시한 도면. 단백질 수준을 제16일에 분화로 측정하였다. 감마(γ) 글로빈(A감마와 G감마 피크의 합) 대 알파(α) 글로빈 비뿐만 아니라 감마 글로빈(A감마와 G감마 피크의 합)/베타-유사 글로빈 비(A감마, G감마, 베타 및 델타-글로빈 피크의 합)를 결정하였다. PB-MR-003 유래 샘플의 불량한 적혈구 분화와 일치되게, 이 공여자로부터 유래된 비형질감염 세포 내 감마-글로빈 수준은 매우 높았고(대략 30%), 따라서 ZFN 처리("+ ZFN")는 비처리 세포("(-) ZFN")에 비해 감마-글로빈 수준의 단지 1.2배 증가를 초래하였다. PB-MR-004는 더 전형적인 비형질감염 수준(대략 9%)을 나타내었고, 마우스를 통한 12주 계대 후에 감마-글로빈 단백질의 대략 2배 증가를 나타내었다.

발명을 실시하기 위한 구체적인 내용

[0052] 본 명세서에서 BCL11A, 감마 글로빈, 및 BCL11A와 감마 글로빈 발현의 조합의 조절을 위해 그리고 이상혈색소증의 치료, 예방 또는 치료와 예방을 위해 계놈 조작을 위한 조성물 및 방법이 개시된다. 특히, 표 1의 단일 행에 나타내는 바와 같은 인식 나선 영역을 갖는 ZFP를 포함하는 뉴클레아제를 이용하는 표적화를 통해, BCL11A의 인핸서의 붕괴가 HSC/PC에서 효율적으로 달성되며, 후속적 적혈구생성 동안 상대적 감마 글로빈 발현의 변화를 초래한다. BCL11A 및 감마 글로빈 발현의 이런 조절은 이상혈색소증(예를 들어, 베타 탈라세미아, 경상 적혈구병)의 치료에 특히 유용하되, 불충분한 베타 글로빈 발현 또는 베타-글로빈의 돌연변이 형태의 발현이 있다. 본 발명의 방법 및 조성물을 이용하여, 비정상 베타 글로빈에 의해 야기되는 합병증 및 질환 관련 후유증은 적혈구 전구체 세포에서의 감마 글로빈의 발현 변경에 의해 극복될 수 있다.

일반

[0053] 본 명세서에 개시된 방법의 실행뿐만 아니라 제제 및 조성물의 용도는 달리 표시되지 않는 한, 당업계의 기술내인 문자 생물학, 생화학, 염색질 구조 및 분석, 컴퓨터 화학, 세포 배양물, 재조합 DNA 및 관련 분야에서 통상적인 기법을 사용한다. 이들 기법은 문헌에서 완전하게 설명된다. 예를 들어, 문헌[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 개신 주기율표]; 문헌[시리즈 METHODS IN ENZYMOLOGY, Academic Press, San Diego; Wolffe,

CHROMATIN STRUCTURE AND FUNCTION, Third edition, Academic Press, San Diego, 1998; METHODS IN ENZYMOLOGY, Vol. 304, "Chromatin" (P.M. Wasserman and 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, 1999] 참조.

[0055] 정의

용어 "핵산", "폴리뉴클레오타이드" 및 "올리고뉴클레오타이드"는 상호 호환적으로 사용되며, 선형 또는 원형 입체배좌로, 그리고 단일- 또는 이중 가닥 형태로 데옥시리보뉴클레오타이드 또는 리보뉴클레오타이드 중합체를 지칭한다. 본 개시내용의 목적을 위해, 이를 용어는 중합체 길이에 대한 제한으로 해석되어서는 안 된다. 상기 용어는 천연 뉴클레오타이드의 공지된 유사체뿐만 아니라 염기, 당, 인산염 모이어티에서 변형된 뉴클레오타이드(예를 들어, 포스포로티오에이트 골격), 및 염기, 당 및 인산염 모이어티로부터 선택된 조합을 포함할 수 있다. 일반적으로, 특정 뉴클레오타이드의 유사체는 동일한 염기쌍 특이성을 가지며; 즉, A의 유사체는 T와 염기쌍을 형성할 것이다.

[0056] 용어 "폴리펩타이드", "펩타이드" 및 "단백질"은 아미노산 잔기의 중합체를 지칭하기 위해 상호 호환적으로 사용된다. 상기 용어는 또한 하나 이상의 아미노산이 대응하는 천연 유래 아미노산의 화학적 유사체 또는 변형된 유도체인 아미노산 중합체에 적용된다.

[0057] [0058] "결합하는"은 거대분자 사이의(예를 들어, 단백질과 핵산 사이의) 서열-특이적, 비공유 상호작용을 지칭한다. 전체로서 상호작용이 서열 특이적이라면, 결합 상호작용의 모든 성분이 서열 특이적(예를 들어, DNA 골격에서 인산염 잔기와 접촉)일 필요는 없다. 이러한 상호작용은 일반적으로 10^{-6} M^{-1} 또는 더 낮은 해리상수(K_d)를 특징으로 한다. "친화도"는 결합 강도를 지칭하며: 증가된 결합 친화도는 더 낮은 K_d 와 상관관계가 있다.

[0059] "결합 단백질"은 다른 분자에 결합할 수 있는 단백질이다. 결합 단백질은, 예를 들어, DNA 분자(DNA-결합 단백질), RNA 분자(RNA-결합 단백질), 단백질 분자(단백질-결합 단백질)에 결합할 수 있거나, 또는 DNA 분자, RNA 분자 또는 단백질로부터 선택되는 분자의 조합에 결합할 수 있다. 단백질-결합 단백질의 경우에, 이는 (동종이량체, 동종삼량체 등을 형성하기 위해) 그 자체에 결합할 수 있거나, 상이한 단백질 또는 단백질들의 하나 이상의 분자에 결합할 수 있거나, 또는 그 자체 또는 상이한 단백질 또는 단백질들의 하나 이상의 분자에 결합할 수 있다. 결합 단백질은 결합 활성의 하나 초과의 유형을 가질 수 있다. 예를 들어, 아연 평거 단백질은 DNA-결합, RNA-결합 및 단백질-결합 활성을 가진다.

[0060] "아연 평거 DNA 결합 단백질"(또는 결합 도메인)은, 구조가 아연 이온의 배위를 통해 안정화되는 결합 도메인 내의 아미노산 서열 영역인 하나 이상의 아연 평거를 통해 서열-특이적 방식으로 DNA에 결합하는 단백질 또는 더 큰 단백질 내의 도메인이다. 용어 아연 평거 DNA 결합 단백질은 종종 아연 평거 단백질 또는 ZFP로서 약칭된다.

[0061] "TALE DNA 결합 도메인" 또는 "TALE"은 하나 이상의 TALE 반복부 도메인/단위를 포함하는 폴리펩타이드이다. 반복부 도메인은 그의 동족 표적 DNA 서열에 대한 TALE의 결합에 연루된다. 단일 "반복부 단위"(또한 "반복부"로서 지칭됨)는 전형적으로 길이가 33 내지 35개의 아미노산이며, 천연 유래 TALE 단백질 내에서 다른 TALE 반복부 서열과 적어도 일부 서열 상동성을 나타낸다.

[0062] [0063] 아연 평거 및 TALE 결합 도메인은, 예를 들어 천연 유래 아연 평거 또는 TALE 단백질의 인식 나선 영역의 조작(하나 이상의 아미노산의 변경)을 통해 사전결정된 뉴클레오타이드 서열에 결합하도록 "조작"될 수 있다. 따라서, 조작된 DNA 결합 단백질(아연 평거 또는 TALE)은 비천연 유래 단백질이다. DNA-결합 단백질을 조작하기 위한 방법의 비제한적 예는 설계 및 선택이다. 설계된 DNA 결합 단백질은 설계/조성물이 원칙적으로 합리적 기준으로부터 초래되는 천연에서 발생하지 않는 단백질이다. 설계를 위한 합리적 기준은 치환 규칙의 적용 및 존재하는 ZFP 및/또는 TALE 설계 및 결합 데이터의 데이터베이스 저장 정보에서 정보를 가공하기 위한 컴퓨터화된 알고리즘을 포함한다. 예를 들어, 미국 특허 제6,140,081호; 제6,453,242호; 제6,534,261호 및 제8,585,526호를 참조하며; 또한 국제특허출원 공개 WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 및 WO 03/016496 참조.

"선택된" 아연 평거 단백질 또는 TALE은 생산이 주로 경험적 과정, 예컨대 파지 디스플레이, 상호작용 트랩 또는 하이브리드 선택으로부터 초래되는 특성에서 발견되지 않는 단백질이다. 예를 들어, 미국 특허 제5,789,538호; 제5,925,523호; 제6,007,988호; 제6,013,453호; 제6,200,759호; 제8,586,526호; 국제특허출원 WO

95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970 WO 01/88197, WO 02/099084 참조.

[0064] "TtAgo"는 유전자 침묵에 수반되는 것으로 생각되는 원핵생물 아거노트 단백질이다. TtAgo는 박테리아 씨무스 씨모필루스(*Thermus thermophilus*)로부터 유래된다. 예를 들어, 문헌[Swarts *et al.* *ibid*, G. Sheng *et al.* (2013) *Proc. Natl. Acad. Sci. U.S.A.* 111, 652] 참조. "TtAgo 시스템"은, 예를 들어, TtAgo 효소에 의한 절단을 위한 가이드 DNA를 포함하는 필요한 모든 성분이다.

[0065] "재조합"은 비상동상 말단 접합(NHEJ) 및 상동성 재조합에 의해 야기되는 공여자 포획을 포함하지만, 이들로 제한되지 않는 두 폴리뉴클레오타이드 사이의 유전자 성분의 교환 과정을 지칭한다. 본 개시내용의 목적을 위해, "상동성 재조합(HR)"은 상동성-지향 수선 메커니즘을 통해 세포 내 이중가닥 파손의 수선 동안 일어나는 이러한 교환의 전문화된 형태를 지칭한다. 이 과정은 뉴클레오타이드 서열 상동성을 필요로 하며, "표적" 분자(즉, 이중 가닥 파손을 경험한 것)의 주형 수선에 대한 "공여자" 분자를 사용하고, "비-크로스오버 유전자 전환" 또는 "짧은 트랙 유전자 전환"으로서 다양하게 알려져 있는데, 그것이 공여자로부터 표적까지 유전자 정보의 전달을 야기하기 때문이다. 임의의 특정 이론에 의해 구속되는 일 없이, 이러한 전달은 파손된 표적과 공여자 사이에 형성되는 이형이중가닥 DNA의 미스매치 보정인 "합성-의존적 가닥 어닐링"을 수반할 수 있는데, 이때 공여자는 표적의 부분, 관련 과정, 또는 이들의 조합이 될 유전자 정보를 재합성하기 위해 사용된다. 이러한 전문화된 HR은 종종 공여자 폴리뉴클레오타이드 서열의 부분 또는 모두가 표적 폴리뉴클레오타이드 내로 흡입되도록 표적 분자 서열의 변경을 초래한다.

[0066] 개시내용의 방법에서, 본 명세서에 기재된 바와 같은 하나 이상의 표적화된 뉴클레아제는 사전결정된 부위에서 표적 서열(예를 들어, 세포의 염색질) 내 이중 가닥 파손(DSB)을 생성한다. DSB는 상동성-지향 수선에 의해 또는 비상동성-지향 수선 메커니즘에 의해 삽입결실을 초래할 수 있다. 결실은 다수의 염기쌍을 포함할 수 있다. 유사하게, 삽입은, 예를 들어, 파손 영역에서 뉴클레오타이드 서열에 대해 상동성을 선택적으로 갖는 "공여자" 폴리뉴클레오타이드의 통합을 포함하는 다수의 염기쌍을 포함할 수 있다. 공여자 서열은 물리적으로 통합될 수 있거나 또는 대안적으로, 공여자 폴리뉴클레오타이드는 상동성 재조합을 통한 파손의 수선을 위한 주형으로서 사용되어, 세포 염색질 내로 공여자 그 자체의 뉴클레오타이드 서열의 모두 또는 부분의 도입을 초래한다. 따라서, 세포 염색질의 제1 서열은 변경될 수 있고, 특정 실시형태에서, 공여자 폴리뉴클레오타이드에 존재하는 서열로 전환될 수 있다. 따라서, 용어 "대체하다" 또는 "대체"의 사용은 하나의 뉴클레오타이드 서열의 다른 뉴클레오타이드 서열로의 대체를 나타내는 것으로 이해될 수 있고(즉, 정보적 의미에서 서열의 대체), 하나의 폴리뉴클레오타이드의 다른 폴리뉴클레오타이드에 의한 물리적 또는 화학적 대체를 반드시 필요로 하는 것은 아니다.

[0067] 본 명세서에 기재된 임의의 방법에서, 아연-핑거 단백질 또는 TALEN의 추가적인 쌍은 세포 내에서 추가적인 표적 부위의 이중-가닥 절단을 위해 사용될 수 있다.

[0068] 본 명세서에 기재된 임의의 방법은 임의의 크기의 공여자의 삽입을 위해, 또는 관심 대상의 유전자(들)의 발현을 방해하는 공여자 서열의 표적화된 통합에 의한 세포 내 하나 이상의 표적 서열의 부분적 또는 완전한 비활성화를 위해 사용될 수 있다. 부분적으로 또는 완전히 비활성화된 유전자를 갖는 세포주가 또한 제공된다.

[0069] 본 명세서에 기재된 임의의 방법에서, 외인성 뉴클레오타이드 서열("공여자 서열" 또는 "이식유전자")은 관심 대상의 영역 내 게놈 서열에 대해 상동성이지만, 동일하지는 않은 서열을 함유할 수 있고, 이에 의해 관심 대상의 영역에서 동일하지 않은 서열을 삽입하도록 상동성 재조합을 자극할 수 있다. 따라서, 특정 실시형태에서, 관심 대상의 영역에서 서열에 대해 상동성인 공여자 서열의 일부는 대체되는 게놈 서열에 대해 약 80 내지 99% (또는 그 사이의 임의의 정수) 서열 동일성을 나타낸다. 다른 실시형태에서, 예를 들어 100개의 연속적 염기쌍에 걸쳐 공여자와 게놈 서열 간에 단지 1개의 뉴클레오타이드가 상이하다면, 공여자와 게놈 서열 사이의 상동성은 99% 보다 더 높다. 특정 경우에, 새로운 서열이 관심 대상의 영역 내로 도입되도록 공여자 서열의 비상동성 부분은 관심 대상의 영역 내에 존재하지 않는 서열을 함유할 수 있다. 이들 예에서, 비상동성 서열은 일반적으로 관심 대상의 영역 내 서열에 대해 상동성이거나 또는 동일한 50 내지 1,000개의 염기쌍(또는 그 사이의 임의의 필수적인 값) 또는 1,000개 초과의 다수의 염기쌍에 측정된다. 다른 실시형태에서, 공여자 서열은 제1 서열에 대해 비상동성이며, 비상동성 재조합 메커니즘에 의해 게놈 내로 삽입된다.

[0070] "절단"은 DNA 분자의 공유 골격의 파괴를 지칭한다. 절단은 포스포다이에스터 결합의 효소적 또는 화학적 가수분해를 포함하지만, 이들로 제한되지 않는 다양한 방법에 의해 개시될 수 있다. 단일 가닥 절단과 이중 가닥 절단이 둘 다 가능하며, 이중 가닥 절단은 2개의 별개의 단일 가닥 절단 사건의 결과로서 일어날 수 있다. DNA 절

단은 평활말단 또는 평활 말단 또는 엇갈린 말단 중 하나의 생산을 초래할 수 있다. 특정 실시형태에서, 융합 폴리펩타이드는 표적화된 이중 가닥 DNA 절단을 위해 사용된다.

[0071] "절단 절반-도메인"은 제2 폴리펩타이드(동일 또는 상이한)와 함께 절단 활성(바람직하게는 이중 가닥 절단 활성)을 갖는 복합체를 형성하는 폴리펩타이드이다. 용어 "제1 및 제2 절단 절반-도메인"; "+ 및 - 절단 절반-도메인" 및 "우측 및 좌측 절단 절반-도메인"은 이량체화하는 절단 절반-도메인의 쌍을 지칭하기 위해 상호 호환적으로 사용된다.

[0072] "조작된 절단 절반-도메인"은 다른 절단 절반-도메인(예를 들어, 다른 조작된 절단 절반-도메인)과의 절대 이형 이량체를 형성하기 위해 변형된 절단 절반-도메인이다. 또한, 전문이 본 명세서에 참고로 포함된 미국 특허 제 7,888,121호; 제7,914,796호; 제8,034,598호; 제8,623,618호 및 미국 특허 공개 제2011/0201055호 참조.

[0073] 용어 "서열"은 DNA 또는 RNA일 수 있고; 선형, 원형 또는 분지형일 수 있으며 단일 가닥 또는 이중 가닥일 수 있는, 임의의 길이의 뉴클레오타이드 서열을 지칭한다. 용어 "공여자 서열"은 계놈 내로 삽입되는 뉴클레오타이드 서열을 지칭한다. 공여자 서열은 임의의 서열, 예를 들어, 길이가 2 내지 100,000,000개의 뉴클레오타이드(또는 그 사이의 또는 그것 초과의 임의의 정수), 바람직하게는 길이가 약 100 내지 100,000개의 뉴클레오타이드(또는 그 사이의 임의의 정수), 더 바람직하게는 길이가 약 2000 내지 20,000개의 뉴클레오타이드(또는 그 사이의 임의의 값) 및 훨씬 더 바람직하게는, 약 5 내지 15kb(또는 그 사이의 임의의 값)를 가질 수 있다.

[0074] "염색질"은 세포 계놈을 포함하는 핵단백질 구조이다. 세포의 염색질은 핵산(주로 DNA), 및 단백질(히스톤 및 비히스톤 염색체 단백질을 포함)을 포함한다. 대다수의 세포 염색질은 뉴클레오솜의 형태로 존재하되, 뉴클레오솜 코어는 히스톤 H2A, H2B, H3 및 H4를 각각 2개씩 포함하는 팔량체와 결합된 DNA의 대략 150개 염기쌍을 포함하고; 그리고 (유기체에 따라서 가변 길이의) 링커 DNA는 뉴클레오솜 코어 사이에서 연장된다. 히스톤 H1의 분자는 일반적으로 링커 DNA와 결합된다. 본 개시내용의 목적을 위해, 용어 "염색질"은 모든 유형의 세포 핵단백질(즉, 원핵생물과 진핵생물 둘 다)을 포함하는 것을 의미한다. 세포의 염색질은 염색체와 에피솜 염색질을 둘다 포함한다.

[0075] "염색체"는 세포 계놈의 모두 또는 일부를 포함하는 염색질 복합체이다. 세포 계놈은 세포 계놈을 포함하는 모든 염색체의 수집물인 그의 핵형을 종종 특징으로 한다. 세포 계놈은 하나 이사의 염색체를 포함할 수 있다.

[0076] "에피솜"은 복제 핵산, 핵단백질 복합체 또는 세포의 염색체 핵형의 부분이 아닌 핵산을 포함하는 다른 구조이다. 에피솜의 예는 플라스미드 및 특정 바이러스 계놈을 포함한다.

[0077] "접근 가능한 영역"은 핵산에 존재하는 표적 부위가 표적 부위를 인식하는 외인성 분자에 의해 결합될 수 있는 세포의 염색질 내 부위이다. 임의의 특정 이론에 의해 결합되는 일 없이, 접근 가능한 영역은 뉴클레오솜 구조 내로 패키징되지 않는 것으로 여겨진다. 접근 가능한 영역의 별개의 구조는 화학적 및 효소적 프로브, 예를 들어, 뉴클레아제에 대한 그의 민감도에 의해 검출될 수 있다.

[0078] "표적 부위" 또는 "표적 서열"은 결합에 충분한 조건이 존재한다면, 결합 분자에 결합될 핵산의 일부를 정하는 핵산 서열이다.

[0079] "외인성" 분자는 정상적으로는 세포에 존재하지 않지만, 하나 이상의 유전적, 생화학적 또는 다른 방법에 의해 세포 내로 도입될 수 있는 분자이다. "세포 내의 정상 존재"는 세포의 특정 발생 단계 및 환경 조건에 대해 결정된다. 따라서, 예를 들어, 근육의 배아 발생 동안에만 존재하는 분자는 성인 근육 세포에 대해 외인성 분자이다. 유사하게, 열충격에 의해 유도되는 분자는 비-열충격 세포에 대한 외인성 분자이다. 외인성 분자는, 예를 들어, 제대로 작동 하지 않는 내인성 분자의 기능화 형태 또는 정상적으로 기능하는 내인성 분자의 제대로 작동 하지 않는 형태를 포함할 수 있다.

[0080] 외인성 분자는, 특히, 예컨대 조합 화학 공정에 의해 생성되는 소분자, 또는 거대분자, 예컨대 단백질, 핵산, 탄수화물, 지질, 당단백질, 지방단백질, 다당류, 상기 분자의 임의의 변형된 유도체, 또는 상기 분자 중 하나 이상을 포함하는 임의의 복합체일 수 있다. 핵산은 DNA 및 RNA를 포함하고, 단일- 또는 이중-가닥일 수 있고; 선형, 분지형 또는 원형일 수 있으며; 임의의 길이를 가질 수 있다. 핵산은 이중가닥을 형성할 수 있는 것뿐만 아니라 삼중가닥-형성 핵산을 포함한다. 예를 들어, 미국 특허 제5,176,996호 및 제5,422,251호 참조. 단백질은 DNA-결합 단백질, 전사 인자, 염색질 리모델링 인자, 메틸화된 DNA 결합 단백질, 중합효소, 메틸라제, 데메틸라제, 아세틸라제, 데아세틸라제, 키나제, 포스파타제, 인테그라제, 재조합효소, 리가제, 토포아이소머라제, 자이라제 및 헬리카제를 포함하지만, 이들로 제한되지 않는다.

- [0081] 외인성 분자는 내인성 분자와 동일한 유형의 분자, 예를 들어, 외인성 단백질 또는 핵산일 수 있다. 예를 들어, 외인성 핵산은 세포 내로 도입된 감염성 바이러스 게놈, 플라스미드 또는 에피솜, 또는 세포 내에 정상적으로 존재하지 않는 염색체를 포함할 수 있다. 세포 내로 외인성 분자의 도입을 위한 방법은 당업자에게 공지되어 있으며, 지질-매개 전달(즉, 중성 및 양이온성 지질을 포함하는 리포좀), 전기천공법, 직접 주사, 세포 융합, 유전자 총, 인산칼슘 공동침전, DEAE-덱스트란-매개 전달 및 바이러스 벡터-매개 전달을 포함하지만, 이들로 제한되지 않는다. 외인성 분자는 또한 내인성 분자와 동일한 유형일 수 있지만, 세포와 상이한 종으로부터 유래될 수 있다. 예를 들어, 인간 핵산 서열은 마우스 또는 햄스터로부터 유래된 세포주 내로 도입될 수 있다.
- [0082] 대조적으로, "내인성" 분자는 특히 특정 환경 조건 하에서 특정 발생 단계의 세포에 정상적으로 존재하는 것이다. 예를 들어, 내인성 핵산은 염색체, 미토콘드리아, 염록체 또는 다른 세포소기관의 게놈, 또는 천연 유래 에피솜 핵산을 포함할 수 있다. 추가적인 내인성 분자는 단백질, 예를 들어, 전사 인자 및 효소를 포함할 수 있다.
- [0083] 본 명세서에서 사용되는 용어 "외인성 핵산의 산물"은 폴리뉴클레오타이드와 폴리펩타이드 산물 둘 다, 예를 들어, 전사 산물(폴리뉴클레오타이드, 예컨대 RNA) 및 번역 산물(폴리펩타이드)을 포함한다.
- [0084] "융합" 분자는 2 이상의 서브유닛 분자가, 바람직하게는 공유적으로 연결되는 분자이다. 서브유닛 분자는 동일한 화학적 유형의 분자일 수 있거나, 또는 상이한 화학적 유형의 분자일 수 있다. 제1 유형의 융합 분자의 예는 융합 단백질 (예를 들어, ZFP 또는 TALE DNA-결합 도메인과 하나 이상의 활성화 도메인 사이의 융합) 및 융합 핵산(예를 들어, 상기 기재된 융합 단백질을 암호화하는 핵산)을 포함하지만, 이들로 제한되지 않는다. 제2 유형의 융합 분자의 예는 삼중가닥-형성 핵산과 폴리펩타이드 사이의 융합, 및 소홈(minor groove) 결합제와 핵산 사이의 융합을 포함하지만, 이들로 제한되지 않는다.
- [0085] 세포 내 융합 단백질의 발현은 세포에 대한 융합 단백질의 전달로부터 또는 융합 단백질을 암호화하는 폴리뉴클레오타이드의 세포에 대한 전달에 의해 초래될 수 있되, 폴리뉴클레오타이드는 전사되고, 전사체는 번역되어 융합 단백질을 생성한다. 트랜스-스플라이싱, 폴리펩타이드 절단 및 폴리펩타이드 결합은 또한 세포 내 단백질의 발현에 연루될 수 있다. 세포에 대한 폴리뉴클레오타이드 및 폴리펩타이드 전달을 위한 방법은 본 개시내용의 다른 곳에 제시된다.
- [0086] 본 개시내용의 목적을 위한 "유전자"는, 조절 서열이 암호 서열, 전사된 서열 또는 암호 서열과 전사된 서열의 조합에 인접하든, 아니든, 유전자 산물을 암호화하는 DNA 영역(이하 참조)뿐만 아니라 유전자 산물의 생성을 조절하는 모든 DNA 영역을 포함한다. 따라서, 유전자는 프로모터 서열, 종결자, 번역 조절 서열, 예컨대 리보솜 결합 부위 및 내부 리보솜 유입 부위, 인핸서, 사일런서, 절연체, 경계 요소, 복제 기점, 기질 부착 부위 및 좌우 제어 영역을 포함하지만, 이들로 반드시 제한되지는 않는다.
- [0087] "유전자 발현"은 유전자에 함유된 정보의 유전자 산물로의 전환을 지칭한다. 유전자 산물은 유전자의 직접 전사 산물(예를 들어, mRNA, tRNA, rRNA, 안티센스 RNA, 마이크로RNA, 라이보자임, 구조적 RNA 또는 임의의 다른 유형의 RNA) 또는 mRNA의 번역에 의해 생성되는 단백질일 수 있다. 유전자 산물은 또한 캡핑, 폴리아데닐화, 메틸화 및 편집과 같은 가공에 의해 변형되는 RNA, 및, 예를 들어, 메틸화, 아세틸화, 포스포릴화, 유비퀴틴화, ADP-리보실화, 미리스틸화 및 글리코실화에 의해 변형되는 단백질을 포함한다.
- [0088] 유전자 발현의 "조절"은 유전자 활성의 변화를 지칭한다. 발현의 조절은 유전자 활성화 및 유전자 억제를 포함할 수 있지만, 이들로 제한되지 않는다. 게놈 편집(예를 들어, 절단, 변형, 비활성화, 무작위 돌연변이)을 사용하여 발현을 조절할 수 있다. 유전자 비활성화는 본 명세서에 기재된 바와 같은 ZFP, TALE 또는 CRISPR/Cas 시스템을 포함하지 않는 세포에 의해 유전자 발현의 임의의 감소를 지칭한다. 따라서, 유전자 비활성화는 부분적이거나 또는 완전할 수 있다.
- [0089] "보호된" mRNA는 mRNA의 안정성 또는 번역을 증가시키는 일부 방식에서 mRNA가 변경된 것이다. 보호의 예는 사이티딘 및 유리딘 잔기의 25%까지의 2-티오유리딘(s2U) 및 5-메틸사이티딘(m5C)으로의 대체의 사용을 포함한다. 얻어진 mRNA는 그의 비변형 상태에 비해 덜 번역원성 그리고 더 안정성을 나타낸다. (문헌[Kariko *et al.* ((2012), *Molecular Therapy*, Vol. 16, No. 11, pages 1833 - 1844] 참조). 다른 변화는 시험관내 생산된 mRNA의 번역 능력을 증가시키는 소위 ARCA 캡의 첨가를 포함한다(미국 특허 제7,074,596호 참조).
- [0090] "관심 대상의 영역"은 세포의 염색질의 임의의 영역, 예를 들어, 외인성 분자에 결합하는 데 바람직한 유전자 내의 또는 유전자에 인접한 유전자 또는 비-암호 서열이다. 결합은 표적화된 DNA 절단, 표적화된 재조합 및 표적화된 DNA 절단과 표적화된 재조합의 조합의 목적을 위한 것일 수 있다. 관심 대상의 영역은, 예를 들어, 염색

체, 에피솜, 세포소기관 계놈(예를 들어, 미토콘드리아, 엽록소) 또는 감염성 바이러스 계놈에 존재할 수 있다. 관심 대상의 영역은 유전자의 암호 영역 내이거나, 전사된 비-암호 영역, 예를 들어, 리더 서열, 비번역 순서(trailer sequence) 또는 인트론 내이거나, 또는 암호 영역의 상류 또는 하류 중 하나에서, 비전사 영역 내일 수 있다. 관심 대상의 영역은 길이가 단일 뉴클레오타이드 쌍 또는 2,000개까지의 뉴클레오타이드쌍, 또는 뉴클레오타이드 쌍의 임의의 정수값만큼 작을 수 있다.

[0091] "진핵" 세포는, 진균 세포(예컨대 효모), 식물 세포, 동물 세포, 포유류 세포 및 인간 세포(예를 들어, T-세포)를 포함하지만, 이들로 제한되지 않는다.

[0092] 용어 "작업적 연결" 및 "작동 가능하게 연결된"(또는 "작동적으로 연결된")은 둘 이상의 성분(예컨대, 서열 요소)의 병치에 대해 상호 호환 가능하게 사용되며, 이때 성분은 성분이 둘 다 정상적으로 기능하고 성분 중 적어도 하나가 다른 성분 중 적어도 하나에 대해 발휘되는 기능을 매개할 수 있는 가능성을 허용하도록 배열된다. 예로서, 전사 조절 서열이 하나 이상의 전사 조절 인자의 존재 또는 부재에 반응하여 암호 서열의 전사 수준을 제어한다면, 전사 조절 서열, 예컨대 프로모터는 암호 서열에 작동 가능하게 연결된다. 전사 조절 서열은 일바적으로 암소 서열과 시스로 작동 가능하게 연결되지만, 그것에 직접적으로 인접할 필요는 없다. 예를 들어, 인핸서는, 인접하지 않는다고 해도, 암호 서열에 작동 가능하게 연결된 전사 조절 서열이다.

[0093] 융합 폴리펩타이드에 대해, 용어 "작동 가능하게 연결된"은 각각의 성분이, 그것이 연결되지 않았다면, 다른 성분과의 연결에서 동일한 기능을 수행한다는 사실을 지칭할 수 있다. 예를 들어, ZFP, TALE 또는 Cas DNA-결합 도메인이 활성화 도메인에 융합된 융합 폴리펩타이드에 대해, ZFP, TALE 또는 Cas DNA-결합 도메인 및 활성화 도메인은 융합 폴리펩타이드에서, Cas DNA-결합 도메인 일부의 ZFP, TALE가 그의 표적 부위, 그의 결합 부위, 및 그의 표적 부위와 결합 부위의 조합에 결합할 수 있다면, 작동성 연결인 반면, 활성화 도메인은 유전자 발현을 상향조절할 수 있다. ZFP, TALE 또는 Cas DNA-결합 도메인이 절단 도메인에 융합되는 융합 폴리펩타이드일 때, ZFP, TALE 또는 Cas DNA-결합 도메인 및 절단 도메인은, 융합 폴리펩타이드에서, ZFP, TALE 또는 Cas DNA-결합 도메인 부분이 그의 표적 부위, 그의 결합 부위, 및 그의 표적 부위와 그의 결합 부위의 조합에 결합할 수 있다면, 작동성 연결인 반면, 절단 도메인은 표적 부위 부근에서 DNA를 절단할 수 있다.

[0094] 단백질, 폴리펩타이드 또는 핵산의 "기능성 단편"은 서열이 전장 단백질, 폴리펩타이드 또는 핵산과 동일하지는 않지만, 아직 전장 단백질, 폴리펩타이드 또는 핵산과 동일한 기능을 보유하는 단백질, 폴리펩타이드 또는 핵산이다. 기능성 단편은 대응하는 천연 분자보다 더 많거나, 더 적거나 또는 동일한 수의 잔기를 가질 수 있고, 하나 이상의 아미노산 또는 뉴클레오타이드 치환을 함유할 수 있으며, 대응하는 천연 분자보다 더 많거나, 더 적거나 또는 동일한 수의 잔기를 갖고, 하나 이상의 아미노산 또는 뉴클레오타이드 치환을 함유하는 조합일 수 있다. 핵산의 기능(예를 들어, 암호 기능, 다른 핵산과 혼성화하는 능력)을 결정하는 방법은 당업계에 잘 공지되어 있다. 유사하게, 단백질 기능을 결정하는 방법은 잘 공지되어 있다. 예를 들어, 폴리펩타이드의 DNA-결합 기능은, 예를 들어, 필터-결합, 전기영동 이동도, 또는 면역침전 분석에 의해 결정될 수 있다. DNA 절단은 젤 전기영동에 의해 분석될 수 있다. 문헌[Ausubel et al., 상기 참조]. 다른 단백질과 상호작용하는 단백질의 능력은, 예를 들어, 공동면역침전, 2-혼성체 분석 또는 상보성에 의해, 유전적 그리고 생화학적으로 결정될 수 있다.

[0095] "벡터"는 표적 세포에 유전자 서열을 전달할 수 있다. 전형적으로, "벡터 작제물", "발현 벡터" 및 "유전자 전달 벡터"는 관심 대사의 유전자 발현을 지시할 수 있고 표적 세포에 유전자 서열을 전달할 수 있는 임의의 핵산 작제물을 의미한다. 따라서, 상기 용어는 클로닝 및 발현 비히클뿐만 아니라 완전한 벡터를 포함한다.

[0096] 상기 용어 "대상체" 및 "환자"는 상호 호환적으로 사용되며, 포유류, 예컨대 인간 환자 및 비인간 영장류뿐만 아니라 실험 동물, 예컨대 돼지, 소, 토끼, 개, 고양이, 래트, 마우스 및 다른 동물을 지칭한다. 따라서, 본 명세서에서 사용되는 용어 "대상체" 또는 "환자"는 본 발명의 줄기 세포가 투여될 수 있는 임의의 포유류 환자 또는 대상체를 의미한다. 본 발명의 대상체는, 예를 들어, 신경 독소를 포함하는 1종 이상의 화학적 독소에 노출된 것을 포함한다.

[0097] "줄기능"은 줄기 세포-유사 방식, 즉, 전형성능, 전분화능, 다능성 또는 소분화능(oligo-potency) 정도 및 임의의 특정 줄기 세포가 가질 수 있는 확장된 또는 무한한 자기 재생으로 작용하는 임의의 세포의 상대적 능력을 지칭한다.

뉴클레아제

[0099] 본 명세서에서 이식유전자가 표적화된 방식으로 계놈 내로 통합되도록 세포 계놈의 절단을 위해 이식유전자 및

뉴클레아제를 운반하는 공여자 분자의 생체내 절단에 유용한 조성물, 특히 뉴클레아제가 기재된다. 특정 실시형태에서, 뉴클레아제 중 하나 이상은 천연 유래이다. 다른 실시형태에서, 뉴클레아제 중 하나 이상은 비천연 유래이며, 즉, DNA-결합 도메인, 절단 도메인, 및 DNA-결합 도메인과 절단 도메인의 조합에서 조작된다. 예를 들어, 천연 유래 뉴클레아제의 DNA-결합 도메인은 선택된 표적 부위에 결합하도록 변경될 수 있다(예를 들어, 동족 결합 부위와 상이한 부위에 결합하도록 조작된 메가뉴클레아제). 다른 실시형태에서, 뉴클레아제는 이종성 DNA-결합 및 절단 도메인(예를 들어, 아연 평거 뉴클레아제; TAL-효과기 도메인 DNA 결합 단백질; 이종성 절단 도메인을 갖는 메가뉴클레아제 DNA-결합 도메인)을 포함한다.

A. DNA-결합 도메인

특정 실시형태에서, 생체내 절단, 세포 계놈의 표적화된 절단, 및 생체내 절단과 세포 계놈의 표적화된 절단의 조합을 위해 사용되는 하나 이상의 뉴클레아제의 DNA 결합 도메인은 아연 평거 단백질을 포함한다. 단일 아연 평거 단백질은 다중 아연 평거 도메인(예를 들어, 3, 4, 5, 6개 이상의 아연 평거 도메인)으로 이루어진다. 각각의 아연 평거 도메인은 표적 서열에 대한 단백질의 결합을 허용하는 특정 입체배좌로 보유되는 베타 회전을 함유하는(2개의 아연 배위 잔기 및 알파 나선을 함유(2개의 비변이체 아연 배위 잔기를 함유)) 길이가 약 30개의 아미노산이다. 베타 회전에서 2개의 시스테인(Cys) 아연 배위 잔기 및 알파 나선에서 2개의 히스티딘(His) 아연 배위 잔기를 갖는 정규(C2H2) 아연 평거 도메인 또는 또는 비정규(CH3)가 사용될 수 있다. 예를 들어, 미국 특허 제9,234,187호 참조. 7-아미노산 인식 나선은 베타 회전의 아연 배위 잔기와 알파 나선의 아연 배위 잔기 사이에 함유된다. 인식 나선 영역은 이 인식 영역 외부의 아연 평거 도메인 및 아미노산 내에서 -1 내지 +6으로 넘버링된다(그리고 아연 배위 잔기를 제외한 것은 골격 잔기로서 지칭된다).

바람직하게는, 아연 평거 단백질은 인식 나선이 선택의 표적 부위에 결합하도록 조작되는 비천연 유래이다. 예를 들어, 문헌[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 제6,453,242호; 제6,534,261호; 제6,599,692호; 제6,503,717호; 제6,689,558호; 제7,030,215호; 제6,794,136호; 제7,067,317호; 제7,262,054호; 제7,070,934호; 제7,361,635호; 제7,253,273호; 및 미국 특허 공개 제2005/0064474호; 제2007/0218528호; 제2005/0267061호를 참조하며, 이들 모두는 본 명세서에 전문이 참고로 포함된다.

조작된 아연 평거 결합 도메인은 천연 유래 아연 평거 단백질에 비해 신규한 결합 특이성을 가질 수 있다. 조작 방법은 합리적 설계 및 다양한 유형의 선택을 포함하지만, 이들로 제한되지 않는다. 합리적 설계는, 예를 들어, 삼중(또는 사중) 뉴클레오타이드 서열 및 개개의 아연 평거 아미노산 서열을 포함하는 데이터베이스를 이용하는 것을 포함하며, 이때 각각의 삼중 또는 사중 뉴클레오타이드 서열은 특정 삼중 또는 사중 서열에 결합하는 아연 평거의 하나 이상의 아미노산 서열과 관련된다. 예를 들어, 본 명세서에 전문이 참고로 포함된 공동 소유된 미국 특허 제6,453,242호 및 제6,534,261호 참조.

더 나아가, 특정 실시형태에서, ZFP DNA-결합 도메인은 아연 평거 도메인의 골격에 대한 하나 이상의 변형을 추가로 포함한다. 표적 DNA 서열에 대한 ZFP의 특이성은 아연 평거 도메인과 특정 DNA 염기 사이의, 특히, 인식 나선 영역과 표적 부위 사이의 서열 특이적 접촉에 의존한다(전형적으로 각각의 인식 나선은 3개의 뉴클레오타이드의 표적 하부부위에 결합한다). 추가로, 아연 평거 도메인은 또한 DNA 골격의 인산염과의 비특이적 상호작용에 참여하는 아미노산 잔기를 포함한다. 문헌[Elrod-Erickson *et al.* ((1996) *Structure* 4:1171)]은 아연 평거 단백질 및 그의 동족 DNA 표적의 공동결정화를 통해 수소 결합의 형성을 통해 DNA 골격 상의 인산염과 상호작용할 수 있는 특정 아미노산이 있다는 것을 입증하였다. 잘 공지된 Zif268 백본을 사용하는 아연 평거 단백질은 전형적으로 또한 제2 변이체 시스테인에 대해 제2 위치 카복실-말단인 β -시트의 그들의 제2 가닥의 아미노산 말단 잔기로서 알기닌을 가진다. 이 위치는 그것이 α -나선의 시작에 선행하는 5번째 잔기(그리고 인식 나선 넘버링 -1 내지 +6에 대해 위치 -5)이기 때문에 각각의 아연 평거 도메인 내에서 (-5)로서 지칭될 수 있다. 이 위치에서 알기닌은 그의 측쇄 구아니디늄기와의 변화된 수소 결합 형성을 통해 DNA 골격 상에서 인산염과 상호작용할 수 있다. Zif268 골격 내 아연 평거 단백질은 또한 제1 비변이체 시스테인에 대해 아미노-말단에 4개의 잔기가 있는 위치에서 라이신을 빈번하게 가진다. 이 위치는 아연 배위 시스테인 잔기(및 인식 나선 영역 넘버링 -1 내지 +6에 대해 위치 -14) 사이의 두 잔기를 갖는 아연 평거에 대해 α -나선의 시작에 선행하는 14번째 잔기이기 때문에, 각각의 평거 내에서 (-14)로서 지칭될 수 있다. 라이신은 그의 측쇄 아미노기와 물-매개 하전된 수소 결합의 형성을 통해 DNA 골격 상에서 인산염과 상호작용할 수 있다. 인산염기는 모두 DNA 골격을 따라서 발견되기 때문에, 아연 평거와 DNA 분자 사이의 이런 유형의 상호작용은 일반적으로 비서열 특이적이 되는

것으로 고려된다(J. Miller, Massachusetts Institute of Technology Ph.D. Thesis, 2002).

[0105] 최근의 연구는 일부 뉴클레아제에서 비특이적 인산염 접촉 측쇄가 해당 뉴클레아제의 소량의 비특이성을 설명할 수 있다는 가설을 세웠다(Kleinstiver *et al.* (2016) *Nature* 529(7587):490-5; Guilinger *et al.* (2014) *Nat Meth*: 429-435). 연구자들은 이를 뉴클레아제는 '과량의 DNA-결합 에너지'를 가질 수 있다는 것을 제안하였는데, 이는 뉴클레아제가 표적 부위에 설질적으로 결합하고 이를 절단하는 데 필요한 것보다 그들의 DNA 표적에 대해 더 큰 친화도를 가질 수 있다는 것을 의미한다. 따라서, TALE DNA 결합 도메인(Guilinger, 이하 참조) 또는 Cas9 DNA 결합 도메인(Kleinstiver, 이하 참조)에서 양이온성 전하를 이들 뉴클레아제의 DNA-결합 에너지보다 더 낮게 감소시키기 위한 시도가 있었는데, 이는 시험판내 증가된 절단 특이성을 초래한다. 그러나, 추가적인 연구(Sternberg *et al.* (2015) *Nature* 527(7576):110-113)는 또한 Cas9 DNA 결합 도메인의 클레인스티버(Kleinstiver) 연구에서 돌연변이된 일부 양이온성 아미노산에 대해 Cas9 뉴클레아제 도메인의 적절한 풀딩 및 활성화에서의 역할을 시사한다. 따라서, Cas9 활성에서 이들 아미노산의 정확한 역할은 알려져 있지 않다.

[0106] 따라서 본 발명의 방법 및 조성물은 DNA 골격 상에서 인산염과 비특이적으로 상호작용할 수 있는 ZFP DNA 결합 도메인('ZFP 골격') 내의 아미노산에 대한 돌연변이를 포함하지만, 그들은 DNA 인식 나선에서 변화를 포함하지 않는다. 따라서, 본 발명은 뉴클레오타이드 표적 특이성에 필요하지 않은 ZFP 골격에서의 양이온성 아미노산 잔기의 돌연변이를 포함한다. 일부 실시형태에서, ZFP 골격에서의 이들 돌연변이는 양이온성 아미노산 잔기의 중성 또는 음이온성 아미노산 잔기로의 돌연변이를 포함한다. 일부 실시형태에서, ZFP 골격에서의 이들 돌연변이는 극성 아미노산 잔기의 중성 또는 비극성 아미노산 잔기로의 돌연변이를 포함한다. 바람직한 실시형태에서, 위치(-5), 위치(-9), 위치(-14)에서 이루어진 돌연변이, 및 DNA 결합 나선에 대해 위치(-5), 위치(-9)와 위치(-14)에서 이루어진 돌연변이로부터 선택된 돌연변이의 조합을 포함한다. 일부 실시형태에서, 아연 평거는 (-5), (-9), (-14)에서의 하나 이상의 돌연변이 및 (-5), (-9), 및 (-14)에서의 돌연변이로부터 선택된 돌연변이의 조합을 포함할 수 있다. 추가 실시형태에서, 다중 평거 아연 평거 단백질에서 하나 이상의 아연 평거(들)는 (-5), (-9), (-14)의 돌연변이 및 (-5), (-9) 및 (-14)로부터 선택되는 조합을 포함할 수 있다. 일부 실시형태에서, (-5), (-9), (-14)에서의 아미노산 및 (-5), (-9) 및 (-14)(예를 들어, 알기닌(R) 또는 라이신(K))으로부터 선택된 조합은 알라닌(A), 류신(L), Ser(S), Asp(D), Glu(E), Tyr(Y) 및/또는 글루타민(Q)으로 돌연변이된다.

[0107] 본 명세서에 기재된 임의의 이들 융합 폴리펩타이드에서, ZFP 상대는 (-5), (-9), (-14) 위치에서 아연 평거 DNA 결합 도메인의 돌연변이, 및 (-5), (-9) 및 (-14)에서의 돌연변이로부터 선택된 돌연변이의 조합을 추가로 포함할 수 있다. 일부 실시형태에서, 위치 -5에서 Arg(R)은 Tyr(Y), Asp(D), Glu(E), Leu(L), Gln(Q) 또는 Ala(A)으로 변화된다. 다른 실시형태에서, 위치 (-9)에서 Arg(R)은 Ser(S), Asp(D), 또는 Glu(E)으로 대체된다. 추가 실시형태에서, 위치(-14)에서 Arg(R)은 Ser(S) 또는 Gln(Q)로 대체된다. 다른 실시형태에서, 융합 폴리펩타이드는 아연 평거 DNA 결합 도메인에서 돌연변이를 포함할 수 있으며, 여기서, (-5), (-9), (-14) 위치에서의 아미노산, 및 (-5), (-9) 및 (-14)에서의 돌연변이로부터 선택되는 돌연변이의 조합은 임의의 조합으로 임의의 상기 열거한 아미노산으로 변화된다.

[0108] 과지 디스플레이 및 2-흔성체 시스템을 포함하는 예시적인 선택 방법은 미국 특허 제5,789,538호; 제5,925,523호; 제6,007,988호; 제6,013,453호; 제6,410,248호; 제6,140,466호; 제6,200,759호; 및 제6,242,568호; 뿐만 아니라 WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 및 GB 2,338,237에 개시되어 있다. 추가로, 아연 평거 결합 도메인에 대한 결합 특이성의 향상은, 예를 들어, 공동 소유된 WO 02/077227에 기재되어 있다.

[0109] 표적 부위의 선택; 융합 단백질(및 이를 암호화하는 폴리뉴클레오타이드)의 설계 및 작제를 위한 ZFP 및 방법은 당업자에게 공지되어 있으며, 미국 특허 제6,140,081호; 제5,789,538호; 제6,453,242호; 제6,534,261호; 제5,925,523호; 제6,007,988호; 제6,013,453호; 제6,200,759호; 국제 특허 출원 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 03/016496에 상세하게 기재되어 있다.

[0110] 거의 임의의 링커(스페이서)는 DNA-결합 도메인 성분 중 하나 이상(예를 들어, 아연 평거) 사이에서, 하나 이상의 DNA-결합 도메인 사이에서, DNA-결합 도메인과 기능성 도메인(예를 들어, 뉴클레아제) 사이에서, 그리고 하나 이상의 DNA-결합 도메인 사이에서 그리고 DNA-결합 도메인과 기능성 도메인 사이에서 사용될 수 있다. 적합한 링커 서열의 비제한적 예는 미국 특허 제8,772,453호; 제7,888,121호; 제6,479,626호; 제6,903,185호; 및 제7,153,949호; 및 미국 특허 공개 제2009/0305419호; 제2015/0064789호 및 제2015/0132269호를 포함한다. 따라서, 본 명세서에 기재된 단백질은 개개의 DNA-결합 성분 사이, DNA-결합 도메인과 기능성 도메인 사이, 또는

하나 이상의 DNA-결합 도메인 사이 및 본 명세서에 기재된 조성물의 DNA-결합 도메인과 기능성 도메인 사이에서 적합한 링커의 임의의 조합을 포함할 수 있다.

[0111] B. 절단 도메인

[0112] 임의의 적합한 절단 도메인은 본 명세서에 기재된 DNA-결합 도메인에 작동 가능하게 연결되어 뉴클레아제를 형성할 수 있다. 절단 도메인은 DNA-결합 도메인, 예를 들어 아연 평거 DNA-결합 도메인 및 뉴클레아제로부터의 절단 도메인에 대해 이종성일 수 있다. 이종성 절단 도메인은 임의의 엔도뉴클레아제 또는 엑소뉴클레아제로부터 얻을 수 있다. 절단 도메인이 유래될 수 있는 예시적인 엔도뉴클레아제는, 제한 엔도뉴클레아제 및 귀소 엔도뉴클레아제를 포함하지만, 이들로 제한되지 않는다. 예를 들어, 문헌[2002-2003 Catalogue, New England Biolabs, Beverly, MA; 및 Belfort *et al.* (1997) Nucleic Acids Res. 25:3379-3388] 참조. DNA를 절단하는 추가적인 효소는 공지되어 있다(예를 들어, S1 뉴클레아제; 녹두 뉴클레아제; 췌장 DNase I; 마이크로코커스 뉴클레아제; 효모 HO 엔도뉴클레아제; 또한 문헌[Linn *et al.* (eds.) Nucleases, Cold Spring Harbor Laboratory Press, 1993] 참조). 이들 효소(또는 이의 기능성 단편) 중 하나 이상은 절단 도메인 및 절단 절반-도메인의 공급원으로서 사용될 수 있다.

[0113] 유사하게, 절단 절반-도메인은 절단 활성의 이량체화를 필요로 하는 상기 제시한 바와 같은 임의의 뉴클레아제 또는 이의 일부로부터 유래될 수 있다. 일반적으로, 융합 단백질이 절단 절반-도메인을 포함한다면, 절단을 위해 2개의 융합 단백질이 필요하다. 대안적으로, 2개의 절단 절반-도메인을 포함하는 단일 단백질이 사용될 수 있다. 2개의 절단 절반-도메인은 동일한 엔도뉴클레아제(또는 이의 기능성 단편)로부터 유래될 수 있거나, 또는 각각의 절단 절반-도메인은 상이한 엔도뉴클레아제(또는 이의 기능성 단편)로부터 유래될 수 있다. 추가로, 2개의 융합 단백질의 그들의 각각의 표적 부위에 대한 결합이 절단 절반-도메인이, 예를 들어, 이량체화에 의해 절단 절반-도메인이 기능성 절단 도메인을 형성하게 허용하는 서로에 대한 공간적 배향으로 절단 절반 도메인을 위치시키도록, 2개의 융합 단백질에 대한 표적 부위는 서로에 대해 바람직하게 배치된다. 따라서, 특정 실시형태에서, 표적 부위 근처의 가장자리는 5 내지 8개의 뉴클레오타이드에 의해 또는 15 내지 18개의 뉴클레오타이드에 의해 분리된다. 그러나, 뉴클레오타이드 또는 뉴클레오타이드 쌍의 임의의 필수적인 수는 2개의 표적 부위(예를 들어, 2 내지 50개 이상의 뉴클레오타이드)를 개재할 수 있다. 일반적으로, 절단 부위는 표적 부위 사이에 놓인다.

[0114] 제한 엔도뉴클레아제(제한 효소)는 다수의 종에 존재하며, (인식 부위에서) DNA에 서열-특이적으로 결합하고, 결합 부위에서 또는 근처에서 DNA를 절단할 수 있다. 특정 제한 효소(예를 들어, IIS형)는 인식 부위로부터 제거된 부위에서 DNA를 절단하고, 분리 가능한 결합 및 절단 도메인을 가진다. 예를 들어, IIS형 효소 Fok I은 하나의 가닥 상의 그의 인식 부위로부터 9개 뉴클레오타이드에서 그리고 다른 가닥 상의 그의 인식 부위로부터 13개의 뉴클레오타이드에서 DNA의 이중 가닥 절단을 촉매한다. 예를 들어, 미국 특허 제5,356,802호; 제5,436,150호 및 제5,487,994호; 뿐만 아니라 문헌[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. 269:31,978-31,982] 참조. 따라서, 일 실시형태에서, 융합 단백질은 적어도 1종의 IIS형 제한 효소 및 하나 이상의 아연 평거 결합 도메인에서 절단 도메인(또는 절단 절반-도메인)을 포함하는데, 이는 조작될 수도 있고 또는 조작되지 않을 수도 있다.

[0115] 서열-선택적(인공) 뉴클레아제에 의한 최적의 절단 특이성을 위해, 표적 상의 결합 및 활성이 포화되지 않도록 조건을 배열하는 것이 바람직하다. 포화 조건 하에서 - 정의에 의해 - 완전한 표적 상의 활성을 달성하는 데 필수적인 것 이상으로 과량의 뉴클레아제가 사용된다. 이 과량은 표적 상의 이점을 제공하지 않지만, 그럼에도 불구하고 표적을 벗어난 부위에서 증가된 절단을 야기할 수 있다. 단량체 뉴클레아제에 대해, 적정 곡선 상의 포화 안정기를 동정하고 회피하기 위한 단순한 용량 반응 연구를 수행함으로써 포화 조건은 용이하게 회피될 수 있다. 그러나, 이량체 뉴클레아제, 예컨대 ZFN, TALEN 또는 dCas-Fok에 대해, 개개의 단량체의 결합 친화도가 유사하지 않다면, 포화 조건을 동정하고 회피하는 것은 더 복잡하게 될 수 있다. 이러한 경우에, 단순한 1:1 뉴클레아제 비를 이용하는 용량 반응 연구는 단지 더 약한 결합 단량체의 포화 지점을 나타낼 것이다. 이러한 시나리오 하에서, 예를 들어, 단량체 친화도가 10배만큼 상이하다면, 1:1 적정 연구에서 동정된 포화 지점에서 더 높은 친화도의 단량체는 그것이 필요로 하는 것보다 10배 더 높은 농도에서 존재한다. 얻어진 과량의 더 높은 친화도 단량체는 결국 의도된 표적에서의 임의의 유리한 증가를 제공하는 일 없이 증가된 표적을 벗어난 활성을 야기하여, 잠재적으로 임의의 주어진 뉴클레아제 쌍에 대해 전반적으로 감소된 특이성을 야기할 수 있다.

[0116] 절단 도메인이 결합 도메인으로부터 분리 가능한 예시적인 IIS형 제한 효소는 FokI이다. 이 특정 효소는 이량체

로서 활성이다. 문헌[Bitinaite *et al.* (1998) Proc. Natl. Acad. Sci. USA 95:10,570-10,575]. 따라서, 본 개시내용의 목적을 위해, 개시된 융합 단백질에서 사용되는 *Fok I* 효소의 일부는 절단 절반-도메인으로 고려된다. 따라서, 표적화된 이중-가닥 절단, 아연 평거-*Fok I* 융합을 이용하는 세포 서열의 표적화된 대체, 및 표적화된 이중-가닥 절단과 아연 평거-*Fok I* 융합을 이용하는 세포 서열의 표적화된 대체의 조합을 위해, 각각 *FokI* 절단 절반-도메인을 포함하는 2개의 융합 단백질은 촉매적으로 활성인 절단 도메인을 재구성하는 데 사용될 수 있다. 대안적으로, 아연 평거 결합 도메인 및 2개의 *Fok I* 절단 절반-도메인을 함유하는 단일 폴리펩타이드 분자가 또한 사용될 수 있다. 아연 평거-*Fok I* 융합을 이용하는 표적화된 절단 및 표적화된 서열 변경을 위한 매개변수는 본 개시내용의 다른 곳에서 제공된다.

[0117] 절단 도메인 또는 절단 절반-도메인은 절단 활성을 보유하거나, 또는 기능성 절단 도메인을 형성하기 위해 다양체화(예를 들어, 이량체화)하는 능력을 보유하는 단백질의 임의의 부분일 수 있다.

[0118] 예시적인 IIS형 제한 효소는 본 명세서에 전문이 포함되는 미국 특허 제7,888,121호에 기재되어 있다. 추가적인 제한 효소는 또한 분리 가능한 결합 및 절단 도메인을 함유하고, 이들은 본 개시내용에 의해 상정된다. 예를 들어, 문헌[Roberts *et al.* (2003) *Nucleic Acids Res.* 31:418-420] 참조.

[0119] 특정 실시형태에서, 절단 도메인은, 예를 들어, 미국 특허 제7,914,796호; 제8,034,598호 및 제8,623,618호(이들 모두의 개시내용은 그들의 전문이 본 명세서에 참고로 포함됨)에 기재된 바와 같이, 동종이량체화를 최소화하거나 또는 방지하는 하나 이상의 조작된 절단-절반 도메인)을 포함한다. *FokI*의 위치 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537 및 538에서 아미노산 잔기는 *FokI* 절단 절반-도메인의 이량체화에 영향을 미치기 위한 모든 표적이며, 여기서 넘버링은 이하에 나타내는 서열을 갖는 결정구조 1FOK.pdb 및 2FOK.pdb에 대한 것이다(문헌[Wah *et al.* (1997) *Nature* 388:97-100] 참조):

[0120] 야생형 *FokI* 절단 절반 도메인(서열번호 18)

QLVKSELEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVY
GYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRY
VEENQTRNKHINPNEWKVPSSVTEFKFLVSGHFKGNYKAQLTRLNHTIN
CNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF

[0121] 절대 이형이량체를 형성하는 *FokI*의 예시적인 조작된 절단 절반-도메인은 제1 절단 절반-도메인이 *FokI*의 위치 490 및 538의 아미노산 잔기에서 돌연변이를 포함하고 제2 절단 절반-도메인이 아미노산 잔기 486 및 499의 아미노산 잔기에서 돌연변이를 포함하는 쌍을 포함한다.

[0122] 따라서, 일 실시형태에서, 490의 돌연변이는 Glu(E)을 Lys(K)으로 대체하며; 538에서의 돌연변이는 Ile(I)를 Lys(K)으로 대체하고; 486에서의 돌연변이는 Gln(Q)를 Glu(E)로 대체하며; 그리고 위치 499에서의 돌연변이는 Ile(I)를 Lys(K)로 대체한다. 구체적으로, 본 명세서에 기재된 조작된 절단 절반-도메인은 "E490K:I538K"로 표기되는 조작된 절단 절반-도메인을 생성하기 위해 하나의 절단 절반-도메인에서 위치 490(E→K) 및 538(I→K)를 돌연변이시킴으로써 그리고 "Q486E:I499L"로 표기되는 조작된 절단 절반-도메인을 생성하기 위해 다른 절단 절반-도메인에서 위치 486 (Q→E) 및 499(I→L)를 돌연변이시킴으로써 제조되었다. 본 명세서에 기재된 조작된 절단 절반-도메인은 ZFN 동종이량체를 통한 비정상적 절단이 최소화되거나 또는 없어지는 절대 이형이량체 돌연변이체이다. 예를 들어, 미국 특허 공개 제2008/0131962호를 참조하며, 이의 개시내용은 본 명세서에 모든 목적을 위해 전문이 참고로 포함된다. 특정 실시형태에서, 조작된 절단 절반-도메인은 위치 486, 499 및 496(야생형 *FokI*에 대해 넘버링됨)에서 돌연변이, 예를 들어 위치 486에서 야생형 Gln(Q) 잔기를 Glu(E) 잔기로, 위치 499에서 야생형 Ile(I) 잔기를 Leu(L) 잔기로 그리고 위치 496에서 야생형 Asn(N) 잔기를 Asp(D) 또는 Glu(E) 잔기로 대체하는 돌연변이(또한 각각 "ELD" 및 "ELE" 도메인으로 지칭됨)를 포함한다. 다른 실시형태에서, 조작된 절단 절반-도메인은 위치 490, 538 및 537에서의 돌연변이(야생형 *FokI*에 대해 넘버링됨), 예를 들어 위치 490에서 야생형 Glu(E) 잔기를 Lys(K) 잔기로, 위치 538에서 야생형 Ile(I) 잔기를 Lys(K) 잔기로, 그리고 위치 537에서 야생형 His(H) 잔기를 Lys(K) 잔기 또는 Arg(R) 잔기로 대체하는 돌연변이(또한 각각 "KKK" 및 "KKR" 도메인으로 지칭됨)를 포함한다. 다른 실시형태에서, 조작된 절단 절반-도메인은 위치 490 및 537에서의 돌연변이(야생형 *FokI*에 대해 넘버링됨), 예를 들어 위치 490에서 야생형 Glu(E) 잔기를 Lys(K) 잔기로 그리고 위치 537에서 야생형 His(H) 잔기를 Lys(K) 잔기 또는 Arg(R) 잔기로 대체하는 돌연변이(또한 각각 "KIK" 및 "KIR" 도메인으로서 지칭됨)를 포함한다. 예를 들어, 미국 특허 제7,914,796호; 제8,034,598호 및 제8,623,618호 참조. 다른 실시형태에서, 조작된 절단 절반 도메인은 "샤키(Sharkey)", "샤키" 돌연변이 및 "샤키"와 "샤키"

돌연변이의 조합을 포함한다(문헌[Guo *et al.* (2010) *J. Mol. Biol.* 400(1):96-107] 참조).

[0124] 따라서, FokI로부터 유래된 절단 절반 도메인은 상기 기재한 바와 같은 이량체화 도메인 내 돌연변이; 촉매 도메인 내 돌연변이, 다른 아미노산 잔기, 예컨대 인산염 접촉 잔기 내 돌연변이, 및 이량체화 도메인 내 돌연변이, 촉매 도메인 내 돌연변이, 및 다른 아미노산 잔기, 예컨대 인산염 접촉 잔기 내 돌연변이로부터 선택되는 돌연변이의 임의의 조합을 포함하는 서열번호 18에 나타내는 바와 같은 아미노산 잔기 중 하나 이상에서의 돌연변이를 포함할 수 있다. 돌연변이는 (상이한 잔기를 갖는 야생형 아미노산 잔기의) 치환, (하나 이상의 아미노산 잔기의) 삽입, (하나 이상의 아미노산 잔기의) 결실, 및 치환, 삽입 및 결실로부터 선택되는 돌연변이의 임의의 조합을 포함한다. 특정 실시형태에서, 잔기 414-426, 443-450, 467-488, 501-502, 521-531 중 하나 이상(서열번호 18에 대해 넘버링됨), 및 이러한 잔기의 임의의 조합은 이들 잔기가 문헌[Miller *et al.* ((2007) *Nat Biotechnol* 25:778-784)]에 기재된 그의 표적 부위에 결합된 ZFN의 분자 모델에서 DNA 골격에 가깝게 위치되기 때문에, 돌연변이된다. 특정 실시형태에서, 위치 416, 422, 447, 448 및 525에서 하나 이상의 잔기가 돌연변이된다. 특정 실시형태에서, 돌연변이는 야생형 잔기의 상이한 잔기, 예를 들어 세린(S) 잔기로의 치환을 포함한다. 특정 실시형태에서, 본 명세서에 기재된 뉴클레아제의 FokI 절단 도메인은 ELD 이량체화 도메인 돌연변이, KKR 이량체화 도메인 돌연변이, K525S 돌연변이 및 ELD 이량체화 도메인 돌연변이 또는 KKR 이량체화 도메인 돌연변이 및 K525S 돌연변이로부터 선택되는 임의의 조합을 포함한다.

[0125] 본 명세서에 기재된 조작된 절단 도메인은 임의의 적합한 방법을 이용하여, 예를 들어, 미국 특허 제7,888,121호; 제7,914,796호; 제8,034,598호 및 제8,623,618호에 기재된 바와 같은 야생형 절단 절반-도메인(Fok I)의 부위 지정 돌연변이 유발에 의해 제조될 수 있다. 더 나아가, 본 명세서에 기재된 절단 도메인은 미국 특허 제9,394,531호 및 제9,567,609호에 기재된 링커를 포함하지만, 이들로 제한되지 않는 임의의 적합한 링커를 이용하여 DNA-결합 도메인(예를 들어, ZFP)에 융합될 수 있다.

[0126] 대안적으로, 뉴클레아제는 소위 "분할-효소" 기법을 이용하여 핵산 표적 부위에서 생체내로 조립될 수 있다(예를 들어, 미국 특허 공개 제2009/0068164호 참조). 이러한 분할 효소의 성분은 별개의 발현 작제물 중 하나에서 발현될 수 있거나 또는 하나의 오픈 리딩 프레임에서 연결될 수 있으며, 여기서 개개의 성분은, 예를 들어, 자기-절단성 2A 웹타이드 또는 IRES 서열에 의해 분리된다. 성분은 개개 아연 평거 결합 도메인 또는 메가뉴클레아제 핵산 결합 도메인의 도메인일 수 있다.

[0127] 뉴클레아제는 사용 전에, 예를 들어 WO 2009/042163 및 20090068164에 기재된 바와 같은 효모-기반 염색체 시스템에서 활성에 대해 선별될 수 있다. 뉴클레아제의 발현은 구성적 프로모터 또는 유도성 프로모터, 예를 들어 라피노스, 갈락토스 및 라피노스와 갈락토스의 조합의 존재 하에 활성화되는(탈억제되는) 그리고 글루코스의 존재 하에 억제되는 갈락토키나제 프로모터의 제어 하에 있을 수 있다.

[0128] 본 명세서에 기재되는 바와 같은 뉴클레아제(들)는 표적 부위에서 하나 이상의 이중 가닥, 하나 이상의 단일-가닥 절단 및 하나 이상의 이중 가닥과 하나 이상의 단일-가닥 절단의 조합을 만들 수 있다. 특정 실시형태에서, 뉴클레아제는 촉매적으로 비활성인 절단 도메인(예를 들어, FokI, Cas 단백질 및 FokI와 Cas 단백질의 조합)을 포함한다. 예를 들어, 미국 특허 제9,200,266호; 제8,703,489호 및 문헌[Guillinger *et al.* (2014) *Nature Biotech.* 32(6):577-582] 참조. 촉매적으로 비활성인 절단 도메인은 촉매적으로 활성인 도메인과 조합하여 단일-가닥 절단을 만들기 위한 틈내기 효소로서 작용할 수 있다. 따라서, 두 틈내기효소는 특정 영역에서 이중-가닥 절단을 만들기 위해 조합하여 사용될 수 있다. 추가적인 틈내기효소는 또한 당업계에 공지되어 있다, 예를 들어, 문헌[McCaffery *et al.* (2016) *Nucleic Acids Res.* 44(2):e11. doi: 10.1093/nar/gkv878. Epub 2015 Oct 19].

표적 부위

[0130] 상기 상세하게 기재한 바와 같이, DNA 도메인은 임의의 선택 서열에 결합하도록 조작될 수 있다. 조작된 DNA-결합 도메인은 천연 유래 DNA-결합 도메인에 비해 신규한 결합 특이성을 가질 수 있다. 특정 실시형태에서, DNA-결합 도메인은 BCL11A 인핸서 서열 내의 서열에 결합하며, 예를 들어 표적 부위(전형적으로 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21 또는 심지어 더 많은 염기쌍)는 표 1에 나타낸 바와 같은 BCL11A 인핸서 서열(예를 들어, +58)에서 DNaseI 초민감성 부위 내의 서열에 결합하는 DNA-결합 도메인을 포함하는 BCL11A의 엑손 2와 엑손 3 사이에 있다. 조작 방법은 합리적 설계 및 다양한 유형의 선택을 포함하지만, 이들로 제한되지 않는다. 합리적 설계는, 예를 들어, 삼중(또는 사중) 뉴클레오타이드 서열 및 개개의 아연 평거 아미노산 서열을 포함하는 데이터베이스를 이용하는 것을 포함하며, 이때 각각의 삼중 또는 사중 뉴클레오타이드 서열은 특정 삼중 또는 사중 서열에 결합하는 아연 평거의 하나 이상의 아미노산 서열과 관련된다. 예를 들어, 본 명세서에 전문

이 참고로 포함되는 공동 소유된 미국 특허 제6,453,242호 및 제6,534,261호 참조. TAL-효과기 도메인의 합리적 설계가 또한 수행될 수 있다. 예를 들어, 미국 특허 공개 제2011/0301073호 참조.

[0131] 예시적인 과자 디스플레이 및 2-흔성체 시스템을 포함하는 DNA-결합 도메인에 적용 가능한 선택 방법은 미국 특허 제5,789,538호; 5,925,523호; 6,007,988호; 6,013,453호; 6,410,248호; 6,140,466호; 6,200,759호; 및 6,242,568호; 뿐만 아니라 국제 특허 출원 공개 WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 및 영국 특허 제2,338,237호에서 개시된다. 추가로, 아연 핑거 결합 도메인에 대한 결합 특이성의 향상은, 예를 들어, 공동 소유된 WO 02/077227에 기재되어 있다.

[0132] 표적 부위의 선택; 융합 단백질(및 이를 암호화하는 폴리뉴클레오타이드)의 설계 및 구성을 위한 뉴클레아제 및 방법은 당업자에게 공지되어 있고, 본 명세서에 전문이 참고로 포함된 미국 특허 공개 제2005/0064474호 및 제2006/0188987호에서 상세하게 기재된다.

[0133] 추가로, 이들 및 다른 참고문헌에 개시된 바와 같이, DNA-결합 도메인(예를 들어, 다중-핑거 아연 핑거 단백질) 및 DNA-결합 도메인(들)과 기능성 도메인(들)의 융합은, 예를 들어, 5개 이상의 아미노산의 링커를 포함하는 임의의 적합한 링커 서열을 이용하여 함께 연결될 수 있다. 미국 특허 제 8,772,453호; 제7,888,121호(예를 들어, "ZC" 링커); 제6,479,626호; 제6,903,185호; 및 제7,153,949호; 미국 특허 공개 제2009/0305419호) 및 제2015/0064789호. 본 명세서에 기재된 단백질은 단백질의 개개 DNA-결합 도메인 사이의 적합한 링커의 임의의 조합을 포함할 수 있다. 또한, 미국 특허 제8,586,526호 참조.

공여자

[0135] 특정 실시형태에서, 본 개시내용은 본 명세서에 기재된 BCL11A 인핸서 영역-결합 분자를 이용하는 세포의 계놈 내로의 외인성 서열의 뉴클레아제-매개 표적화된 통합에 관한 것이다. 상기 언급한 바와 같이, 예를 들어 구체화된 영역의 결실을 위한, 돌연변이체 유전자의 보정을 위한, 구체화된 영역의 결실과 돌연변이체 유전자의 보정의 조합을 위한, 또는 야생형 유전자의 증가된 발현을 위한, 외인성 서열(또한 "공여자 서열" 또는 "공여자" 또는 "이식유전자"로 불림)의 삽입. 공여자 서열은 전형적으로 그것이 위치되는 계놈 서열과 동일하지 않다는 것이 용이하게 명백할 것이다. 공여자 서열은 관심 대상의 위치에서 효율적인 HDR를 허용하는 상동성의 두 영역에 의해 측접되는 비상동성 서열을 함유할 수 있거나 또는 NHEJ와 같은 비상동성 지향 수선 메커니즘을 통해 통합될 수 있다. 추가적으로, 공여자 서열은 세포의 염색질에서 관심 대상의 영역에 대해 상동성이 아닌 서열을 함유하는 백터 분자를 포함할 수 있다. 공여자 분자는 세포의 염색질에 대해 몇몇, 불연속적 상동성 영역을 함유할 수 있고, 예를 들어, 본 명세서에 기재된 뉴클레아제 중 하나에 의해 유도되는 DSB의 수선을 위한 기질로서 사용될 때 BCL11A 인핸서 영역(또는 이의 단편)의 결실을 야기한다. 추가로, 관심 대상의 영역에 정상적으로 존재하지 않는 서열의 표적화된 삽입을 위해, 상기 서열은 공여자 핵산 분자에 존재할 수 있고, 관심 대상의 영역 내 서열에 대한 상동성 영역에 측접될 수 있다.

[0136] 삽입을 위한 폴리뉴클레오타이드는 또한 "외인성" 폴리뉴클레오타이드, "공여자" 폴리뉴클레오타이드 또는 분자 또는 "이식유전자"로서 지칭될 수 있다. 공여자 폴리뉴클레오타이드는 DNA 또는 RNA, 단일-가닥 또는 이중-가닥 일 수 있고, 선형 또는 원형으로 세포 내로 도입될 수 있다. 예를 들어, 미국 특허 출원 공개 제2010/0047805호 및 제2011/0207221호 참조. 공여자 서열(들)은 원형 또는 선형으로 세포 내로 도입될 수 있는 DNA MC 내에 바람직하게 함유된다. 선형으로 도입된다면, 공여자 서열의 말단은 당업자에게 공지된 방법에 의해 (예를 들어, 핵산외부 분해(exonucleolytic degradation)로부터) 보호될 수 있다. 예를 들어, 하나 이상의 다이데옥시뉴클레오타이드 잔기는 선형 분자의 3' 말단에 첨가되고, 자기-상보성 올리고뉴클레오타이드는 말단 중 하나 또는 둘 다에 선택적으로 결찰된다. 예를 들어, 문헌[Chang *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:4959-4963; Nehls *et al.* (1996) *Science* 272:886-889] 참조. 분해로부터 외인성 폴리뉴클레오타이드를 보호하기 위한 추가적인 방법은 말단 아미노기(들)의 첨가 및 변형된 뉴클레오타이드 결합, 예컨대, 포스포로티오에이트, 포스포르아미데이트 및 O-메틸 리보스 또는 데옥시리보스 잔기의 사용을 포함하지만, 이들로 제한되지 않는다. 이중 가닥 형태로 도입된다면, 공여자는 세포 계놈 내로 통합될 이식유전자에 측접하는 하나 이상의 뉴클레아제 표적 부위, 예를 들어, 뉴클레아제 표적 부위를 포함할 수 있다. 예를 들어, 미국 특허 공개 제2013/0326645호 참조.

[0137] 폴리뉴클레오타이드는 추가적인 서열을 갖는 백터 분자의 부분, 예를 들어, 복제 기점, 프로모터 및 항생제 내성을 암호화하는 유전자로서 세포 내로 도입될 수 있다. 게다가, 공여자 폴리뉴클레오타이드는 네이키드 핵산으로서, 리포좀 또는 폴록사미와 같은 제제와 복합체화된 핵산으로서 도입될 수 있거나, 또는 바이러스(예를 들어, 아데노바이러스, AAV, 헤르페스바이러스, 레트로바이러스, 렌티바이러스 및 인테그라제 결합 렌티바이러스(IDLV))에 의해 전달될 수 있다.

- [0138] 특정 실시형태에서, 이중-가닥 공여자는 길이가 1kb 초과, 예를 들어 2 내지 200kb, 2 내지 10kb(또는 그 사이의 임의의 값)의 서열(예를 들어, 또한 이식유전자로서 지칭되는 암호 서열)을 포함한다. 이중-가닥 공여자는 또한, 예를 들어, 적어도 하나의 뉴클레아제 표적 부위를 포함한다. 특정 실시형태에서, 공여자는, 예를 들어 ZFN 또는 TALEN의 쌍에 대해 적어도 2개의 표적 부위를 포함한다. 전형적으로, 뉴클레아제 표적 부위는 이식유전자의 절단을 위해 이식유전자 서열 외부, 예를 들어, 이식유전자 서열에 대해 5' 및/또는 3'이다. 뉴클레아제 절단 부위(들)는 임의의 뉴클레아제(들)에 대해서일 수 있다. 특정 실시형태에서, 이중-가닥 공여자에 함유된 뉴클레아제 표적 부위(들)는 절단된 공여자가 상동성-의존적 방법을 통해 통합된 내인성 표적을 절단하기 위해 사용되는 동일한 뉴클레아제(들)에 대한 것이다.
- [0139] 공여자는 일반적으로 그의 발현이 통합 부위에서 내인성 프로모터, 즉, 공여자가 삽입되는 내인성 유전자의 발현을 유도하는 프로모터(예를 들어, 글로빈, AAVS1 등)에 의해 유도되도록 삽입된다. 그러나, 공여자가 프로모터, 인핸서, 및 프로모터와 인핸서 둘 다의 조합, 예를 들어 구성적 프로모터 또는 유도성 또는 조직 특이적 프로모터를 포함할 수 있다는 것은 명확할 것이다.
- [0140] 공여자 분자는 내인성 유전자 중 모두, 일부가 발현되거나 또는 어떤 것도 발현되지 않도록 내인성 유전자 내로 삽입될 수 있다. 다른 실시형태에서, 이식유전자(예를 들어, 글로빈 암호화 서열이 있거나 또는 없음)는 임의의 내인성 좌위, 예를 들어, 세이프-하버 좌위 내로 통합된다. 예를 들어, 미국 특허 공개 제2008/0299580호; 제2008/0159996호 및 제2010/00218264호 참조.
- [0141] 더 나아가, 발현에 필요하지 않다고 해도, 외인성 서열은 또한 전사 또는 번역 조절 서열, 예를 들어, 프로모터, 인핸서, 절연체, 내부 리보솜 유입 부위, 서열 암호화 2A 펩타이드, 폴리아데닐화 신호, 및 이들의 조합을 포함할 수 있다.
- [0142] 본 명세서에 기재된 공여자 서열 상에서 수행되는 이식유전자는 PCR과 같은 당업계에 공지된 표준 기법을 이용하여 플라스미드, 세포 또는 다른 공급원으로부터 단리될 수 있다. 사용을 위한 공여자는 원형 초나선, 원형 이완, 선형 등을 포함하는 다양한 유형의 위상을 포함할 수 있다. 대안적으로, 그들은 표준 올리고뉴클레오타이드 합성 기법을 이용하여 화학적으로 합성될 수 있다. 추가로, 공여자는 메틸화되거나 또는 메틸화가 결여될 수 있다. 공여자는 박테리아 또는 효모 인공 염색체(BAC 또는 YAC)의 형태일 수 있다.
- [0143] 본 명세서에 기재된 이중-가닥 공여자 폴리뉴클레오타이드는 하나 이상의 비천연 염기, 하나 이상의 골격 및 하나 이상의 비천연 염기와 하나 이상의 골격의 조합을 포함할 수 있다. 특히, 메틸화된 사이토카인과 함께 공여자 분자의 삽입은 관심 대상의 영역에서 전사 정지를 상태를 달성하기 위해 본 명세서에 기재된 방법을 이용하여 수행될 수 있다.
- [0144] 외인성(공여자) 폴리뉴클레오타이드는 관심 대상의 임의의 서열(외인성 서열)을 포함할 수 있다. 예시적인 외인성 서열은 임의의 폴리펩타이드 암호 서열(예를 들어, cDNA), 프로모터 서열, 인핸서 서열, 에피토프 태그, 마커 유전자, 절단 효소 인식 부위 및 다양한 유형의 발현 작제물을 포함하지만, 이들로 제한되지 않는다. 마커 유전자는 항생제 내성(예를 들어, 암페실린 내성, 카나마이신 내성, 네오마이신 내성, G418 내성, 퓨로마이신 내성, 하이그로마이신 내성, 블라스티시딘 내성)을 매개하는 서열 암호화 단백질, 서열 암호화 착색 또는 형광 또는 발광 단백질(예를 들어, 녹색 형광 단백질, 향상된 녹색 형광 단백질, 적색 형광 단백질, 루시피라제), 및 향상된 세포 성장, 유전자 증폭을 매개하는 단백질(예를 들어, 다이하이드로플레이트 환원효소) 및 향상된 세포 성장과 유전자 증폭의 조합을 포함하지만, 이들로 제한되지 않는다. 에피토프 태그는, 예를 들어, FLAG, His, Myc, 칙렬 친화도 정제(tandem affinity purification: TAP), HA, 바이오틴일화 가능한 펩타이드 또는 임의의 검출 가능한 아미노산 서열 중 하나 이상을 포함한다.
- [0145] 바람직한 실시형태에서, 외인성 서열(이식유전자)은, 항체, 항원, 효소, 수용체(세포 표면 또는 핵), 호르몬, 림포카인, 사이토카인, 리포터, 폴리펩타이드, 성장 인자 및 상기 중 임의의 기능성 단편을 포함하지만, 이들로 제한되지 않는 세포 내에서 발현이 요망되는, 임의의 폴리펩타이드를 암호화하는 폴리뉴클레오타이드를 포함한다. 암호 서열은, 예를 들어, cDNA일 수 있다.
- [0146] 예를 들어, 외인성 서열은 임의의 다음의 질환을 포함하지만, 이들로 제한되지 않는 유전자 질환을 갖는 대상체에서 결여되거나 또는 비기능성인 폴리펩타이드를 암호화하는 서열을 포함한다: 연골무형성증, 완전색맹, 산 말 타제 결핍증, 아데노신 테아미나제 결핍증(OMIM No.102700), 부신백질이영양증, 에카르디 증후군, 알파-1 항트립신 결핍증, 알파-탈라세미아, 안드로겐 무감성 증후군, 에이피트 증후군, 부정맥유발성 우심실 형성이상, 이형성증, 혈관확장성 운동실조증, 바쓰 증후군, 베타-탈라세미아, 청색 고무공상모반, 카나반병, 만성 육아종 질

환(CGD), 고양이 울음 증후군, 낭성 섬유증, 텔컴병, 외배엽 이형성증, 판코니 빈혈, 진행성 골화섬유형성이상, 취약 X 증후군, 갈락토스혈증, 고체병, 전신성 강글리오사이드증(예를 들어, GM1), 혈색소증, 베타-글로빈의 6 번째 코돈에서의 헤모글로빈 C 돌연변이(HbC), 혈우병, 헌팅턴병, 후를러 증후군, 저인산효소증, 클라인펠터 증후군, 크라베병, 랑거 기드온 증후군, 백혈구 부착 결핍증(LAD, OMIM No. 116920), 백질이영양증, 긴 QT 증후군, 마르팡 증후군, 뇌비우스 증후군, 점액다당류증(MPS), 순발톱무릎뼈 증후군, 콩팥기원요붕증, 신경섬유종증, 니만-페크병, 골형성 부전증, 포르파린증, 프라더-윌리 증후군, 조로증, 프로테우스 증후군, 망막모세포증, 레트 증후군, 루빈스타인-테이비 증후군, 산필립포증후군, 중증 합병성 면역결핍(SCID), 슈와크만 증후군, 겸상 적혈구병(겸상 적혈구 빈혈), 스미스-마제니스 증후군, 스티클러 증후군, 테이 삭스병, 요골 무형성(TAR) 증후군, 트레처 콜린스 증후군, 삼염색체, 결절성 경화증, 터너 증후군, 요소회로질환, 폰 힙펠-린도우병, 바르덴부르그 증후군, 윌리엄스 증후군, 월슨병, 비스코트 올드리치 증후군, X-연관 림프구증식 증후군(XLP, OMIM No. 308240).

[0147] 표적화된 통합에 의해 치료될 수 있는 추가적인 예시적 질환은 후천성 면역결핍증, 라이소좀 저장병(예를 들어, 고체병, GM1, 파브리병 및 테이 삭스병), 점액다당류증(예를 들어, 헌터병, 혈관병), 이상혈색소증(예를 들어, 겸상 적혈구병, HbC, α -탈라세미아, β -탈라세미아) 및 혈우병을 포함한다.

[0148] 특정 실시형태에서, 외인성 서열은 표적화된 통합을 겪는 세포의 선택을 허용하는 마커 유전자(상기 기재) 및 추가적인 기능성을 암호화하는 연결된 서열을 포함할 수 있다. 마커 유전자의 비제한적 예는 GFP, 약물 선택 마커(들) 등을 포함한다.

[0149] 삽입될 수 있는 추가적인 유전자 서열은 돌연변이된 서열을 대체하기 위해, 예를 들어, 야생형 유전자를 포함할 수 있다. 예를 들어, 야생형 인자 IX 유전자 서열은 유전자의 내인성 복제물이 돌연변이되는 줄기 세포의 계놈 내로 삽입될 수 있다. 야생형 복제물은 내인성 좌위에 삽입될 수 있거나, 또는 대안적으로 세이프 하버 좌위에 표적화될 수 있다.

[0150] 본 명세서의 교시 후에 이러한 발현 카세트의 작제는 문자 생물학 분야에 잘 알려진 방법을 이용한다(예를 들어, 문헌[Ausubel 또는 Maniatis] 참조). 유전자이식 동물을 생성하기 위한 발현 카세트의 사용 전에, 선택된 제어 요소와 관련된 스트레스-유도자에 대한 발현 카세트의 반응성은 발현 카세트를 적합한 세포주(예를 들어, 1차 세포, 형질전환 세포 또는 불멸 세포주) 내로 도입함으로써 시험될 수 있다.

[0151] 더 나아가, 발현에 필요하지 않다고 해도, 외인성 서열은 또한 전사 또는 번역 조절 서열, 예를 들어, 프로모터, 인핸서, 절연체, 내부 리보솜 유입 부위, 서열 암호화 2A 웹타이드, 폴리아데닐화 신호, 및 2A 폴리 웹타이드와 폴리아데닐화 신호의 조합을 포함할 수 있다. 추가로, 관심 대상의 유전자의 제어 요소는 키메라 유전자(예를 들어, 리포터 발현 카세트)를 생성하기 위해 리포터 유전자에 작동 가능하게 연결될 수 있다.

[0152] 비암호 핵산 사열의 표적화된 삽입이 또한 달성될 수 있다. 앤티센스 RNA, RNAi, shRNA 및 마이크로 RNA(miRNA)를 암호화하는 서열이 또한 표적화된 삽입을 위해 사용될 수 있다.

[0153] 추가적인 실시형태에서, 공여자 핵산은 추가적인 뉴클레아제 설계를 위한 특정 표적 부위인 비암호 서열을 포함할 수 있다. 후속적으로, 본래의 공여자 분자가 절단되고 관심 대상의 다른 공여자 분자의 삽입에 의해 변형되도록 추가적인 뉴클레아제는 발현될 수 있다. 이러한 방법에서, 관심 대상의 특정 좌위에서 또는 세이프 하버 좌위에서 형질 스태킹을 허용하는 공여자 분자의 반복적 통합이 생성될 수 있다.

0154] 전달

[0155] 본 명세서에 기재된 바와 같은 뉴클레아제(표 1), 이를 뉴클레아제를 암호화하는 폴리뉴클레오타이드, 공여자 폴리뉴클레오타이드 및 본 명세서에 기재된 단백질, 폴리뉴클레오타이드 및 단백질과 폴리뉴클레오타이드를 포함하는 조성물은 임의의 세포 유형에 임의의 적합한 수단에 의해 생체내 또는 생체외에서 전달될 수 있다.

[0156] 적합한 세포는 진핵(예를 들어, 동물) 및 원핵 세포 및 진핵과 원핵 세포주를 포함한다. 이러한 세포 또는 이러한 세포로부터 생성된 세포주의 비제한적 예는 COS, CHO(예를 들어, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NS0, SP2/0-Ag14, HeLa, HEK293(예를 들어, HEK293-F, HEK293-H, HEK293-T) 및 perC6 세포뿐만 아니라 곤충 세포, 예컨대 스포돕테라푸기페르다(*Spodopteracutipennis*: Sf), 또는 진균 세포, 예컨대 사카로마이세스(*Saccharomyces*), 피키아(*Pichia*) 및 쉬조 사카로마이세스(*Schizosaccharomyces*)를 포함한다. 특정 실시형태에서, 세포주는 CHO, MDCK 또는 HEK293 세포주이다. 적합한 세포는 또한, 예로서, 배아 줄기 세포, 유도만능 줄기세포, 조혈 줄기 세포, 신경 줄기 세포 및

중간엽 줄기 세포를 포함한다.

[0157] 본 명세서에 기재된 바와 같은 뉴클레아제를 전달하는 방법은, 예를 들어, 미국 특허 제6,453,242호; 제6,503,717호; 제6,534,261호; 제6,599,692호; 제6,607,882호; 제6,689,558호; 제6,824,978호; 제6,933,113호; 제6,979,539호; 제7,013,219호; 및 제7,163,824호에 기재되며, 이들 모두의 개시내용은 그들의 전문이 본 명세서에 참고로 포함된다.

[0158] 본 명세서에 기재된 바와 같은 뉴클레아제, 공여자 작제물, 및 뉴클레아제와 공여자 작제물의 조합은 또한 본 명세서에 기재된 ZFN(들) 중 하나 이상을 암호화하는 서열을 함유하는 벡터를 이용하여 전달될 수 있다. 플라스미드 벡터, 레트로바이러스 벡터, 렌티바이러스 벡터, 아데노바이러스 벡터, 폭스바이러스 벡터; 헤르페스바이러스 벡터 및 아데노 연관 벡터 등을 포함하지만, 이들로 제한되지 않는 임의의 벡터 시스템이 사용될 수 있다. 또한, 미국 특허 제6,534,261호; 제6,607,882호; 제6,824,978호; 제6,933,113호; 제6,979,539호; 제7,013,219호; 및 제7,163,824호를 참조하며, 이들은 본 명세서에 그들의 전문이 참고로 포함된다. 더 나아가, 임의의 이들 벡터는 치료에 필요한 서열 중 하나 이상을 포함할 수 있다는 것이 명백할 것이다. 따라서, 하나 이상의 뉴클레아제 및 공여자 작제물이 세포 내로 도입될 때, 뉴클레아제, 공여자 폴리뉴클레오타이드, 및 뉴클레아제와 공여자 폴리뉴클레오타이드의 조합은 동일한 벡터 상에서 또는 상이한 벡터 상에서(DNA MC(들)) 운반될 수 있다. 다중 벡터가 사용될 때, 각각의 벡터는 하나 또는 다중 뉴클레아제, 하나 이상의 공여자 작제물, 및 하나 이상의 뉴클레아제와 하나 이상의 공여자 작제물의 조합을 암호화하는 핵산을 도입하기 위해 사용될 수 있다. 비-바이러스 및 비바이러스 기반 유전자 전달 방법은 세포(예를 들어, 포유류 세포) 및 표적 조직에서 뉴클레아제, 공여자 작제물, 및 뉴클레아제와 공여자 작제물의 조합을 암호화하는 핵산을 도입하기 위해 사용될 수 있다. 비-바이러스 벡터 전달 시스템은 DNA 또는 RNA 플라스미드, DNA MC, 네이키드 핵산, 및 전달 비히클, 예컨대 리포좀 또는 폴록사미와 복합체화된 핵산을 포함한다. 적합한 비-바이러스 벡터는 InCellArt(프랑스)로부터 상업적으로 입수 가능한 벡터를 포함하는 나노탁시스 벡터를 포함한다. 바이러스 벡터 전달 시스템은 세포에 대한 전달 후에 에피솜 또는 통합된 게놈 중 하나를 갖는 DNA 및 RNA 바이러스를 포함한다. 조작된 DNA-결합 단백질 및 이들 결합 단백질을 포함하는 융합 단백질의 생체내 전달의 검토를 위해, 예를 들어, 문헌[Rebar (2004) Expert Opinion Invest. Drugs 13(7):829-839; Rossi *et al.* (2007) Nature Biotech. 25(12):1444-1454]뿐만 아니라 일반적 유전자 전달 참고문헌에 대해 문헌[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 and Neuroscience 8:35-36 (1995); Kremer & Perricaudet, British Medical Bulletin 51(1):31-44 (1995); Haddada *et al.*, Current Topics in Microbiology and Immunology Doerfler and Böhm (eds.) (1995); 및 Yu *et al.*, Gene Therapy 1:13-26 (1994)] 참조.

[0159] 핵산의 비바이러스 전달 방법은 전기천공법, 리포펙션, 미량주사법, 유전자총, 바이로좀, 리포좀, 면역리포좀, 다양이온 또는 지질:핵산 접합체, 네이키드 DNA, 인공 비리온, 막 변형 및 DNA의 제제-향상 흡수를 포함한다. 예를 들어, 소니트론(Sonitron) 2000 시스템(리치-마르(Rich-Mar))을 이용하는 초음파천공법(Sonoporation)이 또한 핵산의 전달을 위해 사용될 수 있다.

[0160] 추가적인 예시적 핵산 전달 시스템은 아맥사 바이오시스템즈(Amaxa Biosystems)(독일 쾨른에 소재), 맥스사이트 인코포레이티드(Maxcyte, Inc.)(메릴랜드주 래빌에 소재), BTX 몰레큘러 딜리버리 시스템즈(BTX Molecular Delivery Systems)(매사추세츠주 홀리스턴에 소재) 및 코페르니쿠스 세라퓨틱스 인코포레이티드(Copernicus Therapeutics Inc.)(예를 들어, 미국 특허 제6008336호 참조)에 의해 제공되는 것을 포함한다. 리포펙션은, 예를 들어, 미국 특허 제5,049,386호; 제4,946,787호; 및 제4,897,355호)에 기재되어 있고, 리포펙션 시약은 상업적으로 시판된다(예를 들어, 트랜스펙탐(Transfectam)(상표명) 및 리포펙틴(Lipofectin)(상표명))에 기재되어 있다. 폴리뉴클레오타이드의 효율적인 수용체-인식 리포펙션에 적합한 양이온성 및 중성 지질은 펠그너(Felgner)의 WO 91/17424, WO 91/16024의 이들을 포함한다.

[0161] 표적화된 리포좀을 포함하는 지질:핵산 복합체, 예컨대 면역지질 복합체의 제조는 당업자에게 잘 공지되어 있다(예를 들어, 문헌[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 제4,186,183호, 제4,217,344호, 제4,235,871호, 제4,261,975호, 제4,485,054호, 제4,501,728호, 제4,774,085호, 제4,837,028호 및 제4,946,787호 참조). 다른 지질:핵산 복합체는 신규한 양이온성 지질, 신규한

폐길화된 지질 및 신규한 양이온성 지질과 신규한 폐길화된 지질의 조합을 포함하는 것을 포함한다(미국 가출원 특허 제62/432,042호 및 제62/458,373호 참조).

[0162] 전달의 추가적인 방법은 EnGeneIC 전달 비하클(EDV)에 전달될 핵산의 패키징의 사용을 포함한다. 이들 EDV는 이 중특이성 항체를 이용하여 표적 조직에 특이적으로 전달되며, 여기서, 항체의 하나의 아암(arm)은 표적 조직에 대한 특이성을 가지며, 다른 아암은 EDV에 대해 특이성을 가진다. 항체는 EDV를 표적 세포 표면으로 가져오며, 이어서, EDV는 내포작용에 의해 세포 내로 운반된다. 일단 세포 내이면, 내용물이 방출된다(문헌[MacDiarmid *et al.*, (2009) *Nature Biotechnology* 27(7):643] 참조).

[0163] 조작된 ZFP, TALE 및 CRISPR/Cas 시스템을 암호화하는 핵산의 전달을 위한 RNA 또는 DNA 바이러스 기반 시스템의 사용은 신체에서 바이러스를 특정 세포에 표적화하기 위한 그리고 바이러스 페이로드를 핵으로 수송하기 위한 고도로 진화된 과정의 이점을 취한다. 바이러스 벡터는 환자에게(생체내에서) 직접 투여될 수 있거나 또는 그들은 시험관내에서 세포를 처리하기 위해 사용될 수 있고, 변형된 세포는(생체외에서) 환자에게 투여된다. ZFP의 전달을 위한 통상적인 바이러스 기반 시스템은 유전자 전달을 위해 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노-연관, 백시니아 및 단순포진 바이러스 벡터를 포함하지만, 이들로 제한되지 않는다. 숙주계놈에서의 통합은 종종 삽입된 이식유전자의 장기간 발현을 초래하는 레트로바이러스, 렌티바이러스 및 아데노연관 바이러스 유전자 전달 방법에 의해 가능하다. 추가적으로, 고 형질도입 효율은 다수의 상이한 세포 유형 및 표적 조직에서 관찰되었다.

[0164] 레트로바이러스의 향성은 표적 세포의 잠재적 표적 집단을 확장시키는 외래 외피 단백질을 혼입함으로써 변경될 수 있다. 렌티바이러스 벡터는 비분화 세포를 형질도입하거나 또는 감염시키고 전형적으로 고바이러스 역가를 생성할 수 있는 레트로바이러스 벡터이다. 레트로바이러스 유전자 전달 시스템의 선택은 표적 조직에 의존한다. 레트로바이러스 벡터는 외래 서열의 6 내지 10kb까지에 대한 패키징 능력을 갖는 시스-작용성 긴 말단 반복부를 포함한다. 최소 시스-작용성 LTR은 벡터의 복제 및 패키징에 충분하며, 이는 이어서 영구한 이식유전자 발현을 제공하기 위해 표적 세포 내로 치료 유전자를 통합하는 데 사용된다. 널리 사용되는 레트로바이러스 벡터는 뮤린 백혈병 바이러스(MuLV), 일본 유인원 백혈병 바이러스(GaLV), 시미안 면역결핍 바이러스(SIV), 인간 면역결핍 바이러스(HIV), 및 이들의 조합에 기반한 것을 포함한다(예를 들어, 문헌[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).

[0165] 일시적 발현이 바람직한 적용에서, 아데노바이러스 기반 시스템이 사용될 수 있다. 아데노바이러스 기반 벡터는 다수의 세포 유형에서 매우 높은 형질도입 효율을 가질 수 있고, 세포 분화를 필요로 하지 않는다. 이러한 벡터에 의해, 고역가 및 고수준의 발현이 얻어졌다. 이 벡터는 상대적으로 단순한 시스템에서 다양으로 생성될 수 있다. 아데노-연관 바이러스("AAV") 벡터는 또한, 예를 들어, 핵산 및 렙타이드의 시험관내 생성에서, 그리고 생체내 및 생체외 유전자 요법 절차에서 표적 핵산을 형질도입하기 위해 사용된다(예를 들어, 문헌[West *et al.*, *Virology* 160:38-47 (1987); 미국 특허 제4,797,368호; WO 93/24641; Kotin, *Human Gene Therapy* 5:793-801 (1994); Muzyczka, *J. Clin. Invest.* 94:1351 (1994)] 참조. 재조합 AAV 벡터의 구성은 미국 특허 제5,173,414호; 문헌[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.* 63:03822-3828 (1989)]을 포함하는 다수의 간행물에 기재되어 있다.

[0166] 적어도 6개의 바이러스 벡터 접근은 형질도입제를 생성하기 위해 헬퍼 세포주에 삽입된 유전자에 의해 결합 벡터의 보완을 수반하는 접근을 이용하는, 임상 시험에서 유전자 전달을 위해 현재 이용 가능하다.

[0167] pLASN 및 MFG-S는 임상 시험에서 사용되는 레트로바이러스 벡터의 예이다(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 시험에서 사용되는 제1 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)).

[0168] 재조합 아데노-연관 바이러스 벡터(rAAV)는 결합 및 비병원성 파보바이러스 아데노-연관 2형 바이러스에 기반한 촉망받는 대안의 유전자 전달 시스템이다. 모든 벡터는 이식유전자 발현 카세트에 측접하는 AAV 145bp 역위 말단 반복부만을 보유하는 플라스미드로부터 유래된다. 형질도입 세포의 계놈 내로의 통합에 기인하는 효율적인 유전자 전달 및 안정한 이식유전자 전달은 본 벡터 시스템에 대한 중요한 특징이다. (Wagner *et al.*, *Lancet*

351:9117 1702-3 (1998), Kearns *et al.*, Gene Ther. 9:748-55 (1996)). AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9 및 AAVrh.10 및 임의의 신규한 AAV 혈청형을 포함하는, 다른 AAV 혈청형은 또한 본 발명에 따라 사용될 수 있다.

[0169] 복제-결합 재조합 아데노바이러스 벡터(Ad)는 고역가에서 생성될 수 있고, 다수의 상이한 세포 유형을 용이하게 감염시킬 수 있다. 이식유전자가 Ad E1a, E1b 또는 E3 유전자를 대체하도록 대부분의 아데노바이러스 벡터는 조작되고; 후속적으로, 복제 결합 벡터는 도중에 결실된 유전자 기능을 공급하는 인간 293 세포에서 증식된다. Ad 벡터는 비분화, 분화 세포, 예컨대 간, 신장 및 근육에서 발견된 것을 포함하는 다양한 유형의 생체내 조직을 형질도입할 수 있다. 통상적인 Ad 벡터는 큰 운반 능력을 가진다. 임상 시험에서 Ad 벡터 사용의 예는 근육내 주사에 의한 항종양 면역화를 위해 폴리뉴클레오타이드 요법을 수반하였다(Sterman *et al.*, Hum. Gene Ther. 7:1083-9 (1998)). 임상 시험에서 유전자 전달을 위한 아데노바이러스 벡터 사용의 추가적인 예는 문헌 [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)]을 포함한다.

[0170] 패키징 세포는 숙주 세포를 감염시킬 수 있는 바이러스 입자를 형성하기 위해 사용된다. 이러한 세포는 AAV 및 아데노바이러스를 패키징하기 위해 사용될 수 있는 HEK293 및 Sf9 세포, 및 레트로바이러스를 패키징하는 ψ2 세포 또는 PA317 세포를 포함한다. 유전자 요법에서 사용되는 바이러스 벡터는 보통 바이러스 입자 내로 핵산 벡터를 패키징하는 생산자 세포주에 의해 생성된다. 벡터는 전형적으로 패키징 및 (적용 가능하다면) 숙주 내로의 후속적인 통합에 필요한 최소 바이러스 서열을 함유하며, 다른 바이러스 서열은 발현될 단백질을 암호화하는 발현 카세트에 의해 대체된다. 상실된 바이러스 기능은 세포주를 패키징함으로써 도중에 공급된다. 예를 들어, 유전자 요법에서 사용되는 AAV 벡터는 전형적으로 숙주 세포 내로 패키징 및 통합을 위해 필요한 AAV 세포주로 부터 역위 말단 반복부(ITR) 서열만을 가진다. 바이러스 DNA는 다른 AAV 유전자를 암호화하는 헬퍼 플라스미드, 즉, *rep* 및 *cap*를 함유하지만, ITR 서열을 결여하는 세포주에서 패키징된다. 세포주는 또한 헬퍼로서 아데노바이러스에 의해 감염된다. 헬퍼 바이러스는 헬퍼 플라스미드로부터 AAV 벡터의 복제 및 AAV 유전자의 발현을 촉진시킨다. 헬퍼 플라스미드는 ITR 서열의 결여에 기인하여 상당한 양으로 패키징되지 않는다. 아데노바이러스에 의한 오염은, 예를 들어, 아데노바이러스가 AAV보다 더 민감한 열 처리에 의해 감소될 수 있다. 일부 실시형태에서, AAV는 바콜로바이러스 발현 시스템을 이용하여 생성된다(예를 들어, 미국 특허 제6,723,551호 및 제7,271,002호 참조).

[0171] 293 또는 바콜로바이러스 시스템으로부터 AAV 입자의 정제는 전형적으로 바이러스를 생성하는 세포의 성장 다음에, 세포 상청액으로부터 바이러스 입자의 수집 또는 세포의 용해 및 조질의 용해물로부터 바이러스의 수집을 수반한다. 이어서, AAV는 이온 교환 크로마토그래피(예를 들어, 미국 특허 제7,419,817호 및 제6,989,264호 참조), 이온 교환 크로마토그래피 및 CsCl 밀도 원심분리(예를 들어, 국제 특허 출원 공개 WO2011094198A10), 면역친화도 크로마토그래피(예를 들어, WO2016128408) 또는 AVB 세파로스를 이용하는 정제(예를 들어, GE 헬스케어 라이프 사이언시즈(GE Healthcare Life Sciences))를 포함하는, 당업계에 공지된 방법에 의해 정제된다.

[0172] 다수의 유전자 요법 적용에서, 유전자 요법 벡터는 특정 세포 유형에 대해 높은 정도의 특이성으로 전달되는 것이 바람직하다. 따라서, 바이러스 벡터는 바이러스의 외면 상에서 바이러스 외피 단백질과의 융합 단백질로서 리간드를 발현시킴으로써 주어진 세포 유형에 대해 특이성을 갖도록 변형될 수 있다. 리간드는 관심 대상의 세포 유형 상에 존재하는 것으로 알려진 수용체에 대해 친화도를 갖도록 선택된다. 예를 들어, 문헌[Han *et al.*, Proc. Natl. Acad. Sci. USA 92:9747-9751 (1995)]은 몰로니 뮤린 백혈병 바이러스가 gp70에 융합된 인간 헤레큘린 융합을 발현시키도록 변형될 수 있고, 재조합 바이러스는 인간 표피 성장 인자 수용체를 발현시키는 특정 인간 유방암 세포를 감염시킨다는 것이 보고되었다. 이 원칙은 다른 바이러스-표적 세포상으로 연장될 수 있으며, 이때 표적 세포는 수용체를 발현시키고, 바이러스는 세포 표면 수용체에 대한 리간드를 포함하는 융합 단백질을 발현시킨다. 예를 들어, 섬유성 파지는 사실상 임의의 선택된 세포 수용체에 대해 특정 결합 친화도를 갖는 항체 단편(예를 들어, FAB 또는 Fv)을 나타내도록 조작될 수 있다. 상기 설명이 바이러스 벡터에 주로 적용된다면, 비바이러스 벡터에 동일한 원칙이 적용될 수 있다. 이러한 벡터는 특정 표적 세포에 의한 흡수를 선호하는 특정 흡수 서열을 함유하도록 조작될 수 있다.

[0173] 유전자 요법 벡터는 개개 환자에 대한 투여에 의해, 전형적으로 전신 투여에 의해(예를 들어, 정맥내, 복강내, 근육내, 피하 또는 두개내 주입) 또는 상기 기재한 바와 같은 국소 적용에 의해 생체내로 전달될 수 있다. 대안적으로, 벡터는 생체외 세포, 예컨대 개개 환자로부터 외식된 세포(예를 들어, 림프구, 골수 흡입물, 조직 생검) 또는 보편적 공여자 조혈 줄기 세포로 전달된 다음에, 보통 벡터 내로 혼입된 세포에 대한 선택 후에 환

자 내로 세포가 재이식될 수 있다.

[0174] 뉴클레아제, 공여자 작제물, 및 뉴클레아제와 공여자 작제물의 조합을 함유하는 벡터(예를 들어, 레트로바이러스, 아데노바이러스, 리포좀 등)는 또한 생체내에서 세포의 형질도입을 위해 유기체에 직접 투여될 수 있다. 대안적으로, 네이키드 DNA가 투여될 수 있다. 투여는 주사, 주입, 국소 적용 및 전기천공법을 포함하지만, 이들로 제한되지 않는, 혈액 또는 조직 세포와의 궁극적 접촉으로 분자를 도입하기 위해 정상적으로 사용되는 임의의 경로에 의한다. 이러한 혁신을 투여하는 적합한 방법을 이용 가능하며, 당업자에게 잘 공지되어 있고, 특정 조성물을 투여하기 위해 하나 초과의 경로가 사용될 수 있지만, 특정 경로는 종종 다른 경로보다 더 즉석의 그리고 더 효과적인 반응을 제공할 수 있다.

[0175] 본 명세서에 기재된 폴리뉴클레오타이드의 도입에 적합한 벡터(예를 들어, 뉴클레아제-암호화, 이중-가닥 공여자, 및 뉴클레아제-암호화 및 이중-가닥 공여자의 조합)는 비-통합 렌티바이러스 벡터(non-integrating lentivirus vector: IDLV)를 포함한다. 예를 들어, 문헌[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 공개 제2009/0117617호 참조.

[0176] 약제학적으로 허용 가능한 담체는 투여 중인 특정 조성물에 의해서뿐만 아니라 조성물을 투여하기 위해 사용되는 특정 방법에 의해 부분적으로 결정된다. 따라서, 이하에 기재하는 바와 같은 약제학적 조성물의 매우 다양한 적합한 제형이 있다(예를 들어, 문헌[Remington's Pharmaceutical Sciences, 17th ed., 1989] 참조).

[0177] 뉴클레아제-암호화 서열 및 공여자 작제물은 동일한 또는 상이한 시스템을 이용하여 전달될 수 있다는 것이 명확할 것이다. 예를 들어, 뉴클레아제 및 공여자는 동일한 DNA MC에 의해 운반될 수 있다. 대안적으로, 공여자 폴리뉴클레오타이드는 MC에 의해 운반될 수 있는 반면, 하나 이상의 뉴클레아제는 표준 플라스미드 또는 AAV 벡터에 의해 운반될 수 있다. 더 나아가, 상이한 벡터는 동일 또는 상이한 경로(근육내 주사, 꼬리 정맥 주사, 다른 정맥내 주사, 복강내 투여 또는 근육내 주사)에 의해 투여될 수 있다. 벡터는 동시에 또는 임의의 순차적 순서로 전달될 수 있다.

[0178] 따라서, 본 개시내용은 치료 단백질을 암호화하는 이식유전자의 삽입을 받아 들이는 질환 및 병태의 생체내 또는 생체외 치료를 포함한다. 조성물은 혈청 또는 표적 기관 또는 세포에서 치료 폴리펩타이드의 목적으로 하는 농도를 얻는 데 효과적인 양으로 인간 환자에게 투여된다. 투여는 폴리뉴클레오타이드가 목적으로 하는 표적 세포에 전달되는 임의의 수단에 의할 수 있다. 예를 들어, 생체내 그리고 생체외 방법이 둘 다 상정된다. 간문맥에 대한 정맥내 주사는 바람직한 투여 방법이다. 다른 생체내 투여 방식은, 예를 들어, 간엽 또는 담도 내로의 직접 주사 및 간동맥을 통하는 것을 포함하는 간의 원위에 대한 정맥내 주사, 간 실질에 대한 직접 주사, 간동맥을 통한 주사 및 담관가지를 통한 역행 주사를 포함한다. 생체외 투여 방식은 절제된 간세포 또는 간의 다른 세포의 시험관내 형질도입 다음에, 문맥 맥관구조, 인간 환자의 간 실질 또는 담관가지 내로 다시 형질도입, 절제된 간세포의 주입을 포함하며, 예를 들어, 문헌[Grossman *et al.*, (1994) Nature Genetics, 6:335-341] 참조.

[0179] 투여될 유효량의 뉴클레아제(들) 및 공여자는 환자에 따라 그리고 관심 대상의 치료적 폴리펩타이드에 따라 다를 것이다. 따라서, 유효량은 조성물을 투여하는 의사에 의해 가장 잘 결정되고, 적절한 투약량은 당업자에 의해 용이하게 결정될 수 있다. 통합 및 발현을 위한 충분한 시간(예를 들어, 전형적으로 4 내지 15일) 동안 충분한 시간을 허용한 후에, 치료적 폴리펩타이드의 혈청 또는 다른 조직 수준의 분석 및 투여 전 초기 수준과의 비교는 투여될 양이 정확한 범위 내에서 너무 낮든 또는 너무 높든 결정될 것이다. 초기 및 후속적 투여를 위한 적합한 요법은 또한 가변적이지만, 필요하다면 초기 투여 다음에 후속적 투여에 의해 전형화된다. 후속적 투여는 1일 내지 1년 내지 몇년 범위의 가변적 간격으로 투여될 수 있다. 당업자는 적절한 면역억제 기법이 전달 벡터의 면역억제에 의한 형질도입의 저해 또는 차단을 회피하도록 권장될 수 있다는 것을 인식할 것이다, 예를 들어, 문헌[Vilquin *et al.*, (1995) Human Gene Ther., 6:1391-1401] 참조.

[0180] 생체외와 생체내 투여 둘 다를 위한 제형은 액체 또는 유화된 액체 중에서 혼탁액을 포함한다. 활성 성분은 종종 약제학적으로 허용 가능하고, 활성 성분에 적합한 부형제와 혼합된다. 적합한 부형제는, 예를 들어, 물, 식염수, 텍스트로스, 글리세롤, 에탄올 등 및 이들의 조합을 포함한다. 추가로, 조성물은 소량의 보조 물질, 예컨대, 습윤 또는 유화제, pH 완충제, 안정제 또는 약제학적 조성물의 유효성을 향상시키는 다른 시약을 함유할 수 있다.

[0181] 세포

- [0182] 또한 본 명세서에서 내인성 BCL11A 인핸서 서열이 본 명세서에 기재된 뉴클레아제에 의해 변형되는 세포 및 세포주가 기재된다(표 1). 변형은, 예를 들어, 세포의 야생형 서열에 비교될 수 있다. 세포 또는 세포주는 변형을 위해 이형접합적 또는 동형접합적일 수 있다. BCL11A 서열에 대한 변형은 삽입결실을 포함할 수 있다.
- [0183] 변형은, 예를 들어, 절단 부위(들)의 상류 또는 하류의 1 내지 300개(또는 그 사이의 임의의 값) 염기쌍 내에서, 더 바람직하게는 결합 부위(들), 절단 부위(들), 또는 결합 부위(들)와 절단 부위(들) 중 한 측면의 1 내지 100개의 염기쌍(또는 그 사이의 임의의 값) 내에서, 훨씬 더 바람직하게는 결합 부위(들), 절단 부위(들), 또는 결합 부위(들) 및 절단 부위(들) 중 하나의 측면 상에서 1 내지 50개의 염기쌍(또는 그 사이의 임의의 값) 내에서, 바람직하게는 뉴클레아제(들) 결합 부위(들), 절단 부위(들) 및 결합 부위(들)와 절단 부위(들)의 조합에서 또는 근처에 있다. 특정 실시형태에서, 변형은 BCL11A 인핸서의 "+58" 영역에 또는 근처에, 예를 들어, 표 1의 임의의 첫 번째 열에 나타낸 뉴클레아제 결합 부위에 또는 근처에 있다.
- [0184] 임의의 세포 또는 세포주는 변형된, 예를 들어 줄기 세포, 예를 들어 배아 줄기 세포, 유도 만능 줄기 세포, 조혈 줄기 세포, 신경 줄기 세포 및 중간엽 줄기 세포일 수 있다. 본 명세서에 기재된 바와 같은 세포의 다른 비제한적 예는 T-세포(예를 들어, CD4+, CD3+, CD8+ 등); 수지상 세포; B-세포를 포함한다. 복수의 또는 완전히 분화된 세포를 포함하는 줄기세포의 후손이 또한 제공된다(예를 들어, RBC 또는 RBC 전구체 세포). 변형된 BCL11A 서열을 포함하는 다른 세포주의 비제한적 예는 COS, CHO(예를 들어, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NS0, SP2/0-Ag14, HeLa, HEK293(예를 들어, HEK293-F, HEK293-H, HEK293-T) 및 perC6 세포뿐만 아니라 곤충 세포, 예컨대 스포돕테라파기페르다(*Spodoptera frugiperda*: Sf), 또는 진균 세포, 예컨대 사카로마이세스(*Saccharomyces*), 퍼키아(*Pichia*) 및 쉬조사카로마이세스(*Schizosaccharomyces*)를 포함한다.
- [0185] 본 명세서에 기재된 바와 같은 세포는, 예를 들어, 생체외 요법에 의해 장애를 치료하거나 또는 예방하는 데 유용하다. 뉴클레아제-변형 세포는 확장될 수 있고, 이어서, 표준 기법을 이용하여 환자에게 재도입될 수 있다. 예를 들어, 문헌[Tebas *et al.* (2014) *New Eng J Med* 370(10):901] 참조. 줄기 세포의 경우에, 대상체에게 주입 후에, 이들 전구체의 기능성 이식유전자를 발현시키는 세포로의 생체내 분화가 또한 일어난다. 본 명세서에 기재된 바와 같은 세포를 포함하는 약제학적 조성물이 또한 제공된다. 추가로, 세포는 환자에 대한 투여 전에 동결보존될 수 있다.
- [0186] 본 명세서에 개시된 임의의 변형된 세포 또는 세포주는 감마 글로빈의 증가된 발현을 나타낼 수 있다. 본 명세서에 기재된 바와 같은 유전자 변형된 세포를 포함하는 약제학적 조성물과 같은 조성물이 또한 제공된다.
- [0187] **적용분야**
- [0188] 본 명세서에 개시된 방법 및 조성물은 단백질 발현을 변형시키거나, 또는 유전자 질환, 예컨대 겹상 적혈구병 또는 탈라세미아에서 발현되는 단백질을 암호화하는 비정상적 유전자 서열을 보정하기 위한 것이다. 따라서, 상기 방법 및 조성물은 이러한 유전자 질환의 치료 및 예방을 제공한다. 예를 들어 줄기 세포의 계놈 편집은 비정상적 유전자의 보정, 야생형 유전자의 삽입 또는 내인성 유전자의 발현의 변화를 위해 사용될 수 있다. 비제한적 예로서, 예를 들어, 적어도 하나의 글로빈(예를 들어, α 글로빈, γ 글로빈, β 글로빈과 이들의 조합)을 암호화하는 야생형 유전자는 세포 내로(예를 들어, 본 명세서에 기재된 하나 이상의 뉴클레아제를 이용하여 내인성 BCL11A 인핸서 서열 내로) 삽입되어 세포 내에서 결핍 또는 결여된 글로빈 단백질을 제공하고, 이에 의해 불완전한 글로빈 발현에 의해 야기되는 유전자 질환, 예를 들어, 이상혈색소증을 치료할 수 있다. 대안적으로 또는 추가로, 적절한 공여자의 투여에 의한 또는 투여가 없는 계놈 편집은 불완전한 내인성 유전자를 보정할 수 있다(유전자의 발현을 회복하기 위한 또는 유전자 질환, 예를 들어, 겹상 적혈구병을 치료하기 위한, 임의의 직접 또는 간접 글로빈 조절 유전자의 넉아웃 또는 변경(과발현 또는 억제)하기 위한(예를 들어, γ 글로빈-조절 유전자 BCL11A 또는 BCL11A-조절자 KLF1의 비활성화) α - 또는 β -헤모글로빈에서의 점 돌연변이 보정). 구체적으로, 본 발명의 방법 및 조성물은 이상혈색소증의 치료 또는 예방에서의 용도를 가진다.
- [0189] 본 발명의 뉴클레아제는 적혈구 생성 동안 BCL11A의 발현에 필요한 것으로 알려진 BCL11A 인핸서 영역으로 표적화되고, 그리고 따라서 감마 글로빈 발현을 하향 조절한다. 이 인핸서 영역의 변형은 증가된 감마 글로빈 발현을 갖는 적혈구를 초래할 수 있고, 따라서, 겹상 적혈구병 또는 베타 탈라세미아의 치료 또는 예방에 도움을 줄 수 있다.
- [0190] 다음의 실시예는 뉴클레아제가 아연 평거 뉴클레아제(ZFN)를 포함하는 본 개시내용의 예시적 실시형태에 관한 것이다. 이는 단지 예시의 목적을 위한 것이며, 다른 뉴클레아제는, 예를 들어 조작된 DNA-결합 도메인을 갖는

TtAgo 및 CRISPR/Cas 시스템, 귀소 엔도뉴클레아제(메가뉴클레아제), 조작된 귀소 엔도뉴클레아제(메가뉴클레아제) DNA-결합 도메인의 천연 유래의 융합(귀소 엔도뉴클레아제(메가뉴클레아제)와 조작된 DNA-결합 도메인의 조합을 포함) 및 조작된 귀소 엔도뉴클레아제(메가뉴클레아제) DNA-결합 도메인, 및 이종성 절단 도메인의 천연 유래의 융합, 메가뉴클레아제와 TALE 단백질의 융합(이종성 절단 도메인과 메가뉴클레아제와 TALE 단백질의 융합의 조합을 포함하는 단백질)이 사용될 수 있다는 것이 인식될 것이다.

[0191]

실시예

[0192]

실시예 1: 아연 평거 뉴클레아제의 조립

[0193]

ZFN을 인간 BCL11A 유전자에 대해 조립하였고, 상기 기재한 형질감염 세포로부터 단리된 DNA의 심층 서열분석에 의해 활성을 시험하였다. 인해서 영역의 +58 영역에 특이적인 ZFN을 기재한 바와 같이 제조하였다. ZFN 쌍 51857/51949는 이전에 기재하였다(WO 2016/183298 참조).

[0194]

실시예 2: 표적을 벗어난 분석

[0195]

ZFN 쌍에 의한 표적을 벗어난 절단을 분석하기 위해, 2단계 편향 특이적 분석을 수행하였다. 제1 단계에서(도 1), 각각의 ZFN에 대한 표적을 벗어난 부위 후보를 문현[Tsai *et al.* ((2015), *Nat Biotechnol* 33(2):187-197. doi: 10.1038/nbt.3117)]에 기재된 것과 유사한 절차를 이용하는 올리고뉴클레오타이드 이중가닥 통합 부위 분석을 통해 동정하였다.

[0196]

올리고뉴클레오타이드 이중가닥 통합 부위 분석은 표적 세포 내로 뉴클레아제와 이중가닥 DNA의 짧은 세그먼트의 공동도입은 NHEJ DNA 수선 경로를 통한 게놈 절단 사건의 분획 수선 동안 이중가닥 통합을 초래한다는 관찰에 기반한다(Orlando *et al.*, (2010), *Nucleic Acids Res*, 38(15) e152. doi: 10.1093/nar/gkq512; Gabriel *et al.*, (2011), *Nat Biotechnol*. 2011 Aug 7;29(9):816-23. doi: 10.1038/nbt.1948; Tsai *et al.*, 이하 참조). 통합 시, 이중 가닥은 절단 사건의 영구적 태그를 제공한다. 이어서, 통합 부위는 전단 게놈 DNA에 대한 올리고뉴클레오타이드 어댑터의 결찰, 다음에 25주기의 네스티드 PCR, 및 얻어진 공여자-게놈 접합의 심층 서열 분석의 2회 라운드를 통해 동정한다. 이 분석은 게놈 내에서 모든 잠재적 통합 부위의 평가를 허용한다.

[0197]

공여자 전달, ZFN 발현 및 공여자 통합을 최대화하기 위해 K562 세포에서 통합 부위 분석을 수행하였다. 게다가, K562 세포는 빠르게(대략 24시간의 배가시간) 분해되기 때문에, 그들은 세포 표적을 절단하는 ZFN의 능력에 대해 최소 후성적 제한을 부과하는 것으로 예상된다. ZFN의 최대 표적상의 활성을 대해 최적화된 아맥사(Amaxa) 셔틀 및 세팅을 이용하여 세포(2×10^5 개)를 $0.47\mu\text{g}$ 의 올리고뉴클레오타이드 이중가닥 공여자 및 400ng 의 각각의 ZFN-암호화 mRNA으로 전기천공시켰다. 올리고와 mRNA의 각각의 조합에 대해 4개의 복제물 샘플을 제조하였다. 형질감염 후 제7일에, 게놈 DNA를 각각의 샘플에 대해 단리시키고(퀴아젠(Qiagen) DNeasy 혈액 및 조직 키트), 도 1에서 약술하는 증식 프로토콜을 위한 유입으로서 400 ng (133000 반수체 게놈)을 사용하였다. 이어서, 샘플을 기재한 바와 같이 본질적으로 가공하였다(Tsai *et al.* 이하 참조). 최종 산물을 풀링하고 나서, 정량화하고, 그리고 앰플리콘의 각각의 말단 상에서 샘플 바코드를 검출하기 위해 쌍별-말단 150bp 판독 및 8bp/16bp 이중 색인 판독으로 v2 300 주기 서열분석 키트를 이용하여 MiSeq 기기(일루미나) 상에서 서열분석하였다.

[0198]

표적을 벗어난 부위 후보 목록을 생성하기 위해, 정확한 프라이밍 서열에 대해 서열분석 데이터를 여과한 후에, 어댑터 서열을 트리밍(trimming)하고 게놈에 대해 맵핑하였다. 다음에, 접합 좌표를 맵핑하고 나서, 이중가닥-게놈 접합뿐만 아니라 DNA 전단에 의해 야기되는 과순 위치를 사용하여 별개의 통합 사건을 동정하였다. 이어서, 통합 사건을 가공하여 게놈 내에서 근위에 통합 클러스터를 동정하였다(서로 100bp 내의 최소 4개의 별개 통합 사건, 모든 복제물에 걸친 합계). hg38 조립체(즉, hg38 내 chrUn)에서 맵핑 가능하지 않은 콘티(contig) 상에 남아있는 클러스터를 추가 분석으로부터 제거하였다. 이전의 경험은 이들이 증폭 인공물이 되는 것을 나타냈기 때문에 반복적 좌위(클러스터 내 모든 서열에 걸쳐 게놈에 대한 3개 이상의 히트의 중앙값)에 대한 클러스터 맵핑이 또한 제거되었다. 남아있는 클러스터는, 그들이 (총 4개 중에서) 적어도 2개의 복제물 ZFN 처리 샘플로부터 유래되고, 대조군에 비해 ZFN 처리 샘플에서 5배 초과의 과량의 통합 사건을 나타내었다면, 후보 ZFN 절단 부위로서 스코어링하였다. 후보 절단 부위를 ZFN 처리 샘플에서 독특한 통합의 총 수에 따라 랭크하였다. 이 분석을 통해 동정한 후보 좌위를 구성요소 계수에 따라 랭크한 ZFN 쌍 51857/51949에 대해 도 2에서 제공한다.

[0199]

실시예 3: ZFN의 최적화

- [0200] 표적을 벗어난 절단을 감소시키기 위해, 비특이적 인산염 접촉이 선택적으로 제거되어 표적을 벗어난 절단의 전반적인 억제를 유발하는 뉴클레아제 최적화에 대한 전략(Guilingher *et al.* (2014) *Nat Methods.* 11(4):429-35. doi: 10.1038/nmeth.2845; Kleinstiver *et al.* (2016) *Nature* 529(7587):490-5. doi: 10.1038/nature16526; Slaymaker *et al.* (2016) *Science* 351(6268):84-8. doi: 10.1126/science.aad5227)을 채택하였다(미국 기출원 특허 제62/443,981호 및 제62/378,978호 참조). DNA의 인산염 결격과 상호작용하는 아연 평거 프레임워크 내의 중요한 위치에서(Pavletich and Pabo, (1991) *Science* 252(5007):809-17; Elrod-Erickson *et al.* (1996) *Structure* 4(10):1171-80)(도 3a 내지 도 3b)뿐만 아니라 인산염 접촉을 만드는 것으로 예측되는 우측 ZFN FokI 도메인 내 단일 위치에서(도 3c) 아미노산 치환이 이루어졌다.
- [0201] 전달비의 최적화를 가능하게 하는 두 분리된 mRNA로부터의 각각의 ZFN의 독립적 발현을 허용함으로써 특이성은 추가로 개선되었다. 이들 노력은 본래의 하나와 고도로 관련된 최적화된 ZFN 쌍을 수득하였는데, 이는 DNA 인산염 결격과의 상호작용 에너지를 감소시키지만, 서열 특이적 염기 인식에 최소로 영향을 미치거나 또는 영향을 미치지 않는 치환만큼 상이하였다. 이와 일치되게, 통합 부위 분석은 본래의 51857/51949 쌍에 대한 ZFN 절단의 잠재적 표적을 위한 455 좌위를 수득하였다. 최적화된 쌍에 대해, 이 분석에 의한 ZFN 절단의 잠재적 표적(총 72개)으로서 추가 시험을 위해 훨씬 더 소수의 좌위를 동정하였다. 쌍 둘 다에 대해, BCL11A 인핸서 내에서의 의도된 표적은 상부 랭크된 좌위였다. 게다가, 훨씬 더 높은 분율의 통합 사건을 최적화된 쌍에 대한 BCL11A 인핸서에서 주목하였고, 이는 그의 더 큰 특이성과 일치된다.
- [0202] 서열 데이터 가공 파이프라인을 정함에 있어서, 그것을 필터링하는 대신에 실현하기 쉬운 표적을 벗어난 좌위의 다수의 후보로서 포함하는 측면 상의 오류에 대해 중요한 매개변수를 보존적으로 선택하였다는 것을 주목하는 것은 중요하다. 이는 심지어 거짓 양성이 되는 훨씬 다수를 받아들이는 대가로 최적화된 ZFN에 대한 진실된 절단 부위를 나타낼 수 있는 모든 좌위를 후속 삽입결실 연구에서 동정하고 시험하였다는 것을 보장하기 위해 행하였다. 분석의 제1 단계는 각각의 ZFN 쌍에 대한 큰 세트의 후보 좌위를 수득하였으며, 이 중 대다수는 (특히 최적화된 ZFN에 대해) 후속 삽입결실 연구에서 절단에 대해 음성으로 증명된 배경 사건보다는 진정한 표적을 벗어난 절단 부위를 나타내지 않는다는 것이 예상되었다.
- [0203] 분석의 제2 단계에서, 통합 부위 분석을 통해 동정한 표적을 벗어난 좌위의 후보를 ZFN-처리 CD34+ HSPC에서 변형의 증거(예를 들어, 삽입결실의 존재)에 대해 선별하였다.
- [0204] 특히, 동원된 말초 혈액으로부터 유래된 인간 CD34+ HSPC를 RNA 형질감염을 위해 임상 규모 및 임상 병태를 이용하여 본래의 그리고 최적화된 ZFN 쌍으로 처리하였다(본래의 ZFN 쌍에 대해 120 μ g/ml의 mRNA 및 최적화된 쌍에 대해 100 μ g/ml의 mRNA). 형질감염 후 2일에 계놈 DNA를 단리시킨 후, 표적을 벗어난 좌위 후보의 PCR 증폭 및 심층 서열분석으로 삽입결실 수준을 정량화하였다. 본래의 그리고 최적화된 ZFN 쌍 둘 다에 대해, 본래의 ZFN을 이용하는 초기 연구에서 다른 방법을 통해 동정된 더 소수의 표적을 벗어난 부위 후보와 함께 137개의 표적을 벗어난 좌위 후보의 동일한 세트를 이 단계에서 선별하였다.
- [0205] 결과는 최적화된 ZFN이 본래의 쌍보다 현저히 더 특이적이라는 것을 나타내었다. 이는 ZFN 절단의 증거에 대해 양성으로 스코어링된 좌위의 수로부터(본래의 쌍에 대해 52 대 최적화된 쌍에 대해 3)뿐만 아니라, 최적화된 쌍이 훨씬 더 낮은 관찰된 삽입결실 수준으로부터 명확하다. 도 4는 본 연구에서 ZFN의 증거를 나타내는 모든 좌위에서 삽입결실의 플롯을 나타낸다(y 축의 log 규모를 주목). 모든 이러한 좌위에 걸친 표적을 벗어난 삽입결실의 응집은 300배의 표적을 벗어난 활성의 감소(본래의 쌍에 대해 46.5% 응집의 표적을 벗어난 삽입결실, 대 최적화된 쌍에 대해 0.15%의 표적을 벗어난 삽입결실)를 나타낸다. 표적을 벗어난 활성의 이러한 감소는 의도된 표적 부위에서 활성의 임의의 상실 없이 달성되었다(본래의 쌍에 대해 72.5% 삽입결실 대 최적화된 ZFN에 대해 81.9%). 이들 연구에서, 본래의 쌍(또는 모 쌍)은 51857/51949인 반면, 최적화된 ZFN 쌍은 63014/65722였다(이하 참조).
- [0206] 이하의 표 1에서 뉴클레아제 설계를 나타낸다:

표 1

+58 BCL11A 인핸서 영역에 특이적인 ZFN 쌍

SBS # (표적 부위, 5'-3')	설계 [나선 서열, 서열번호]						링커 [Fok 돌연변이 체]
	[핑거 골격에 대한 돌연변이]						
	F1	F2	F3	F4	F5	F6	
좌측 상대							
51857 aaAGCAACTG TTAGCTTGCA Ctagacta (서열번호 1)	DQSNLRA (서열번호 19)	RNFSLTM (서열번호 20)	STGNLTN (서열번호 21)	TSGSLTR (서열번호 22)	DQSNLRA (서열번호 19)	AQCCLFH (서열번호 23)	L7c5
	없음	없음	없음	없음	없음	없음	ELD
63014 aaAGCAACTG TTAGCTTGCA Ctagacta (서열번호 1)	DQSNLRA (서열번호 19)	RNFSLTM (서열번호 20)	STGNLTN (서열번호 21)	TSGSLTR (서열번호 22)	DQSNLRA (서열번호 19)	AQCCLFH (서열번호 23)	L7c5
	Qm5	없음	Qm5	없음	Qm5	없음	ELD
65459 aaAGCAACTG TTAGCTTGCA Ctagacta (서열번호 1)	DQSNLRA (서열번호 19)	RNFSLTM (서열번호 20)	STGNLTN (서열번호 21)	TSGSLTR (서열번호 22)	DQSNLRA (서열번호 19)	AQCCLFH (서열번호 23)	L7c5
	Qm14Qm5	없음	Qm5	없음	Qm5	없음	ELD
우측 상대							
51949 caCAGGCTCC AGGAAGGgtt tggcctct (서열번호 2)	RNDHRTT (서열번호 24)	QKAHLIR (서열번호 25)	QKGTLGE (서열번호 26)	RGRDLSR (서열번호 26)	RRDNLHS (서열번호 27)	N/A	L0
	없음	없음	없음	없음	없음	N/A	KKR
65722 caCAGGCTCC AGGAAGGgtt tggcctct (서열번호 2)	RNDHRTT (서열번호 24)	QKAHLIR (서열번호 25)	QKGTLGE (서열번호 26)	RGRDLSR (서열번호 26)	RRDNLHS (서열번호 27)	N/A	L0
	Qm5	Qm5	없음	Qm5	없음	N/A	KKR K525S
65526 caCAGGCTCC AGGAAGGgtt tggcctct (서열번호 2)	RNDHRTT (서열번호 24)	QKAHLIR (서열번호 25)	QKGTLGE (서열번호 26)	RGRDLSR (서열번호 26)	RRDNLHS (서열번호 27)	N/A	L0
	Qm5	Qm5	없음	Qm5	Qm5	N/A	KKR R416S
65549 caCAGGCTCC AGGAAGGgtt tggcctct (서열번호 2)	RNDHRTT (서열번호 24)	QKAHLIR (서열번호 25)	QKGTLGE (서열번호 26)	RGRDLSR (서열번호 26)	RRDNLHS (서열번호 27)	N/A	L0
	Qm5	Qm5	Qm5	Qm5	없음	N/A	KKR K525S
65550 caCAGGCTCC AGGAAGGgtt tggcctct (서열번호 2)	RNDHRTT (서열번호 24)	QKAHLIR (서열번호 25)	QKGTLGE (서열번호 26)	RGRDLSR (서열번호 26)	RRDNLHS (서열번호 27)	N/A	L0
	Qm5	Qm5	없음	Qm5	Qm5	N/A	KKR K525S

[0207]

표 1은 각각의 ZFN에 관한 정보의 특성규명을 나타낸다. 좌측에서부터 시작하여, SBS 번호(예를 들어, 51857)를 DNA 표적에 따라 나타내는데, ZFN은 SBS 번호 아래로 나타내도록 결합한다. 다음에, 펩거 1 내지 6 또는 1 내지 5에 대한 아미노산 인식 나선 설계를 나타낸다(표 1의 2열로 다시 분할됨). 또한 미국 가출원 특허 제 62/378,978호 및 제62/443,981호에 기재한 바와 같이 표시된 펩거의 ZFP 골격 서열로 이루어진 표 1에서 적절한 나선 설계 하에서 돌연변이를 나타낸다. 표 1에서 사용하는 명명법에서, "Qm5"는 표시된 펩거의 위치 - 5에서(-1 내지 +6으로 넘버링된 나선에 대해), 이 위치에서 알기난은 글루타민(Q)으로 대체된다는 것을 의미하는 반면, "Qm14"는 -14에서 정상적으로 존재하는 알기닌(R)이 글루타민(Q)으로 대체된다는 것을 의미한다. "없음"은 인식 나선 영역 외부에서 변화 없음을 나타낸다. 따라서, 예를 들어, SBS# 63014는 펩거 1, 3 및 5에서 Qm5 돌연변이를 포함하는 반면, 펩거 2, 4 및 6은 아연 펩거 골격(예를 들어, 인식 나선 영역 외부의 아연 펩거 서열)에 대한 돌연변이를 갖지 않는다.

[0209]

최종적으로, 표 1의 가장 우측 열은 FokI 절단 도메인(예를 들어, "L7c5"(LRGSISRARPLNPHP (서열번호 5), 예를

들어 미국 특허 제9,567,609호에 기재된 바와 같음)에 DNA 결합 도메인을 연결하기 위해 사용되는 링커를 열의 상부선 상에 나타내고, FokI 인산염 접촉 돌연변이 및 이량체화 돌연변이 부위를 링커 표기 아래의 박스에 나타낸다. 구체적으로, FokI 돌연변이체 박스의 상부선 상에 이량체화 도메인에서 발견되는 돌연변이 유형을 표시한다(예를 들어, 미국 특허 제8,962,281호에 기재한 바와 같은, 예를 들어, ELD 또는 KKR). 이하에서, 이량체화 돌연변이체 표기를 하부에 나타낸 비-특이적 인산염 접촉을 제거하도록 이루어진 FokI 도메인에 존재하는 임의의 돌연변이를 나타낸다(예를 들어, K525S 또는 R416S, 여기서, 아미노산 위치 525 또는 416에서의 세린 잔기는 미국 가출원 특허 제62/378,978호 및 제62/443,981호에 각각 기재하는 바와 같이 라이신 또는 알기닌으로 치환함). 따라서, 예를 들어, SBS# 63014에서, 링커는 L7c5 링커이고, FokI 절단 도메인은 ELD 이량체화 돌연변이체 및 인산염 접촉 돌연변이 없음을 포함한다. 추가로, SBS# 65722에 대해, 링커는 L0 링커(LRGSQLVKS(서열번호 6), 또한 '표준' 링커로서 지칭되며, 미국 특허 제9,567,609호 참조)이고, FokI 절단 도메인은 KKR 이량체화 돌연변이 및 K525S FokI 인산염 접촉 돌연변이를 포함한다.

[0210] 모든 ZFN은 기능성(이하의 실시예 4에 기재한 바와 같은 삽입결실에 대해 분석함으로써 결정된 바와 같은 절단 활성)에 대해 시험하였고, 활성인 것을 발견하였다.

[0211] 더 나아가, ZFN 설계가 가장 특이적이라는 것을 결정하기 위해, 본래의 ZFN 쌍에 의한 표적을 벗어난 절단의 알려진 부위의 삽입결실 분석을 ZFN-처리 CD34+ HSPC에서 수행하였다. 이를 달성하기 위해, 동원 말초 혈액으로부터 유래된 인간 CD34+ HSPC를 임상 조건 및 mRNA 농도(본래의 ZFN 쌍에 대해 $120\mu\text{g}/\text{mL}$ 및 최적화된 쌍에 대해 $100\mu\text{g}/\text{mL}$)를 이용하여 본래의 그리고 최적화된 ZFN 쌍으로 처리하였다. 이들 세포 및 비처리 대조군으로부터 형질감염 후 2일에 게놈 DNA를 단리시킨 후에, 각각의 후보 좌위의 PCR 증폭 및 심풍 서열분석으로 삽입결실 수준을 정량화하였다.

[0212] 각각의 좌위에서의 변형 수준을 300 주기 카트리지를 이용하는 일루미나(Illumina) MiSeq 상에서 쌍별-말단 심층 서열분석함으로써 결정하였다. 쌍별 서열을 병합하고, SeqPrep를 통해 트리밍한 어댑터를 모든 염기에 걸쳐 15 이상의 품질 스코어로 필터링하고, 이어서, 인간 게놈(hg38 조립체)에 맵핑하였다. 부정확한 좌위로 맵핑된 서열을 버렸다. 프라이머-이량체 산물을 최소화하기 위해 70 bp 초과 또는 70% 초과만큼 야생형 앰플리콘보다 더 짧은 서열을 제거하였다. 삽입결실을 맵핑하기 위해 표적 앰플리콘과 각각의 MiSeq 판독 간에 니들만-분위(Needleman-Wunsch) 정렬(Needleman and Wunsch, (1970), *J Mol Biol* 48(3):443-53))을 수행하였다. 실제 사건보다 적게 계수되는 것을 피하기 위해 삽입결실 1bp 길이를 허용한 것을 제외하고 정렬된 서열에서의 삽입결실을 문헌[Gabriel et al. 2011](이하 참조)에 기재하는 바와 같이 정의하였다. 좌위 중 하나의 분획은 서열을 증폭시키지 않았거나 또는 서열분석하지 않았거나, 또는 고배경(대조군 샘플에서 1% 초과의 변형) 또는 불충분한 서열분석 깊이(10000개 미만의 판독)에 기인하여 분석으로부터 거부되었다는 것을 주목한다. 이 분석의 결과 및 '모' 51857/51949 ZFN 쌍에 대한 비교를 이하의 표 2에 제공한다.

표 2

표적을 벗어난 절단 분석

좌	우	μg RNA L:R	BCL11A 삽입결실%	표적을 벗어난 좌위의 수			삽입결실-양성 OT 좌위의 수		
				표적	PCR'd	분석 가능	P<0.05	수동	모체에 있는가? 포획/확인
51857	51949	60:60	73.0	23	21	17	15	17	-
63014	65722	60:15	82.2	31	24	14	0	4	4/3
63014	65526	60:15	81.2	23	22	9	3	4	4/4
63014	65527	60:60	81.4	30	24	10	4	4	4/2
63014	65549	60:60	80.0	30	24	13	0	2	2/1
63014	65550	60:60	79.8	30	24	9	0	1	1/1
65459	65526	60:15	76.9	23	19	14	0	2	2/2

[0213] 실시예 4: 인간 CD34+ 세포에서 ZFN의 활성

[0215] 시험관내 검사를 위해, 뉴클레아제를 CD34+ 세포에서 시험하였다. ZFN을 mRNA로서 공급하였고, 여기서, mRNA를 다음과 같이 시험관내에서 생성하였다: ZFN을 암호화하는 유전자를 포함하는 플라스미드를 선형화하고 나서, 엠메세지 엠머신(mMessage mMachine)(등록상표) T7 울트라 키트(엠비온(Ambion)/어플라이드 바이오시스템즈(Applied Biosystems))를 이용하여 시험관내 mRNA 전사에 대해 사용하였다. 이어서, RNeasy(등록상표) 미니 키트(큐아젠)를 이용하여 mRNA를 정제하였다.

[0216] CD34+ 세포를 동원 말초 혈액으로부터 단리시키고 나서, 페니실린, 스트렙토마이신 및 글루타민뿐만 아니라 스템스판(StemSpan) CC110으로 보충한 X-VIVO 10 배지에서 유지하고, 37°C, 및 5% CO₂에서 인큐베이션시켰다. 단리후 또는 해동후 48시간에 세포를 형질감염시켰다. 소량의 분취액을 PBS(코닝(Corning)) 중에서 트립판 블루 용액 0.4%(w/v)과 1:1로 혼합하고 나서, 세포 수를 TC20 자동 세포 계수기(바이오-라드(Bio-Rad)) 상에서 결정하였다.

[0217] 대규모 형질감염을 위해, 세포를 MaxCyte 전기천공 완충제(맥스사이트(Maxcyte))로 세척하고 나서, 100μl로 전기천공 완충제 중에서 ml 당 3 내지 5e7개의 세포로 재현탁시켰다. 전형적으로, 60μg/ml 내지 120μg/ml의 mRNA 농도를 사용하여 후보 ZFN 세트를 선별하였다. 이어서, 세포를 30°C에서 18시간 동안 ml 당 3e6개의 세포로 성장 배지에서 성장시켰고, 이어서, 37°C에서 추가 24시간 동안 ml 당 1e6 세포로 희석시켰다. 절단 활성의 결정을 위해, 형질감염 후 2 내지 3일에 게놈 DNA를 단리시키고 나서, MiSeq 서열분석기(일루미나) 상에서 심층 서열분석을 통해 BCL11A 인핸서 좌위에서 유전자 변형 수준을 측정하였다.

[0218] 표 1로부터의 ZFN 쌍을 CD34+ 세포에서 시험하였고, 활성 결과를 이하의 표 3에 나타낸다.

표 3

BCL11A 표적에 대한 ZFN 쌍의 활성

우측 ZFN	좌측 ZFN	R ZFN 농도 (μg)	L ZFN 농도 (μg)	삽입 결실 (%)
51857	51949	60	60	72.98
63014	65722	15	60	80.62
63014	65722	60	15	82.19
63014	65526	60	15	81.22
63014	65527	60	60	81.43
63014	65549	60	60	79.82
63014	65550	60	60	79.96
GFP 대조군				0.07

[0219]

[0220] 적혈구 분화 전에 CD34+ 세포에서 뉴클레아제 활성을 분석하는 것에 추가로, 편집된 세포는 또한 시험관내에서 적혈구 세포로 분화되었다. 이어지는 프로토콜은 문헌[Giarratana *et al.*, ((2011) *Blood* 120 (15):2945-53)]에 기반하였다. 간략하게, 이하의 프로토콜을 따랐다:

[0221] 제0일 내지 제7일: 4x10⁴ 개의 CD34+ 세포를 10⁻⁶ M 하이드로코티손, 100ng/ml 줄기 세포 인자(SCF), 5ng/ml IL 3, 및 3 IU/ml 에리트로포이에틴(EPO)의 존재 하에 분화 배지(EDM)(이스코브 변형 둘베코 배지(Iscove's Modified Dulbecco's Medium)[IMDM], 330μg/ml 트랜스페린, 10μg/ml 인간 인슐린, 2U/ml 혜파린 나트륨, 5% 인간 AB+ 혈장)에서 2x104개/ml의 밀도로 배양시켰다.

[0222] 제4일: SCF, IL-3, EPO, 및 하이드로코티손을 함유하는 신선한 EDM에서 세포를 재현탁시켰다.

[0223] 제7일 내지 제11일: SCF 및 EPO로 보충한 EDM의 신선한 ml 당 1.5 × 10⁵개의 세포 밀도로 세포를 재현탁시켰다.

[0224] 제11일 내지 제21일: 제11일에, 세포를 EPO로 보충한 신선한 EDM 중에서 1 × 10⁶개/ml로 재플레이팅시켰다. 세포를 후속적으로 제14일에 5 × 10⁶개/ml로 동일한 배지에서 재플레이팅시켰다. 제21일에 배양의 종결까지 세포 생존도가 하락하기 시작할 때 제14일 내지 제18일의 이 시간 기간 동안 성장은 안정 상태를 유지하였다.

- [0225] 700 밀리초로 설정한 형광 채널 1(AO) 및 5000 밀리초로 설정한 형광 채널 2(PI)를 이용하는 AOPi 적혈구 분석 방식과 함께 넥스셀롬 바이오사이언스 셀로미터 K2(Nexcelom Bioscience Cellometer K2)를 이용하여 아크리딘 오렌지 확실성 및 요오드화프로파듬 배제(AOPi)를 측정함으로써 파종 시 그리고 분화 전체적으로 세포 계수를 취하였다.
- [0226] 다음의 프로토콜을 이용하여 분화 제21일에 제핵 세포의 백분율을 결정하였다. 제핵 비율은 비형질감염 대조군 및 ZFN-형질감염 샘플에서 이들 두 그룹으로부터 59 내지 63%의 백분율로 비슷하였다:
1. 세포 계수
 2. 100,000개의 세포, 450 x g에서 스핀 다운, 5분, RT.
 3. 50 μ l PBS-BSA + 1 μ l GlyA-FITC(DAKO) 중에서 재현탁시킴.
 4. 냉장고에서 15분 동안 염색.
 5. 1ml PBS-BSA를 첨가하고 나서, 교반 후, 스핀 다운시켰다.
 6. 250 μ l의 PBS-BSA-NucRed(2방울의 NucRed/ml) 중에서 재현탁시켰다.
 7. NucRed에 대해 APC를 이용하여 FACS Canto 상에서 획득함.
 8. 유핵 적혈구 세포는 GlyA 양성 NucRed 음성/저분획에 있을 것이고, 적아구는 이중 GlyA-NucRed 양성 분획에 있을 것이다.
- [0235] a) 전기천공법 48시간 후에 b) 시험관내 분화가 시작되었을 때 세포 해동일에 그리고 c) 시험관내 적혈구의 제14일에 채취한 DNA 샘플에서 MiSeq 심층 서열분석에 의해 BCL11A 유전자 변형을 측정하였다. 21일 동안 분화를 수행하였지만, DNA 분석 동안 제14일 시점을 선택하였는데, 이는 DNA 회수의 손실을 초래한 적혈구 세포의 대형 분획의 제핵 전이기 때문이다. BCL11A 인핸서에서 관찰된 변형 백분율을 형질감염 조건의 상세한 설명과 함께 표 4에 열거한다.

표 4

MiSeq 분석에 의한 BCL11A 유전자 변형 수준

CD34+ 세포 프레파라트	형질감염	BCL11A 유전자 변형 (%)		
		TF 후 제2일	해동 후	분화 제14일
프레파라트 #1	80 μ g/mL 63014 + 20 μ g/mL 65722	78.6	81.4	72.0
	비형질감염	0.1	0.2	0.2
프레파라트 #2	80 μ g/mL 63014 + 20 μ g/mL 65722	75.4	77.3	72.3
	비형질감염	0.1	0.1	0.0

- [0236] [0237] 이들 데이터는 최적화된 쌍 63014 및 65722 mRNA에 의한 CD34+ 세포 형질감염이 BCL11A 인핸서 표적 부위에서 매우 효율적인 유전자 변형을 야기하고(변형된 대립유전자의 75% 초과), 세포의 냉동 및 해동 후에 그리고 적혈구 분화 후에 변형이 매우 잘 유지된다는 것(변형의 90% 초과 체류)을 나타낸다.
- [0238] 추가 분석을 위해 ZFN 쌍 63014/65722를 선택하였다. 이들 ZFN에 대한 아미노산 서열을 이하에 나타내며, 여기서, 각각은 핵 국소화 신호(nuclear localization signal: NLS, Kalderon *et al.* (1984) *Cell* 39 (3 Pt 2):499-509) 및 표적 상의 ZFN 활성을 향상시키는 친수성 펩타이드(Hopp *et al.* (1988) *Nat Biotechnol* 6:1204-10)를 포함하며, 둘 다 N-말단의 암호 서열에 융합된다. 따라서, ZFN의 mRNA 및 아미노산 서열은 다음과 같다:

63014 mRNA (1725 nt)

5' gggagacaagcuuugaauuacaagcuugcuuguuuuuugcagaagcucagaaauaacgcuacaacuuugg
 cagaucgaauucgccauggacuacaagaccacugacggugauuaaagaucuacuacuacaggaaug
 acgaugacaagauggcccccaagaagaagaggacggcauccacggguacccggcguauggcugagag
 gcccuccagugcgaucugcgaucacuucagugaccaguccaaccugcgcccccacauccgcacccaca
 ccggcgagaagccuuuugccugugacauuugugggaggaaauuugccgcaacuucuuccugaccaugcauac
 caagauacacacggcagccaaagccuuccagugcgaucugcgaugcagaacuucaguuccacccggcaacc
 ugaccaaccacauccgcaccacccggcagaagccuuuugccugugacauuugugggaggaaauuugccac
 cuccggcucccugacccgccaucacaagauacacacgcacccgccccgaugccuuccagugc
 gaaucugcugcagaacuucagugaccaguccaaccugcgcccccacauccgcacccacccggcagaagccu
 uuugccugugacauuugugggaggaaauuugccgcccaguguugcugugaccaagauacaccugc
 gggauccauacgcagagccagaccacugaacccgcacccggagcuggaggagaagaaguccgagcugcggcac
 aacgugaaguacgugccacagaguacaucgagcugacuacgagcaggacagcaccaggaccgcauccu

[0239] ggagaugaaggugauggaguuuucuicaugaagguguacggcuacagggaaagcaccugggcggaagcagaaa
 ggcugacggcgccaucuauacagugggcagcccaucgauuacggcgugaucuguggacacaaaggccuacagc
 gggcgcuacaucugccuaucggcaggccgacgagaaggagagauacguggaggagaaccagacccgggaua
 agcaccucaaccccaacgagugggaggagaagguguacccuagcagcugacccgaguucaaguuccugugag
 cggccacucaaggcaacuacaaggcccagcugaccaggcugaaccacuacuacugccaauggccgugc
 ugagcuguggaggagcugcugauccggcgagaucaaaagccggacccugacacuaggaggugcggc
 gcaaguucaacaacggcgagaucaacuicagauuacuacugcagacuagaagcucugcuiucuugcugucc
 aauuucuauuaaggguuccuuuuguuccuaaguacuacuacuacuacuacuacuacuacuacuacuacu
 aucuggauucugccuaauaaaaacauuuuuuuucauugcugcugcugcugcugcugcugcugc
 ucuauuaaggguuccuuuuguuccuaaguacuacuacuacuacuacuacuacuacuacuacuacu
 ggauuucugccuaauaaaaacauuuuuuuucauugcugcggacauuucuuaauuaaaaaaaaaaaaaaaa
 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaacuag

[0240] (서열번호 28) .

[0241] 63014 아미노산 서열(인식 나선 영역을 밑줄 표시하고; 링커를 대문자 이탈릭체로 나타내며; 평거 1, 3 및 5 골격 잔기로의 돌연변이를 이중 밑줄 표시하고; 이량체화 도메인 돌연변이(ELD)를 볼드체 및 이탈릭체로 나타내며; 친수성 웨타이드를 소문자 텍스트로 나타내고; 핵 국소화 신호(NLS)를 소문자 이탈릭체로 나타냄):

MdykdhdgdykdhdidykdddkMApkkrkvGIHGVPAAMAERPFQCRICM**QNFSDQS**
NLRAHIRTHTGEKPFACDICGRKFARNFSLTMHTKIHTGSQKPFQCRICM**QNF**
SSTGNLTNHIRHTHTGEKPFACDICGRKFATSGSLTRHTKIHTHPRAPIPKPFQCR
ICMQNFSDQS**NLRAHIRTHTGEKPFACDICGRKF**AAQCCLFHHTKIHLRGSI**R**
ARPLNPHPELEEKKSELRHKLK YVPEYIELIEIARNSTQDRILEMKVMEFFMK
VYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEME
RYVEENQTRD**KHLNPNEWWKVYPSSVTEFKFLVSGHFKGNYKAQLTRLNH**

ITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFRS

[0242] (서열번호 29)

65722 mRNA (1680 뉴클레오파이드):

5'gggagacaagcuugaaacaagcuugcuuucuuuuugcagaagcucagaauaaacgcucaacuuuggca
gaucgaauucgccuagagaucuggcggcggagagggcagaggaagucuucuaaccugcggugacguggagga
gaaucggccuaggaccauaggacuacaaagaccauugcggugauuaagaaucaugacaucgauuaacaagg
augacgaugacaagaauuggccccaagaagaagaggaaggugcggcauuucauggguacccgcccuauggcuga
gaggcccuccagugucgaaucugcaugcagaaguugccgcaacgaccacccauaccaagauac
acacgggcgagaagccuuccagugugcgaaucugcaugcagaacuucagcagaaggcccaccugauccgcccac

[0243] auccgcacccacccggcggagaagccuuuugccugugacauuugugggaggaaauuugcccagaagggcacc
ugggcgagcauaccaagauacacacgggaucucagaagccuuccagugugcgaaucugcaugcagaacuucag
ucgcggccgcgaccuguccggccacauccgcacccacccggcggagaagccuuuugccugugacauuuguggg
agggaaauuugccggccgacaaccugcacucccauaccaagauacaccuugccgggaauccagcuggugaagag
cgagcuggaggagaagaaguccgagcugccgacaaggcugaaguacugcccccacgaguacaucgagcugauc
gagaucgccaggaacagcacccaggaccgcauccuggagaugaaggugugiuucaugaagguguacg
gcuacagggaaagcaccuggggcgaaagcagaaagccugacggcgccaucuauacagugggcagcccaucga
uuacggcgugaucuguggacacaaaggccuacagccggccuacauucugccuauccgcccaggcugaccagg
cugaaccgcaaaaccaacugcaaugccggccugcugagcugccggaggcugcugauccggccgagaugauca
aagccggcacccugacacuggaggugccggcaaguacaacaacggcgagaucaacuucugauaacucga
gcuuagaagcucgcuuiucuugcuguccaauuuucauaagguiuccuuuguiuccuaaguacauacuacugg
gggauuaauauggaaggccuugagcaucuggauucugccuaauaaaaacauuuuauuuucauugcugccggac
auucuuauuaacuag

[0244] (서열 번호 30).

[0245] 65722 아미노산 서열(인식 나선 영역을 밑줄 표시하고; 링커를 대문자 이탈릭체로 나타내며; 친수성 펩타이드는 소문자이고; 뉴클레아제 국소화 신호는 소문자 이탈릭체이며; 평거 1, 2 및 4 골격 잔기로의 돌연변이를 이중 밑줄로 나타내고; 이량체화 도메인 돌연변이(ELD)를 볼드체 및 이탈릭체로 나타내고; 그리고 FokI 인산염 접촉 돌연변이(들)를 물결 밑줄로 나타냄):

MdykdhdgdykdhdidykdddkMApkkrkvGIHGVPAAMAERPFQCRICMQKFARND
HRTTHTKIHTGEKPFQCRICMQNFSQKAHLIRHIRTHTGEKPFACDICGRKFAQ
KGTLGEHTKIHTGSQKPFQCRICMQNFSRGRDLSRHIRTHTGEKPFACDICGR
KFARRDNLHSHTKIHLRGSQLVKSELEEKKSELRHKLKYVPHEYIELIEARNST
QDRILEMKVMEFFMKVGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKA
YSGGYNLPIGQADEMQRVVKENQTRNKHINPNEWWKVYPSSVTEFKFLFVS
GHFSGNYKAQLTRLN**R**KTNCNGAVLSVEELLIGGEMIKAGTLTLEEVRKFN

[0246] NGEINF (서열 번호 31).

실시예 5: 적혈구 자손에서 글로빈 수준의 평가

[0247] (전반적인 mRNA 수준이 제핵 및 적혈구 성숙 과정에서 극적으로 감소되기 전에) 분화 제14일에 단리시킨 세포 mRNA에서의 α - , β - 및 γ - 글로빈 mRNA 수준을 상기 표 4에서 나타낸 2개의 세포 프레파라트에 대해 RT-qPCR에 의해 결정하였다. γ -글로빈 mRNA 값을 (1로 정의한 비형질감염 RT-PCR 표준에서의 비를 기준으로 한 임

의의 단위를 이용하여) 동일한 샘플로부터의 β -글로빈 mRNA에 대해(도 5A) 또는 α -글로빈 mRNA 값에 대해(도 5B) 정규화하여 나타낸다.

- [0249] 적혈구 분화의 종말점인 제21일에 단리시킨 단백질 샘플의 역상 HPLC를 사용하여 BCL11A 적혈구 인핸서의 ZFN 매개 변형이 단백질 수준에서 태아 혜모글로빈을 상승시키는지의 여부를 결정하였다. 감마 글로빈(A감마와 G감마 피크의 합) 대 알파 글로빈 비뿐만 아니라 감마 글로빈(A감마와 G감마 피크의 합)/베타-유사 글로빈 비(A감마, G감마, 베타 및 델타-글로빈 피크의 합)를 결정하였고, 도 6에 나타낸다.
- [0250] 이 실험에서, 약 15% 내지 20% 수준까지 태아 글로빈 단백질 백분율의 대략 3 내지 4배 상승이 BCL11A 인핸서의 63014/65722-매개 봉괴 시 HSPC의 적혈구 자손에서 관찰되었다.
- [0251] **실시예 6: NSG 마우스에서 편집 세포의 생착**
- [0252] 이어서, 편집 인간 CD34+ 세포를 NSG 마우스에 주사하여 생착을 평가하였다. 이식 후 8 및 12, 16 및 20주에 수집한 말초 혈액에서, 그리고 12주 및 20주에 수집한 골수에서 형광 활성화 세포 분류(fluorescence activated cell sorting: FACS)를 이용하여 인간 키메라 현상 정도(즉, 인간 CD45+ 세포의 백분율)를 측정하였다. 추가로, ZFN 변형 세포의 생착 수준을 시험하기 위해, BCL11A 인핸서 좌위에서 유전자 봉괴 수준을 ZFN 표적 좌위의 직접 고속 대량 서열분석에 의해 평가하였고, 유입 물질에서 측정한 표적 유전자 변형 수준과 비교하였다.
- [0253] 두 건강한 공여자(PB-MR-003 및 PB-MR-004로 칭함)로부터의 HSPC를 G-SCF 및 플레리사포르(Plerixafor)를 이용하여 동원시키고 나서, 문헌[Yannaki *et al.* ((2012) *Mol Ther* 20(1):230-8. doi: 10.1038/mt.2011)]에 기재된 바와 같이 정제하였다. 밀테니 바이오테크 클리니맥스 플러스(Miltenyi Biotech CliniMACS Plus) 기기를 이용하여 CD34+ 세포에 대해 농축시키기 전에 프레세니우스-카비 로보(Fresenius-Kabi Lovo) 장치를 이용하여 백혈구 성분채집술 생성물에 대해 혈소판 고갈을 수행하였다. 이어서, 정제된 세포를 형질감염을 위해 배양물에 퍼종시켰다.
- [0254] CD34+ 세포 정제 2일 후에, 모 ZFN 쌍을 암호화하는 $120\mu\text{g}/\text{ml}$ 의 단일 mRNA, 63014/65722 또는 최적화된 ZFN 쌍 $80\mu\text{g}/\text{ml}$ 63014 및 $20\mu\text{g}/\text{ml}$ 65722를 암호화하는 최적화된 양의 두 별개의 mRNA 중 하나의 존재 하에서 맥시사이트(Maxcyte) 기기를 이용하여 세포를 전기천공시켰다. 형질감염 전에, 세포의 분취액을 비형질감염 대조군으로서 두었다. 9천 5백만개의 세포를 PB-MR-003으로부터 형질감염시키고 나서, 1억 2천만개의 세포를 PB-MR-004로부터 형질감염시켰다.
- [0255] 전기천공법 후에, 30°C 에서 일시적으로 밤새 배양을 수행하였고, 이어서, 추가 24시간 동안 37°C 에서 배양시켰다. 전기천공법 후 2일에, DNA 분석을 위한 세포 분취액을 취하고 나서, 남아있는 세포를 채취하고, 액체 질소에 저장하였다.
- [0256] 조건화: 방사선 조사 1 내지 2일 전에 $10\text{mg}/\text{kg}/\text{일}$ 바이트릴(Baytril) 워터로 마우스를 처리하고, 후속적으로 이식 전 16 내지 24시간에 300 RAD로 방사선 조사하였다. 꼬리 정매 주사를 통해 이식을 수행하였다(이하 참조). 이어서, 마우스는 신선한 바이트릴 워터를 받았다. 바이트릴 워터를 1주 후에 대체하고, 이식 14일 후에 바이트릴 워터 첨가를 중단하였다.
- [0257] 이식: 이식일에, 사전가온시킨 X-Vivo 10/1%PSG + 3 사이토카인 쟝테일(재조합 인간 줄기 세포 인자(Recombinant Human Stem Cell Factor: SCF), 재조합 인간 트롬보포이에틴(Recombinant Human Thrombopoietin: TPO), 및 재조합 인간 F1t-3 리간드(F1t-3L))로 37°C 에서 주위 온도(멸균/여과)에서 신선한 PBS/0.1% BSA를 제조한다. 동결보존한 세포를 37°C 에서 해동시키고, 펠릿화하고 나서, 사전 가온시킨 X-Vivo 배지에서 재현탁시키고 나서, 다시 펠릿화하고, PBS/0.1% BSA에서 재현탁시키고 나서, 계수하였다. 다른 펠릿화 후에, 세포 펠릿을 $550\mu\text{l}/\text{PBS}/0.1\%\text{BSA}$ 의 마우스(세포 계수를 기준으로 2×10^6 개 세포/ ml) 중에서 재현탁시켰다. 이어서, 세포를 25 개이지 바늘을 이용하여 마우스 꼬리 정맥 내로 실온에서 주사하였다. 연구 그룹을 이하의 표 5에 나타낸다.

표 5

편집된 hCD34+ 세포의 생착을 위한 투약 그룹

종: 마우스				성별, 연령, 수: 60 마리 암컷 NSG 마우스			
그룹 번호	N / 그룹	시험 항목	삽입결실% (제 2 일)	해동 후 제 1 일에 생존도	용량 (세포/마우스)	N / 희생	
						제 12 주	제 20 주
1	10 (M1-10)	63014 및 65722 mRNA로 처리한 PB-MR-003 공여자 세포.	79%	83%	1백만	5	5
3	10 (M21-30)	PB-MR-003.비형질감염	0.1%	95%	1백만	5	5
4	10 (M31-40)	63014 및 65722 mRNA로 처리한 PB-MR-004 공여자 세포.	75%	77%	1백만	5	5
6	10 (M51-60)	PB-MR-004.비형질감염	0.1%	92%	1백만	5	5

[0258]

[0259] 동물을 일반적 건강 상태에 대해 매일 관찰하였고, 처음 2주 동안 매일 체중을 재고, 이후에 2주마다 체중을 쟐다. 이식 후 8, 12, 16 및 20주에 하약 정맥으로부터($100\mu\text{l}$) 또는 이식 후 12 및 20주에 희생시킨 동물에 대해 심장천자를 통해(1ml) 말초 혈액을 수집하였다. 각각의 그룹(5마리 마우스/그룹)에서 동물의 절반을 이식 후 12주에 안락사시키고 나서, 분석을 위해 골수 및 말단 혈액을 수집하였다. 각각의 그룹(5마리 마우스/그룹) 내 남아있는 동물을 이식 후 20주에 희생시켰다.

[0260]

혈액 수집, 세포 채취 및 가공: 악하 정맥 또는 심장 천자를 통해 EDTA 관 내로 말초 혈액을 수집하고 나서, $500 \times g$ 에서 5분 동안 원심분리시켜 혈장을 제거하였다. 인산염-완충 식염수(PBS) 소 혈청 일부민(BSA) 세척 및 원심분리 후에, 10X 용적의 용혈성 완충제를 펠릿에 첨가하고 나서, 혼합물을 37°C 에서 15분 동안 인큐베이션시키고, 원심분리 후에 다시 세척하였다. 1ml PBS BSA 중에서 펠릿화된 분획을 재구성하고 나서; 분취액을 제거하고, $1,000 \times g$ 에서 5분 동안 원심분리시키고, 얻어진 펠릿을 유전형질분석을 위해 보존하였다. FACS 분석을 위해 상청액 분획을 이용하였다.

[0261]

골수, 대퇴골, 경골 및 골반뼈를 송아지 태아 혈청(FCS)을 함유하는 이스코브 변형 둘째코 배지(IMDM)에서 수집하고 나서; 전체 골수를 PBS BSA 용액 내로 풀러싱하고 나서, $70\mu\text{m}$ 나일론 스트레이너를 이용하여 여과시켰다. PBS BSA를 이용하여 용적을 10ml 로 조절하고 나서, 세포 계수(셀로미터(Cellometer))를 위해 분취액을 사용하였다.

[0262]

MiSeq 심층 서열 분석을 이용하여 ZFN 활성을 분석하였다. 간략하게, 비형질감염 대조군 CD34+ HSPC 또는 인핸서 표적화 ZFN mRNA으로 형질감염된 CD34+ HSPC 중 하나를 주사한 마우스로부터의 게놈 DNA를 제8주 및 제12주에 얻은 샘플로부터 또는 주사 후 제12주에 골수로부터 단리시켰다. 관심 대상의 영역(BCL11A 좌위 내에서 ZFN 결합 부위를 함유)을 PCR 증폭시키고 나서, 일루미나 플랫폼(MiSeq) 상에서 쌍별 말단 심층 서열분석에 의해 변형 수준을 결정하였다.

[0263]

일루미나 MiSeq 서열분석 플랫폼에 적합한 라이브러리를 생성하기 위해, 어댑터, 바코드 및 유동세포 결합제(짧은 DNA 서열)를 순차적 PCR에서 융합 프라이머의 2개 세트를 이용하여 표적 특이적 앰플리콘에 부착하였다. 마우스 혈액 및 골수 샘플에서 인간 BCL11A 인핸서 변형의 MiSeq 평가를 위해, 이를 샘플에서의 낮은 표적 DNA 양에 기인하여 프로토콜을 조절하여야 했다.

[0264]

MiSeq 어댑터 PCR을 위해 다음의 프라이머를 사용하였다: PRJIYLFN-f2:

ACA CGA CGC TCT TCC GAT CTN NNN AGT CCT CTT CTA CCC CAC CCA (서열번호 32) 및

[0266]

PRJIYLFN-r4:

GAC GTG TGC TCT TCC GAT CTC TAC TCT TAG ACA TAA CAC ACC AGG G (서열번호 33).

[0268]

분석을 위해, 마우스 골수 샘플로부터의 DNA를 DNeasy에 의해 단리시키고 나서, 각각의 PCR 반응에서 대략 100ng 의 DNA를 사용하였다. 마우스 혈액 샘플로부터의 DNA를 Tissue XS에 의해 단리시키고 나서, 각각의 반응에서 $15\mu\text{l}$ 단리된 DNA 중 $10\mu\text{l}$ 를 사용하였다. DNA에 추가로, 각각의 MiSeq PCR 반응에 다음을 첨가하였다: $25\mu\text{l}$ 의

핫스타 태그 믹스(HotStar Taq mix)(퀴아젠), 상기 열거한 $0.5\mu\text{l}$ 의 각각의 BCL11A 인핸서 프라이머(100nM의 농도) 및 $50\mu\text{l}$ 의 총 용적으로 물. 전형적인 MiSeq PCR 조건은 다음과 같다: 15' 동안 95°C, 변성, 및 30" 동안 94°C, 30" 동안 62°C, 및 40" 동안 72°C, 다음에 72°C,에서 10'로 30주기. MiSeq PCR 후에, PCR 산물을 물을 이용하여 1:50 내지 1:200으로 희석시키거나, 또는 매우 낮은 시작 세포 수로 샘플에 대해 희석시키지 않고 남겨 두었다. 상기 기재한 바와 같이 희석시킨 $1\mu\text{l}$ 의 MiSeq PCR 산물, $25\mu\text{l}$ 핫스타 태그 믹스, $1\mu\text{l}$ 정방향 바코드 프라이머, $1\mu\text{l}$ 역방향 바코드 프라이머(둘 다 10nM의 농도) 및 $50\mu\text{l}$ 의 총 반응 용적으로 물을 이용하여 바코드 PCR을 수행하였다. 바코드 PCR 조건은 다음과 같다: 15' 동안 95°C, 변성, 및 30" 동안 94°C, 30" 동안 60°C 및 30" 동안 72°C, 다음에 72°C,에서 10'로 18주기. 바코드 PCR 산물을 풀링하고 나서, 일루미나 MiSeq 서열분석기 상에서 서열분석하였다. 결과를 상기 표 5에 나타낸다.

[0269] 키메라 현상을 위한 FACS 분석 및 세포 계통 결정. 인간 키메라 현상 정도를 평가하기 위해, (생착 후 8, 12, 16 및 20주에) 말초 혈액 내 세포 분획 및 (생착 후 12 및 20주에) 골수를 각각 hCD45-APC Cy7(바이오레전드(Biolegend)) 및 hCD45-BV510(BD 바이오사이언시즈(BD Biosciences)) 항체로 염색하고 나서, FACS 분석을 추가로 수행하고, 이하의 표 6에 기재하는 특정 항체를 이용하여 조혈 계통 분석을 수행하였다:

표 6

세포 마커에 대한 항체 공급원

CD3-FITC: BD 561807 (클론 UCHT1)	BD 561807
CD19-PE: BD 340364 (클론 SJ25C1)	BD 340364
CD45-BV510	BD 563204
Lin-APC (CD3/UCHT1, CD14/HCD14, CD16/3G8, CD19/HIB19, CD20/2H7, CD56/HCD56)	BIOLEGEND 348803
CD33-PE-CF594	BD 562492
GlyA-FITC	DAKO 0870
CD38-PerCP Cy5.5	BD 551400
CD14-PE	BD 555398
CD34-PE Cy7	BD 560710
CD71-APC Cy7/H7	BD 563671
CD15-BV650	BD 564232
CD8-PerCP Cy5.5	BD 341051
CD4-PE Cy7	BD 344612
CD56-APC	BD 318310
IgM-APC Cy7	BD 314520
CD20-BV650	BD 563780

[0270]

추가로, HSPC 집단을 정제하고 분류하기 위해, 본 발명자들은 자기 세포 분리(magnetic cell separation: MACS)를 이용하는 농축/고갈 전략을 사용하였다. 골수 세포를 CD19-바이오틴, CD3-바이오틴, B220-바이오틴, TER119-바이오틴 및 m-ckit-바이오틴(BD 바이오사이언시즈)를 이용하여 처음 염색하고 나서, 이어서, 항-바이오틴 비드(밀테니 바이오텍)와 함께 인큐베이션시켰다. MACS의 자기장에 둔 LS 칼럼(밀테니 바이오텍)을 이용하여 양성 분획 및 고갈 분획을 분리시켰다. 분리 후에, 양성 분획을 스트렙타비딘-APC, CD3-FITC, CD19-PE, CD45-BV510(BD 바이오사이언시즈)으로 염색하였고, 고갈된 분획은 CD34-FITC(BD 바이오사이언시즈), Gly-A-PE(DAKO), CD19-APC(BD), Lin-APC(바이오레전드), 스트렙타비딘-APC, CD45-BV510, CD33-PE-CF594(BD) 및 CD38-PECy-7(바이오레전드)을 이용하여 염색하였다.

[0272] 상기 기재한 바와 같은 표준 절차를 이용하여 NSG 마우스 내로 비형질감염 HSPC 및 63014/65722-형질감염 HSPC를 생착시켰다. FACS를 이용하여 hCD45 양성 세포 분획을 측정함으로써 생착 후 이들 마우스에서 인간 키메라 현상 정도를 평가하였다.

[0273] 도 7은 이식 후 제8주 및 제12주에 수집한 말초 혈액에서 인간 CD45+ 세포의 백분율을 나타내며, 도 8은 제12주에 채취한 골수에서의 백분율을 나타낸다. 나타낸 바와 같이, 본 연구에서의 생착 수준은 비형질감염 대조군 및 63014/65722 형질감염 HSPC의 생착 후 인간 키메라 현상과 비슷하였다. 그룹을 통해 분포된 60마리 중에서 3마리의 마우스만이 CD45+ 세포를 갖지 않았다.

[0274] 다양한 조혈 세포 계통의 재구성을 표준 절차를 이용하는 계통 특이적 세포 표면 마커를 이용하여 제12주에 얻은 골수 세포의 FACS 분석에 의해 시험하였다. 도 9에 나타낸 바와 같이, BCL11A 특이적 ZFN 암호화 mRNA 처리

한 CD34+ 세포 자손과 비형질감염 세포의 자손 간에 주사 후 12주에 골수 내 모든 분석 인간 조혈 계통의 비슷한 표현을 관찰하였다. 생착 후 12주에 희생시킨 마우스의 골수를 단리시키고, 표시된 계통 마커를 인식하는 항체를 이용하는 FACS에 의해 다양한 조혈 계통의 분포를 분석하였다. 적혈구 세포는 CD45 양성이 아니기 때문에, 전체 집단에서 양성 염색 세포의 백분율로서 제공된 도 9c에서 적혈구 마커 Cd71+(Ter119)를 발현시키는 세포를 제외하고, 모든 수를 표시된 계통 마커에 대해 양성인 세포 염색 대 인간 CD45 양성 세포의 백분율의 비로서 제공한다.

[0275] BCL11A 적혈구 인핸서에서 유전자 변형 수준(삽입과 결실[삽입결실]을 갖는 대립유전자의 %)을 상기 기재한 바와 같은 MiSeq 서열분석 플랫폼을 이용하는 ZFN 표적 영역의 심층 서열분석에 의해 평가하였다. 데이터를 제8주 및 제12주로부터의 혈액 샘플에 대해 도 10에 나타내고, 제12주로부터의 골수 샘플 및 63014/65722 처리 세포의 12주 골수 세포 샘플로부터 유래된 분류된 계통에 대해 도 11에 나타낸다. 비교를 위해, 형질도입 후 2일에 측정한 삽입결실 백분율(표 5에서 열거)을 도 10 및 도 11의 그래프 상에 나타낸다.

[0276] 추가로, 다양한 시점에 그리고 다양한 계통에서 설정한 63014/65722 -처리 HSPC 공여자 세트 둘 다에 대해 BCL11A 적혈구 인핸서에서 유전자 변형의 양호한 체류를 발견하였다. BCL11A 의존적(B 세포, 'CD19'; 원시 전구체, 'CD38H') 계통과 BCL11A 독립적(골수 'CD33') 계통 둘 다에서 비슷한 변형을 관찰하였다. 유입 유전자 변형 수준은 PB-MR-004 공여자 샘플에서보다 PB-MR-003 공여자 샘플에서 더 높았지만, PB-MR-004 유래 세포는 일관되게 더 높은 변형 수준, 즉, PB-MR-003으로부터 유래된 것보다 마우스에서 변형의 더 양호한 보유를 나타낸다.

[0277] 전반적으로, 마우스의 BCL11A 적혈구 인핸서에서 유전자 변형의 관찰된 체류는 다양한 유전자 표적을 표적화하는 다수의 ZFN을 이용하는 마우스 실험에서 관찰된 것과 일치되었다.

[0278] 더 나아가, 인간 적혈구 전구체는 마우스에서 분화할 수 없기 때문에, 이들 세포에서 발생된 BCL11A 표적화된 유전자 변형의 양을 결정하기 위해, 골수 세포를 마우스로부터 제거하였고, 시험관내에서 분화시켰다. 이들 실험에서, 골수 유래 인간 세포를 생착 후 12주에 희생 마우스로부터 제거하고 나서, 상기 기재한 바와 같이 시험관내에서 분화시켰다. 분화 제14일에 세포로부터 단리시킨 DNA의 고속대량 Miseq 서열분석에 의해 BCL11A 표적 유전자 변형을 측정하였다.

[0279] 변형 데이터(삽입결실)를 도 12에 제시하였는데, 이는 적혈구 분화의 제14일에 변형 수준을 나타낸다. 시험관내 분화의 제14일에 삽입결실 백분율은 한 마리의 마우스로부터 단리시킨 세포로부터 생성된 각각의 배양물에 대해 현저하게 달랐는데, 이는 이들 조건 하에서 얻은 확장의 올리고세포 특성을 반영한다. 데이터는 63014/65722 ZFN에 의해 매개된 BCL11A 인핸서 변형이 적혈구 분화 동안 현저하게 변경되지 않았다는 것을 나타낸다. 혈액 및 골수 샘플에서 관찰된 바와 같이, PB-MR-004 유래 세포의 적혈구 자손 샘플은 PB-MR-003 유래 세포의 적혈구 자손보다 더 높은 평균 수준의 변형을 나타내었다.

[0280] 다양한 글로빈 mRNA의 상대적 양을 시험관내 적혈구 분화의 제14일에 세포로부터 단리시킨 RNA의 RT-PCR 분석에 의해 결정하였고, 데이터를 도 13A에 제시하며, 여기서 상대적 γ -글로빈 대 β 글로빈 mRNA 및 γ -글로빈 대 α -글로빈 mRNA 비(도 13B)를 각각의 그룹으로부터의 5개 적혈구 배양물에 대해 평균내었다. 비형질감염과 63014/65722 처리 샘플 γ -글로빈 대 β 글로빈 또는 γ -글로빈 대 α 글로빈 mRNA 비는 둘 다 동일한 그룹으로부터의 개개 마우스의 적혈구 자손 간에 크게 다르다. 공여자 PB-MR-004 유래 배양물은 PB-MR-004 유래 샘플에 대해 관찰된 양호한 성숙과 비슷하게 공여자 PB-MR-003으로부터의 γ -글로빈 비보다 평균적으로 더 낮은 비를 나타낸다. 그러나, 이 가변성에도 불구하고, ZFN 처리 샘플 평균은 그들의 각각의 비형질감염 대상에 비해 γ 글로빈 mRNA 수준의 대략 1.5 내지 2배 증가를 나타낸다.

[0281] HPLC 분석에 의해 글로빈 단백질 수준을 평가하였다. 도 14는 분화의 제16일에 샘플의 글로빈 단백질 분석을 나타낸다. 감마 글로빈(A감마와 G감마 피크의 합) 대 알파 글로빈 비뿐만 아니라 감마 글로빈(A감마와 G감마 피크의 합)/베타-유사 글로빈 비(A감마, G감마, 베타 및 텔타-글로빈 피크의 합)를 결정하였고, 각각의 그룹에 대한 평균을 각각의 막대 위에 나타낸다. PB-MR-003 유래 샘플의 불량한 적혈구 분화와 비슷하게, 이 공여자로부터 유래된 비형질감염 세포의 감마-글로빈 수준은 매우 높았고(대략 30%), 따라서 ZFN 처리는 감마-글로빈 수준의 단지 1.2배 증가를 초래하였다. PB-MR-004는 더 전형적인 비형질감염 수준(대략 9%)을 나타내었고, 마우스를 통한 12주 계대 후에 감마-글로빈 단백질의 대략 2배 증가를 나타내었다.

[0282] 자연적으로 8.6% 초과의 γ 글로빈을 갖는 환자는 8.6% 미만의 γ 글로빈 수준을 갖는 환자에 비해 이점이 있는 것으로 생각된다(Platt *et al.* (1994) *N Engl J Med*, 330:1639-44). 사실, 편집 세포의 생착을 통해 비경상 세포 RBC의 키메라의 10 내지 20% 백분율을 달성하는 것은 임상적 개선을 야기할 수 있다(Chang *et al.* (2017)

Mol Ther Methods Clin Dev 4:137-148. doi 10.1016/j.omtm.2016.12.009). 따라서, 시험관내 적혈구 분화 과정을 거침에도 불구하고, 키메라 세포의 백분율 및 검출될 γ글로빈 단백질의 수준은 치료 효율을 나타낸다.

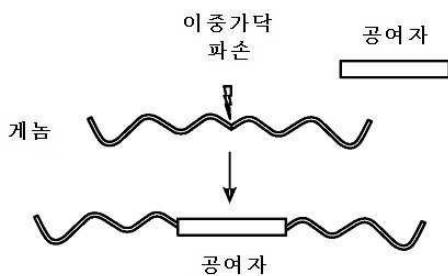
[0283] 본 명세서에서 언급되는 모든 특허, 특허 출원 및 간행물은 본 명세서에 그들의 전문이 참고로 포함된다.

[0284] 개시내용은 이해의 명확을 기하기 위하여 예시 및 실시예의 방법으로 일부 상세하게 제공하였지만, 다양한 변화 및 변형이 본 개시내용의 정신 또는 범주로부터 벗어나는 일 없이 실행될 수 있다는 것은 당업자에게 분명할 것이다. 따라서, 앞서 언급한 설명 및 실시예는 제한으로서 해석되어서는 안 된다.

도면

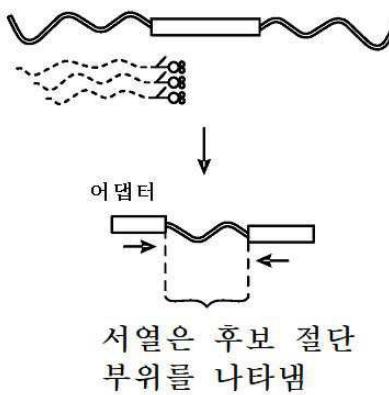
도면1

개놈을 절단하여 공여자에 통합함



A

공여자에 인접한 서열 개놈



B

도면2a

염색체	염기번호	구성요소 계수	염색체	염기번호	구성요소 계수	염색체	염기번호	구성요소 계수	
BCL11A→	2	60495266	1064	22	20508058	15	17	33940450	9
	8	119856442	939	1	167637598	14	18	20578334	9
	2	62164814	156	1	211655462	14	19	850496	9
	10	132654828	105	9	129611194	14	20	29445862	8
	14	67422056	102	16	46399310	14	1	51299472	8
	4	7829724	93	19	38665086	14		201507258	8
	2	23702832	88	14	50504	14	2	16715068	8
	14	73740920	85	14	50604	14	2	20996616	8
	1	29182166	82	2	98873178	13	2	36785314	8
	10	69571176	75	2	107763998	13	2	226916064	8
	16	15141540	68	2	127820122	13	2	237467876	8
	8	94988042	66	6	1596312	13	6	49519624	8
	16	2122340	66	6	35468034	13	6	160955686	8
	19	45406680	64	6	144064118	13	10	26734304	8
	16	67184172	59	8	107386748	13	10	79114636	8
	16	46394744	58	10	98910092	13	12	50505884	8
	13	75549350	57	12	104048816	13	14	73746866	8
	16	34593646	57	14	89466992	13	17	43805774	8
	16	46390726	54	1	21635648	12	17	82718744	8
	1	54506814	53	2	135453242	12	18	20575274	8
	19	55115766	53	6	7899998	12	20	5788510	8
	6	53805956	52	8	13105966	12	21	25629778	8
	17	29624234	52	9	128604032	12	22	19983370	8
	1	26938300	48	11	2992880	12	X	47067946	8
	5	145864942	40	13	91576712	12	1	19916860	7
	20	37707466	39	14	18295748	12	1	143187676	7
	21	33308800	36	16	34587146	12	1	148402426	7
	1	204975770	35	21	25607680	12	1	223224616	7
	5	132776212	35	22	15226546	12	2	119198990	7
	11	33894468	34	1	70765014	11	3	45419400	7
	10	101133378	33	1	120942294	11	4	39364694	7
	17	28854518	33	1	200702950	11	4	99547294	7
X		30280458	31	3	19942238	11	5	38056212	7
	2	84459580	30	5	2057482	11	6	6715482	7
	7	131503656	30	6	3196958	11	6	15090744	7
	2	112583534	27	9	98755462	11	6	89830096	7
	3	49724756	27	11	30406776	11	6	89888012	7
	12	55928112	27	1	177943758	10	7	113383212	7

도면2b

염색체	염기번호	구성요소 계수	염색체	염기번호	구성요소 계수	염색체	염기번호	구성요소 계수
8	66165666	26	2	218419648	10	8	37700430	7
11	119316888	26	6	136793200	10	8	88289848	7
X	134914730	26	8	25069948	10	9	101660500	7
8	95226212	25	10	46426956	10	9	112857552	7
12	124995298	25	10	47957704	10	10	42097318	7
16	10992070	25	10	52321784	10	10	91591618	7
3	47439806	24	10	52321864	10	10	120431936	7
11	70518294	23	11	102685874	10	11	5225468	7
16	10839466	22	12	54411358	10	12	155086	7
X	66004390	22	12	76936636	10	12	51951306	7
1	92672120	21	15	19782342	10	15	64356576	7
9	91430052	21	16	15220070	10	15	65185434	7
16	30692064	21	16	15220090	10	16	34573028	7
1	155276858	20	16	18131696	10	16	34573058	7
2	173018750	20	16	18131716	10	16	46401528	7
16	88359110	20	18	5447734	10	16	81738062	7
2	149709484	19	19	43779827	10	17	836804	7
8	124924782	19	19	51327816	10	17	75814568	7
13	112237576	19	21	33199996	10	19	17579762	7
20	44683654	19	1	26569750	9	19	38478442	7
	228985770	18	1	57734650	9	19	40578628	7
7	13991304	18	1	115988958	9	19	41165952	7
7	101110944	18	1	149636460	9	20	58585134	7
11	108224088	18	3	51942092	9	22	36368680	7
12	89974464	18	3	149936952	9	22	36423722	7
1	208115440	17	4	84425848	9	1	2914438	6
5	150115792	17	4	138895946	9	1	91089046	6
12	7096136	16	6	3737388	9	1	115328550	6
15	28018428	16	6	164378486	9	1	125169466	6
19	1279936	16	9	93003938	9	1	125169552	6
20	48636010	16	9	104239026	9	1	155610614	6
1	165379386	15	10	67764450	9	1	155746194	6
3	71103574	15	10	69789200	9	1	201664794	6
5	103961570	15	13	98822320	9	2	131443890	6
7	128077200	15	15	31392604	9	3	4323698	6
8	29495584	15	15	34422898	9	3	58414862	6
16	463399748	15	15	34569124	9	3	66129874	6
16	57534378	15	15	78243618	9	3	87800708	6

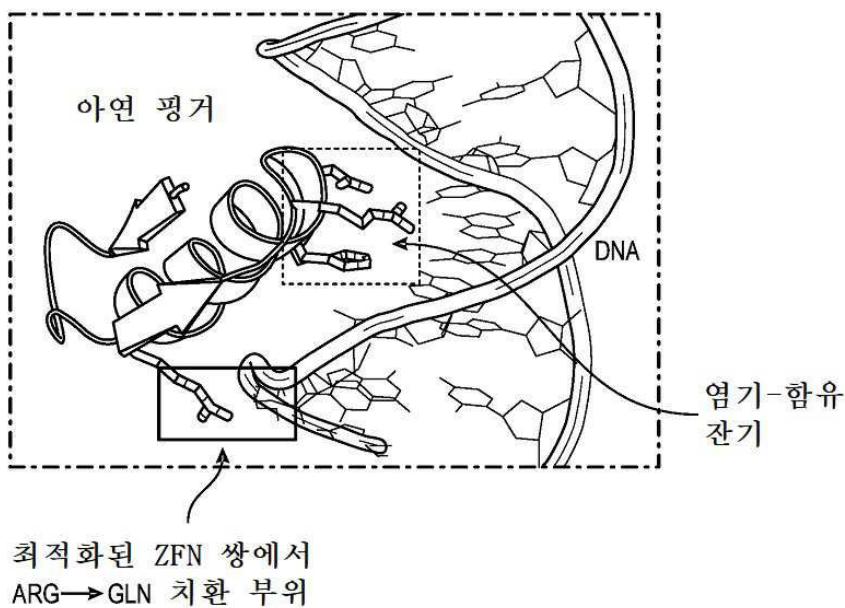
도면2c

염색체	염기번호	구성요소 계수	염색체	염기번호	구성요소 계수	염색체	염기번호	구성요소 계수
3	141228736	6	13	79775976	5	9	92277906	4
4	151031930	6	15	40114388	5	9	97934932	4
5	178121148	6	15	68307028	5	9	121566378	4
6	28931180	6	15	82539916	5	9	124392210	4
6	108893514	6	15	86572492	5	9	126591182	4
6	136088656	6	16	21645824	5	9	138212048	4
6	163725808	6	16	53042972	5	10	25357460	4
7	162658	6	16	58873960	5	10	41847448	4
7	35016604	6	16	67841642	5	10	41847504	4
8	30408876	6	16	78980108	5	10	42103514	4
8	91876290	6	17	17025784	5	10	64508254	4
8	119945566	6	17	26884866	5	10	76825618	4
9	4090722	6	17	72902448	5	10	109893420	4
9	104968230	6	18	11456592	5	11	35944902	4
10	72293352	6	18	33538826	5	11	72431102	4
10	79866034	6	19	7159860	5	12	6232920	4
11	118997014	6	19	11922082	5	12	7242432	4
11	121069264	6	19	18462114	5	12	49404046	4
12	3073480	6	19	45686048	5	12	75511410	4
12	26186106	6	20	29881056	5	12	88031364	4
13	43878548	6	20	43558622	5	12	95073610	4
15	41284346	6	20	47575112	5	13	113153372	4
15	58555840	6	22	19545504	5	13	18964478	4
15	60478340	6	22	23284226	5	13	34779902	4
16	3106742	6	14	33684	5	13	113600728	4
16	4518046	6	1	6359676	4	14	92616966	4
16	29292606	6	1	18556502	4	15	33229354	4
16	77988246	6	1	24862574	4	15	64142544	4
16	88266652	6	1	25442520	4	15	67498752	4
16	88732066	6	1	30801948	4	15	75356538	4
17	36938710	6	1	51631166	4	15	76355710	4
19	267174	6	1	72774624	4	15	684294	4
22	39690408	6	1	109213944	4	15	14285910	4
X	40259106	6	1	153609522	4	15	85317250	4
X	40835932	6	1	175878760	4	17	9710280	4
1	1724636	5	1	201154009	4	17	21372082	4
1	6161778	5	1	224687886	4	17	21906776	4
1	17419988	5	2	137774808	4	17	26603948	4

도면2d

염색체	염기번호	구성요소 계수	염색체	염기번호	구성요소 계수	염색체	염기번호	구성요소 계수
1	36401004	5	2	166910124	4	17	26603992	4
1	143208542	5	2	234688446	4	17	40557500	4
1	150528982	5	3	49802584	4	17	43796376	4
1	154750274	5	3	195856500	4	17	49883740	4
2	86553402	5	3	198148214	4	18	13272880	4
2	105039552	5	4	1688620	4	18	30136846	4
2	117954094	5	4	3480800	4	18	30845084	4
2	130230372	5	4	99869180	4	18	74019902	4
2	176188548	5	4	138090616	4	19	40266218	4
3	2041446	5	4	158876704	4	19	42741976	4
3	16078092	5	5	10727038	4	19	42767202	4
3	75669968	5	5	36318878	4	19	42881190	4
3	178998880	5	5	36322678	4	19	49683046	4
3	184689238	5	5	65562414	4	20	4673782	4
4	49119414	5	5	100709408	4	20	24683868	4
4	73573948	5	5	142193948	4	20	44911634	4
4	78700116	5	5	149564992	4	20	48621130	4
5	7057496	6	6	164896712	4	20	49279476	4
5	11767140	6	6	176740968	4	20	49694942	4
5	56314802	6	6	2643624	4	20	56914278	4
5	132314286	6	6	27782516	4	20	62819166	4
6	58116958	6	6	45156450	4	21	10442640	4
6	60870558	6	6	146750040	4	22	22533154	4
7	128792974	6	6	169090710	4	22	22648222	4
7	133081754	7	7	1214740	4	22	22956152	4
9	92835866	7	7	37357350	4	22	23111094	4
9	131373156	7	7	37357420	4	22	23162790	4
10	50701042	7	7	62355868	4	22	24629320	4
10	87382350	7	7	71679460	4	22	32136074	4
11	49652236	7	7	106778616	4	22	45216526	4
11	61612468	7	7	13077880	4	22	6134666	4
11	62795536	8	8	5753270	4	22	39688532	4
11	67277754	8	8	5753290	4	22	52934730	4
11	82032970	8	8	8244554	4	22	71283192	4
11	88966978	8	8	8371950	4	22	90323678	4
12	6326346	8	8	101955394	4	22	106863278	4
12	104138422	8	8	119000446	4	22	3900438	4
12	121500124	5	9	38068006	4			

도면3a



도면3b

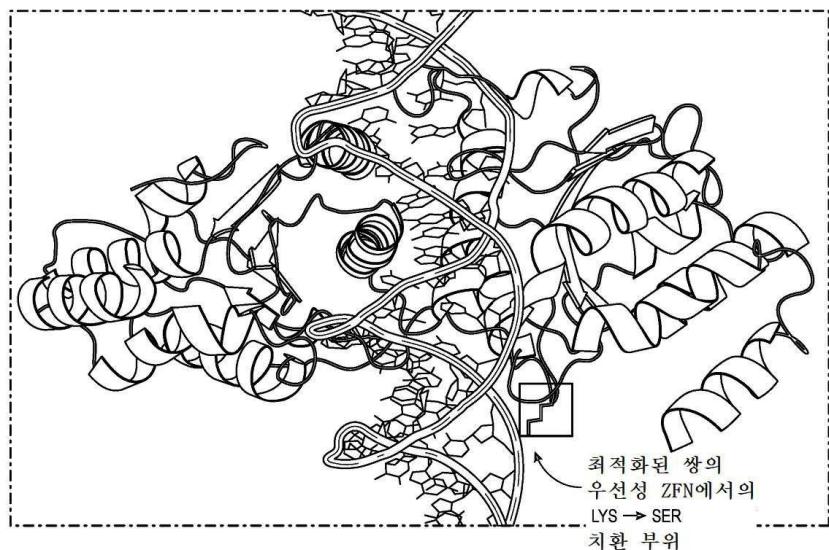
최적화된 ZFN 쌍에서의
ARG GLN 치환 부위

염기-함유
잔기

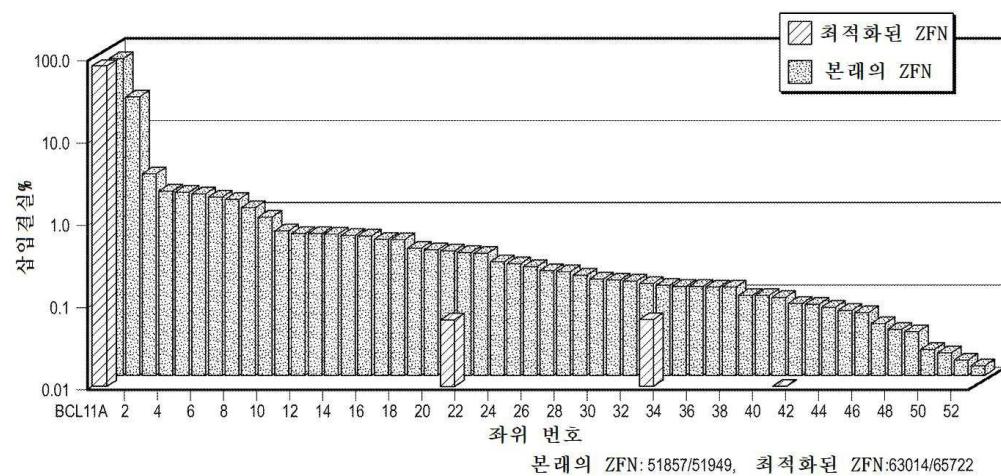
좌 ZFP	F1	VPAAMAERP	FQCRICM R NF	S	DQSNL R AHIRTH	
	F2	TGEKP	FACDICGRKF	A	RNFS	LTMHTKIH
	F3	TGSQKP	FQCRICM R NF	S	STGN	LTNHIRTH
	F4	TGEKP	FACDICGRKF	A	TSGS	LTRHTKIH
	F5	THPRAPIPKP	FQCRICM R NF	S	DQSN	LRAHIRTH
	F6	TGEKP	FACDICG R KF	A	AQCC	L F HHTKIH LRGS
<hr/>						
우 ZFP	F1	VPAAMAERP	FQCRICM R NF	A	RNDHRTH	IHTKIH
	F2	TGEKP	FQCRICM R NF	S	QKAH	LIRHIRTH
	F3	TGEKP	FACDICGRKF	A	QKGT	LG E HTKIH
	F4	TGSQKP	FQCRICM R NF	S	RGRD	LSRHIRTH
	F5	TGEKP	FACDICGRKF	A	RRDN	LHS I HTKIH LRGS

→ → → →

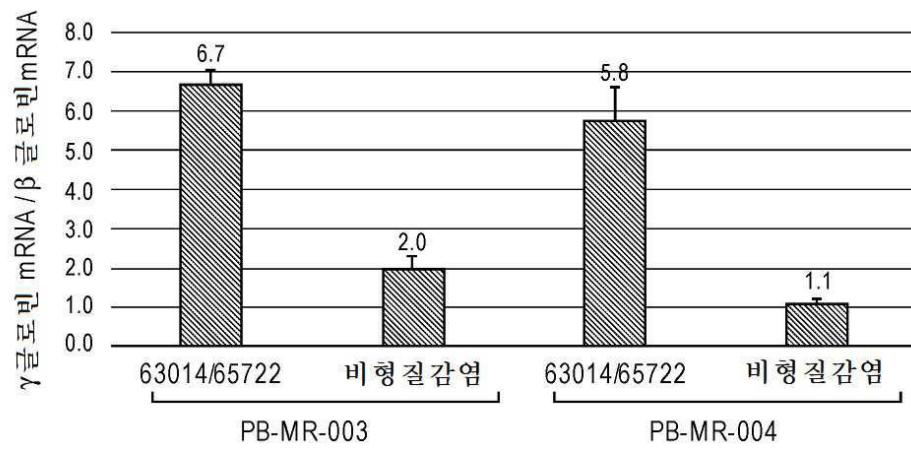
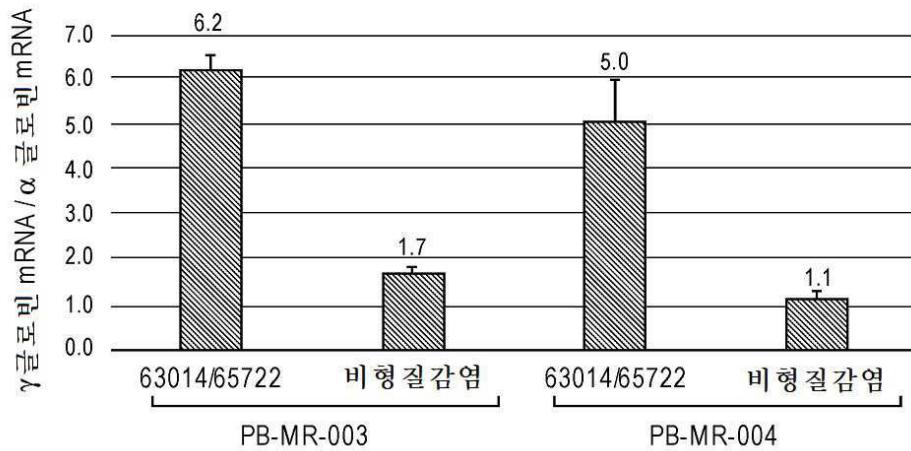
도면3c



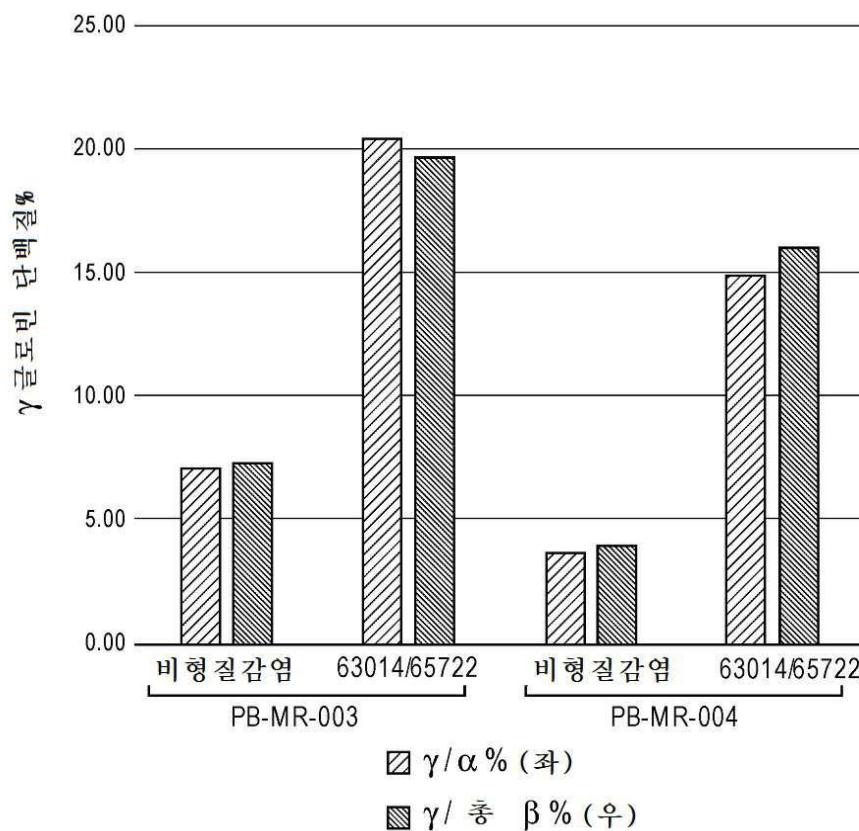
도면4



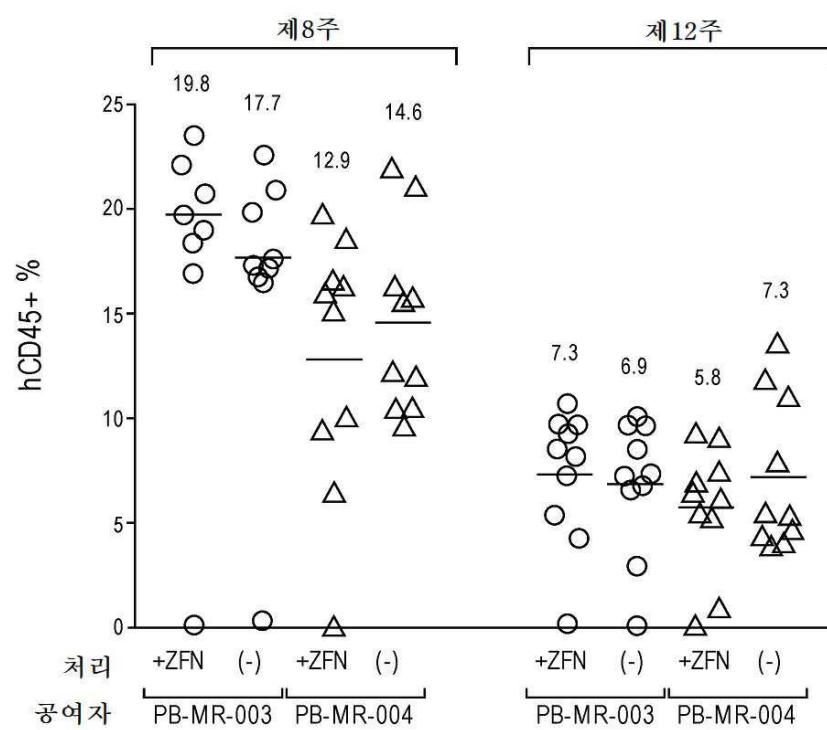
도면5

**A****B**

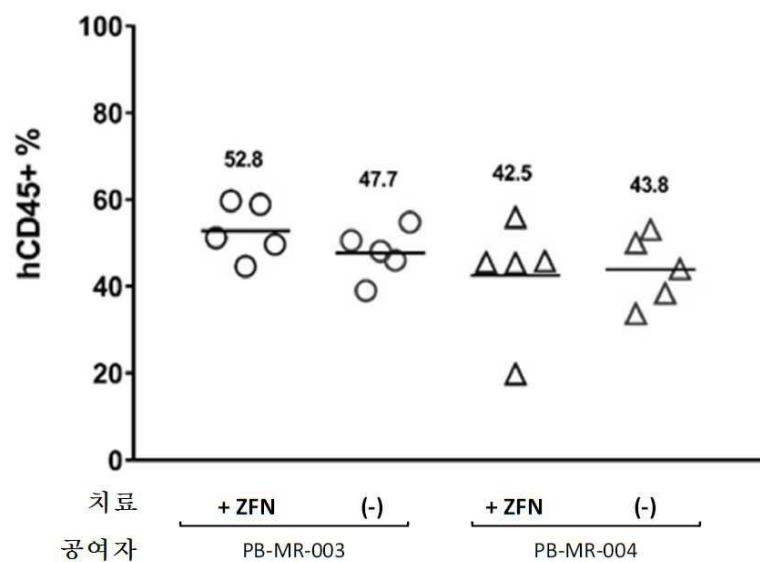
도면6



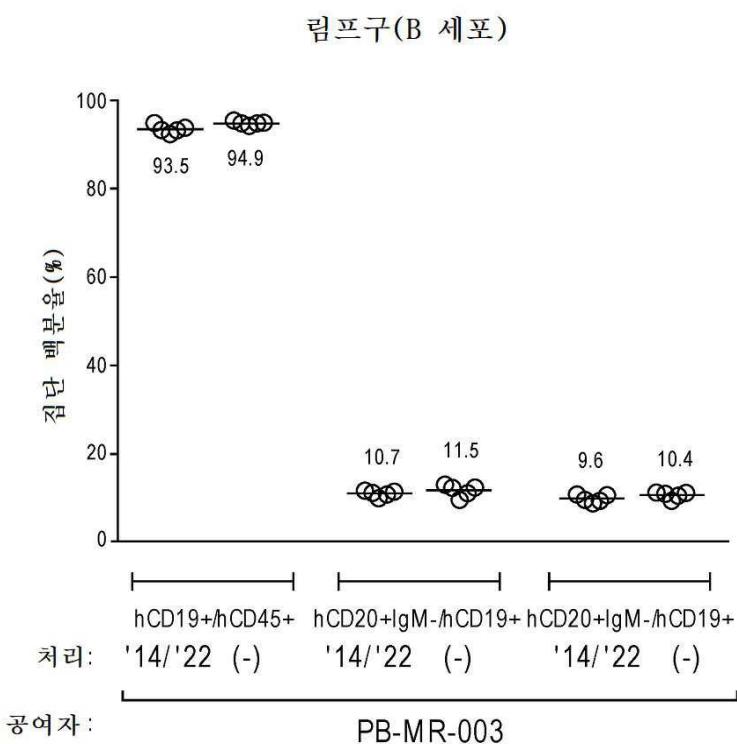
도면7



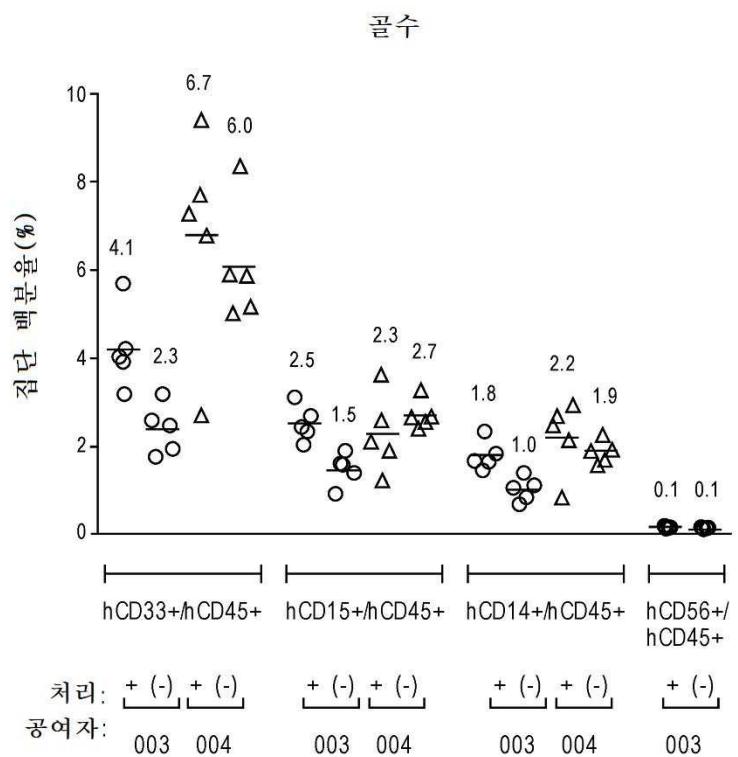
도면8



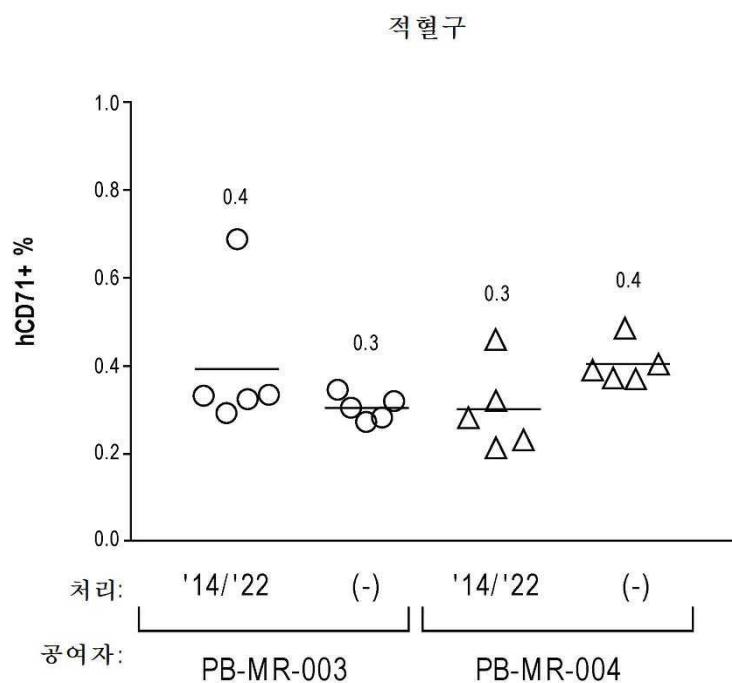
도면9a



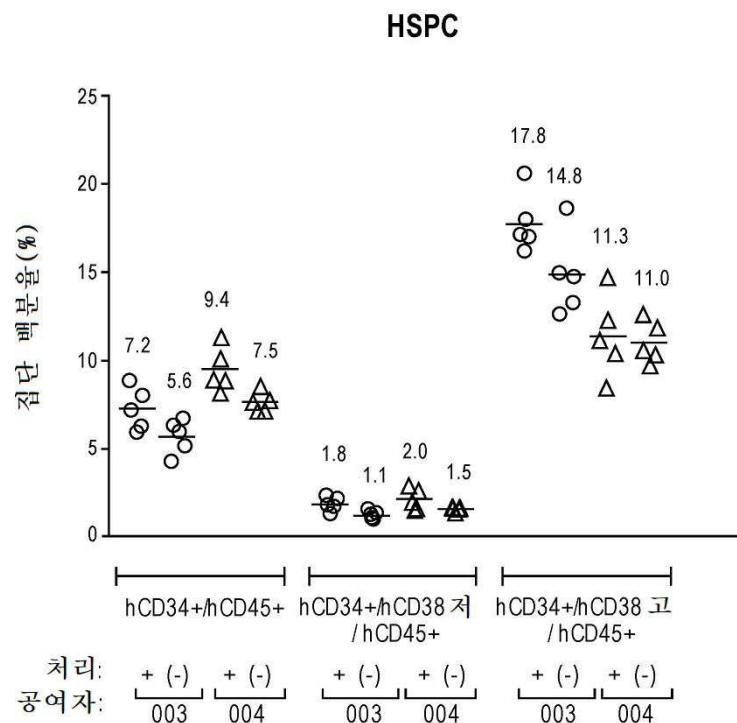
도면9b



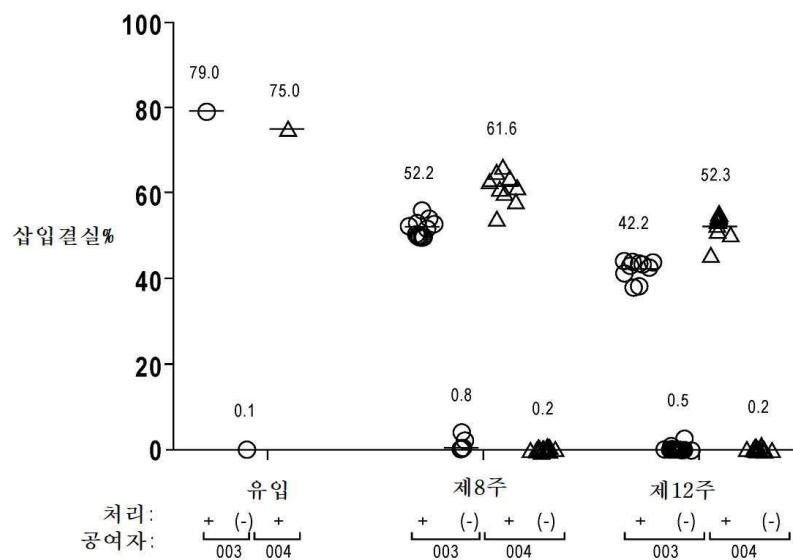
도면9c



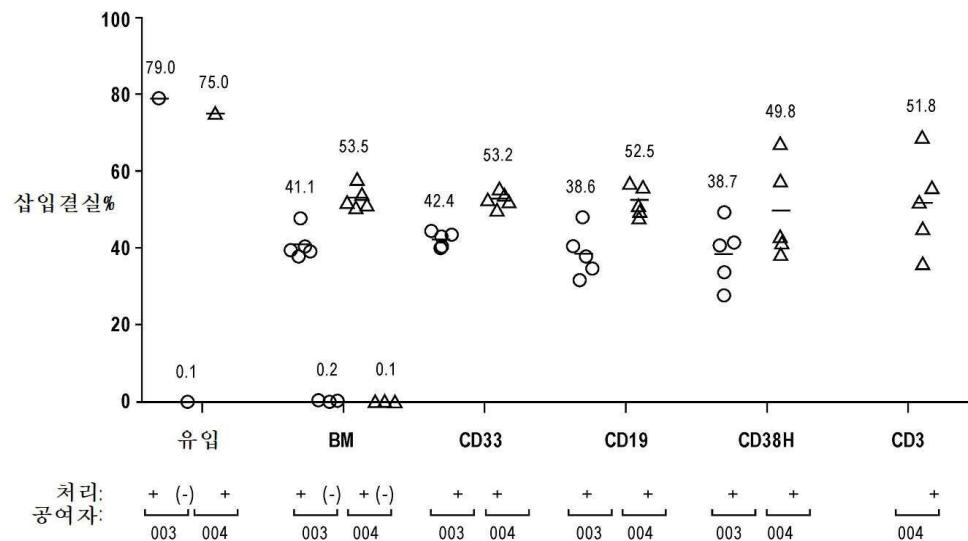
도면9d



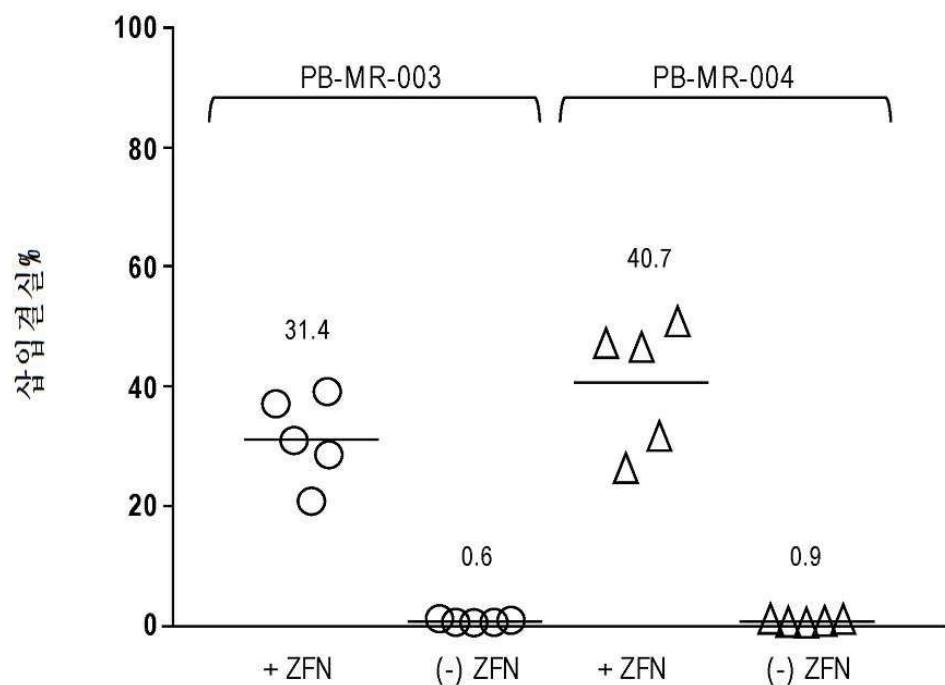
도면10



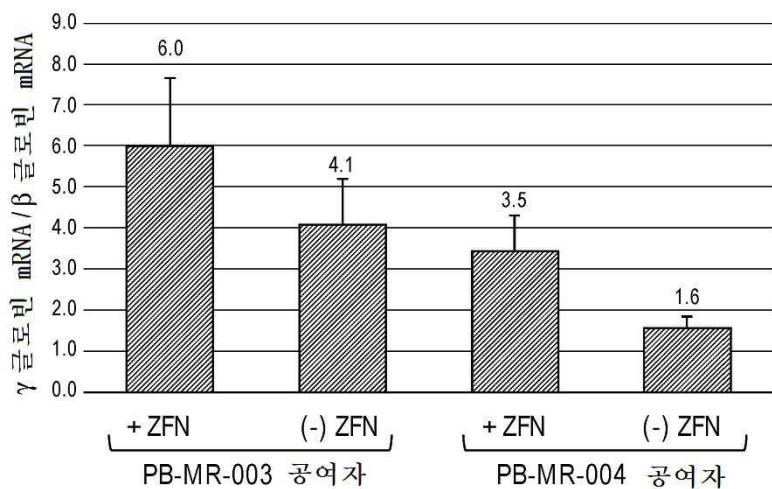
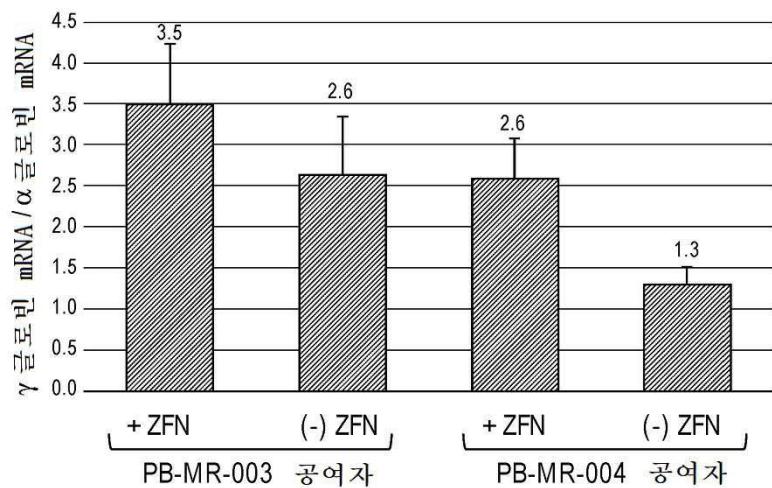
도면11



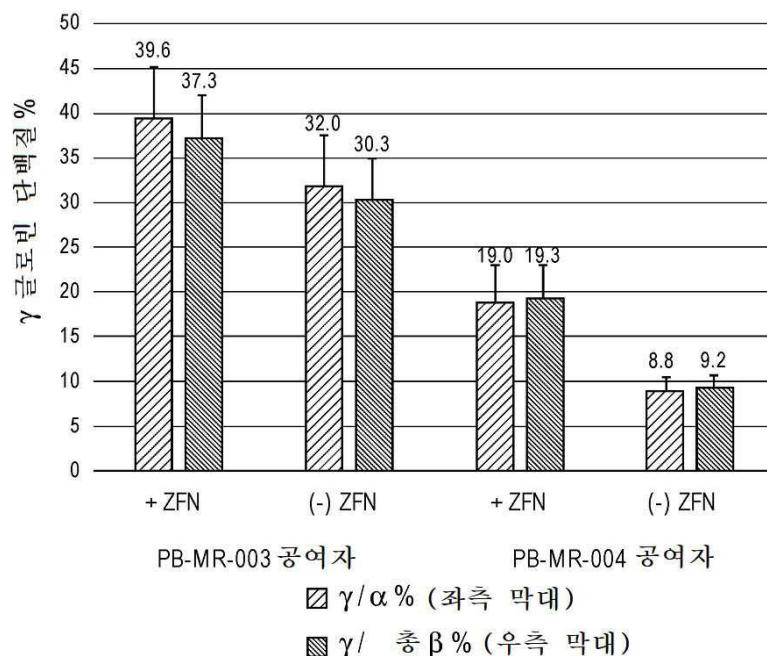
도면12



도면13

**A****B**

도면14



서 열 목 록

SEQUENCE LISTING

<110> SANGAMO THERAPEUTICS, INC.

<120> REGULATION OF GENE EXPRESSION USING ENGINEERED NUCLEASES

<130> WO/2018/039440

<140> PCT/US2017/048397

<141> 2017-08-24

<150> US 62/545,778

<151> 2017-08-15

<150> US 62/443,981

<151> 2017-01-09

<150> US 62/378,978

<151> 2016-08-24

<160> 34

<170> Patent In version 3.5

<210> 1

<211> 28

<212> DNA

<213> Artificial Sequence

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

<400> 1

aaagcaactg tttagtttgc ctagacta 28

<210> 2

<211> 28

<212> DNA

<213> Artificial Sequence

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

<400> 2

cacaggctcc aggaagggtt tggcctct 28

<210> 3

<211> 407

<212> PRT

<213> Artificial Sequence

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

<220><221> NON_CONS

<222> (214)..(215)

<223> Residues at these positions are separated by any linker
sequence known in the art

<400> 3

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1 5 10 15

Tyr Lys Asp Asp Asp Lys Met Ala Pro Lys Lys Lys Arg Lys Val

20 25 30

Gly Ile His Gly Val Pro Ala Ala Met Ala Glu Arg Pro Phe Gln Cys

35 40 45

Arg Ile Cys Met Gln Asn Phe Ser Asp Gln Ser Asn Leu Arg Ala His

50 55 60

Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly

65	70	75	80
Arg Lys Phe Ala Arg Asn Phe Ser Leu Thr Met His Thr Lys Ile His			
85	90	95	
Thr Gly Ser Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Gln Asn Phe			
100	105	110	
Ser Ser Thr Gly Asn Leu Thr Asn His Ile Arg Thr His Thr Gly Glu			
115	120	125	
Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Thr Ser Gly			
130	135	140	
Ser Leu Thr Arg His Thr Lys Ile His Thr His Pro Arg Ala Pro Ile			
145	150	155	160
Pro Lys Pro Phe Gln Cys Arg Ile Cys Met Gln Asn Phe Ser Asp Gln			
165	170	175	
Ser Asn Leu Arg Ala His Ile Arg Thr His Thr Gly Glu Lys Pro Phe			
180	185	190	
Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Ala Gln Cys Cys Leu Phe			
195	200	205	
His His Thr Lys Ile His Glu Leu Glu Glu Lys Ser Glu Leu Arg			
210	215	220	
His Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile			
225	230	235	240
Ala Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu			
245	250	255	
Phe Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser			
260	265	270	
Arg Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr			
275	280	285	
Gly Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro			
290	295	300	
Ile Gly Gln Ala Asp Glu Met Glu Arg Tyr Val Glu Glu Asn Gln Thr			
305	310	315	320

Arg Asp Lys His Leu Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser
 325 330 335

Ser Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly
 340 345 350

Asn Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn
 355 360 365

Gly Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile
 370 375 380

Lys Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn
 385 390 395 400

Gly Glu Ile Asn Phe Arg Ser
 405

<210> 4

<211> 372

<212> PRT

<213> Artificial Sequence

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

<220><221> NON_CONS

<222> (181)..(182)

<223> Residues at these positions are separated by any linker
 sequence known in the art

<400> 4

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1 5 10 15

Tyr Lys Asp Asp Asp Lys Met Ala Pro Lys Lys Arg Lys Val
 20 25 30

Gly Ile His Gly Val Pro Ala Ala Met Ala Glu Arg Pro Phe Gln Cys
 35 40 45

Arg Ile Cys Met Gln Lys Phe Ala Arg Asn Asp His Arg Thr Thr His
 50 55 60

Thr Lys Ile His Thr Gly Glu Lys Pro Phe Gln Cys Arg Ile Cys Met

65	70	75	80
Gln Asn Phe Ser Gln Lys Ala His Leu Ile Arg His Ile Arg Thr His			
85	90	95	
Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala			
100	105	110	
Gln Lys Gly Thr Leu Gly Glu His Thr Lys Ile His Thr Gly Ser Gln			
115	120	125	
Lys Pro Phe Gln Cys Arg Ile Cys Met Gln Asn Phe Ser Arg Gly Arg			
130	135	140	
Asp Leu Ser Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala			
145	150	155	160
Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Arg Asp Asn Leu His Ser			
165	170	175	
His Thr Lys Ile His Glu Leu Glu Lys Lys Ser Glu Leu Arg His			
180	185	190	
Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala			
195	200	205	
Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe			
210	215	220	
Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser Arg			
225	230	235	240
Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly			
245	250	255	
Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile			
260	265	270	
Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Lys Glu Asn Gln Thr Arg			
275	280	285	
Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser			
290	295	300	
Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Ser Gly Asn			
305	310	315	320

Tyr Lys Ala Gln Leu Thr Arg Leu Asn Arg Lys Thr Asn Cys Asn Gly

325 330 335

Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys

340 345 350

Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly

355 360 365

Glu Ile Asn Phe

370

<210> 5

<211> 15

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 5

Leu Arg Gly Ser Ile Ser Arg Ala Arg Pro Leu Asn Pro His Pro

1 5 10 15

<210> 6

<211> 9

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 6

Leu Arg Gly Ser Gln Leu Val Lys Ser

1 5

<210> 7

<211> 32

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 7

Val Pro Ala Ala Met Ala Glu Arg Pro Phe Gln Cys Arg Ile Cys Met

1 5 10 15

Arg Asn Phe Ser Asp Gln Ser Asn Leu Arg Ala His Ile Arg Thr His

20 25 30

<210> 8

<211> 28

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 8

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala

1 5 10 15

Arg Asn Phe Ser Leu Thr Met His Thr Lys Ile His

20 25

<210> 9

<211> 29

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 9

Thr Gly Ser Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe

1 5 10 15

Ser Ser Thr Gly Asn Leu Thr Asn His Ile Arg Thr His

20 25

<210> 10

<211> 28

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 10

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala

1 5 10 15

Thr Ser Gly Ser Leu Thr Arg His Thr Lys Ile His

20 25

<210> 11

<211> 33

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 11

Thr His Pro Arg Ala Pro Ile Pro Lys Pro Phe Gln Cys Arg Ile Cys

1 5 10 15

Met Arg Asn Phe Ser Asp Gln Ser Asn Leu Arg Ala His Ile Arg Thr

20 25 30

His

<210> 12

<211> 32

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 12

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala

1 5 10 15

Ala Gln Cys Cys Leu Phe His His Thr Lys Ile His Leu Arg Gly Ser

20 25 30

<210> 13

<211> 32

<212> PRT

<213> Artificial Sequence

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

<400> 13

Val Pro Ala Ala Met Ala Glu Arg Pro Phe Gln Cys Arg Ile Cys Met
1 5 10 15
Arg Lys Phe Ala Arg Asn Asp His Arg Thr Thr His Thr Lys Ile His
20 25 30

<210> 14

<211> 28

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 14

Thr Gly Glu Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe Ser
1 5 10 15
Gln Lys Ala His Leu Ile Arg His Ile Arg Thr His
20 25

<210> 15

<211> 28

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 15

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
1 5 10 15

Gln Lys Gly Thr Leu Gly Glu His Thr Lys Ile His
20 25

<210> 16

<211> 29

<212> PRT

<213> Artificial Sequence

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

<400> 16

Thr Gly Ser Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Arg Asn Phe
 1 5 10 15
 Ser Arg Gly Arg Asp Leu Ser Arg His Ile Arg Thr His
 20 25

<210> 17

<211> 32

<212> PRT

<213> Artificial Sequence

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

<400> 17

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala
 1 5 10 15
 Arg Arg Asp Asn Leu His Ser His Thr Lys Ile His Leu Arg Gly Ser
 20 25 30

<210> 18

<211> 196

<212> PRT

<213> Unknown

<220><223> Description of Unknown:

Wild type FokI cleavage half domain

<400> 18

Gln Leu Val Lys Ser Glu Leu Glu Lys Lys Ser Glu Leu Arg His
 1 5 10 15
 Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala
 20 25 30
 Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe
 35 40 45
 Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Gly Ser Arg

50	55	60
Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly		
65	70	75
Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile		
85	90	95
Gly Gln Ala Asp Glu Met Gln Arg Tyr Val Glu Glu Asn Gln Thr Arg		
100	105	110
Asn Lys His Ile Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser		
115	120	125
Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn		
130	135	140
Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn Gly		
145	150	155
Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys		
165	170	175
Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly		
180	185	190
Glu Ile Asn Phe		
195		
<210> 19		
<211> 7		
<212> PRT		
<213> Artificial Sequence		
<220><223> Description of Artificial Sequence: Synthetic peptide		
<400> 19		
Asp Gln Ser Asn Leu Arg Ala		
1	5	
<210> 20		
<211> 7		
<212> PRT		
<213> Artificial Sequence		

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

<400> 20

Arg Asn Phe Ser Leu Thr Met

1 5

<210> 21

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 21

Ser Thr Gly Asn Leu Thr Asn

1 5

<210> 22

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 22

Thr Ser Gly Ser Leu Thr Arg

1 5

<210> 23

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 23

Ala Gln Cys Cys Leu Phe His

1 5

<210> 24

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 24

Arg Asn Asp His Arg Thr Thr

1 5

<210> 25

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 25

Gln Lys Ala His Leu Ile Arg

1 5

<210> 26

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

peptide

<400> 26

Gln Lys Gly Thr Leu Gly Glu

1 5

<210> 27

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 27

Arg Arg Asp Asn Leu His Ser

1 5

<210> 28

<211> 1727

<212> RNA

<213> Artificial Sequence

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

polynucleotide

<400> 28

gggagacaag cuuugaauua caagcuugcu uguuucuuuuu gcagaagcuc agaauaaacg 60

cuacaacuuug gcagaucgaa uucgccaugg acuacaaaga ccaugacggu gauuauaaag 120

aucaugacau cgauuacaag gaugacgaug acaagauggc ccccaagaag aagaggaagg 180

ucggcaucca cggguuaccc gccgcuaugg cugagaggcc cuuccagugu cgaaucugca 240

ugcagaacuu cagugaccag uccaaccugc gcgcacacau ccgcacccac accggcgaga 300

agccuuuugc cugugacauu ugugggagga aauuugcccg caacuucucc cugaccaugc 360

auaccaagau acacacgggc agccaaaagc cciuccagug ucgaucugc augcagaacu 420

ucaguuccac cggcaaccug accaaccaca uccgcaccca caccggcgag aagccuuuug 480

ccugugacau uugugggagg aauuugcca ccuccggcuc ccugacccgc cauaccaaga 540

uacacacgca cccgcgcgcc ccgaucccga agccciucca gugucgaauc ugcaugcaga 600

acuucaguga ccaguccaac cugcgcgcac acauccgcac ccacaccggc gagaagccuu 660

uugccuguga cauuuugggg agggaaauuug ccgcacagug uugucuguuc caccauacca 720

agauacaccu gcggggaucc aucagcagag ccagaccacu gaacccgcac ccggagcugg 780

aggagaagaa guccgagcug cggcacaagc ugaaguacgu gccccacgag uacaucgagc 840

ugaucgagau cggcaggaac agcaccagg accgcauccu ggagaugaag gugauggagu 900

ucuucaugaa gguguacggc uacagggaa agcaccuggg cggaaggcaga aagccugacg 960

gcccacaucua uacagugggc agccccaucg auuacggcgu gaucguggac acaaaggccu 1020

acagcggcgg cuacaaucug ccuaucggcc aggccgacga gauggagaga uacguggagg 1080

agaaccagac cgggauaag caccucaacc ccaacgagug guggaaggug uacccuagca 1140

gcgugaccga guucaaguuc cuguucguga gcggccacuu caagggcaac uacaaggccc 1200

agcugaccag gcugaaccac aucaccaacu gcaauggcgc cgugcugagc guggaggagc 1260

ugcugaucgg cggcgagaaug aucaaagccg gcacccugac acuggaggag gugcggcgca 1320

aguucaacaa cggcgagaauc aacuucagau cuugauaacu cgagucuaga agcucgcuu 1380

cuugcugucc aauuuucuauu aaagguuccu uuguuuccua aguccaacua cuaaacuggg	1440
ggauauuaug aagggccuug agcaucugga uucugccuaa uaaaaaacau uuauuuucau	1500
ugcugcgcua gaagcucgcu uucuugcugu ccaauuuucua uuaaggguuc cuuuguucc	1560
uaaguccaac uacuaaacug ggggauauua ugaaggccu ugagcaucug gauucugccu	1620
aauaaaaaac auuuauuuuc auugcugcgg gacauucuua auuaaaaaaaaaaaaaaa	1680
aaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaaaaaaaaaa aaacuag	1727

<210> 29

<211> 422

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 29

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1 5 10 15

Tyr Lys Asp Asp Asp Lys Met Ala Pro Lys Lys Arg Lys Val

20 25 30

Gly Ile His Gly Val Pro Ala Ala Met Ala Glu Arg Pro Phe Gln Cys

35 40 45

Arg Ile Cys Met Gln Asn Phe Ser Asp Gln Ser Asn Leu Arg Ala His

50 55 60

Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly

65 70 75 80

Arg Lys Phe Ala Arg Asn Phe Ser Leu Thr Met His Thr Lys Ile His

85 90 95

Thr Gly Ser Gln Lys Pro Phe Gln Cys Arg Ile Cys Met Gln Asn Phe

100 105 110

Ser Ser Thr Gly Asn Leu Thr Asn His Ile Arg Thr His Thr Gly Glu

115 120 125

Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Thr Ser Gly

130 135 140

Ser Leu Thr Arg His Thr Lys Ile His Thr His Pro Arg Ala Pro Ile

145	150	155	160
Pro Lys Pro Phe Gln Cys Arg Ile Cys Met Gln Asn Phe Ser Asp Gln			
165	170	175	
Ser Asn Leu Arg Ala His Ile Arg Thr His Thr Gly Glu Lys Pro Phe			
180	185	190	
Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala Ala Gln Cys Cys Leu Phe			
195	200	205	
His His Thr Lys Ile His Leu Arg Gly Ser Ile Ser Arg Ala Arg Pro			
210	215	220	
Leu Asn Pro His Pro Glu Leu Glu Lys Lys Ser Glu Leu Arg His			
225	230	235	240
Lys Leu Lys Tyr Val Pro His Glu Tyr Ile Glu Leu Ile Glu Ile Ala			
245	250	255	
Arg Asn Ser Thr Gln Asp Arg Ile Leu Glu Met Lys Val Met Glu Phe			
260	265	270	
Phe Met Lys Val Tyr Gly Tyr Arg Gly Lys His Leu Gly Ser Arg			
275	280	285	
Lys Pro Asp Gly Ala Ile Tyr Thr Val Gly Ser Pro Ile Asp Tyr Gly			
290	295	300	
Val Ile Val Asp Thr Lys Ala Tyr Ser Gly Gly Tyr Asn Leu Pro Ile			
305	310	315	320
Gly Gln Ala Asp Glu Met Glu Arg Tyr Val Glu Glu Asn Gln Thr Arg			
325	330	335	
Asp Lys His Leu Asn Pro Asn Glu Trp Trp Lys Val Tyr Pro Ser Ser			
340	345	350	
Val Thr Glu Phe Lys Phe Leu Phe Val Ser Gly His Phe Lys Gly Asn			
355	360	365	
Tyr Lys Ala Gln Leu Thr Arg Leu Asn His Ile Thr Asn Cys Asn Gly			
370	375	380	
Ala Val Leu Ser Val Glu Glu Leu Leu Ile Gly Gly Glu Met Ile Lys			
385	390	395	400

Ala Gly Thr Leu Thr Leu Glu Glu Val Arg Arg Lys Phe Asn Asn Gly

405

410

415

Glu Ile Asn Phe Arg Ser

420

<210> 30

<211> 1680

<212> RNA

<213> Artificial Sequence

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

polynucleotide

<400> 30

gggagacaag cuugaaauaca agcuugcuug uucuuuuugc agaagcucag aauaaacgcu	60
caacuuuggc agaucgaaau cgccuagaga ucuggcggcg gagagggcag aggaagucuu	120
cuaaccugcg gugacgugga ggagaaauccc ggcccuagga ccauggacua caaagaccau	180
gacggugauu auaaagauga ugacaucgau uacaagggaa acgaugacaa gauggcccc	240
aagaagaaga ggaaggucgg cauucauggg guacccgcgg cuauggcuga gaggccciuc	300
cagugucgaa ucugcaugca gaaguuugcc cgcaacgacc accgcaccac ccauaccaag	360

auacacacgg gcgagaagcc cuuccagugu cgaaucugca ugcagaacuu cagucagaag	420
gcccaccuga uccgcccacau ccgcacccac accggcgaga agccuuuugc cugugacauu	480
ugugggagga aauuugccca gaagggcacc cugggcgagc auaccaagau acacacggg	540
ucucagaagc cciuccagug ucgaaucugc augcagaacu ucagucgcgg ccgcgaccug	600
ucccgccaca uccgcaccca caccggcgag aagccuuuug ccugugacau uugugggagg	660
aaauuugccc gccgcgacaa ccugcacucc cauaccaaga uacaccugcg gggauccc	720
cuggugaaga gcgagcugga ggagaagaag uccgagcugc ggcacaagcu gaaguacug	780

ccccacgagu acaucgagcu gaucgagauc gccaggaaca gcacccagga ccgcauccug	840
gagaugaagg ugauggaguu cuucaugaag guguacggcu acagggaaa gcaccuggc	900
ggaagcagaa acccugacgg cgccaucuau acagugggca gcccuaucga uuacggcgug	960
aucguggaca caaaggccua cagcggcggc uacaucugc cuaucggcca ggccgacgag	1020
augcagagau acgugaagga gaaccagacc cgaaauaagc acaucaaccc caacgagugg	1080
uggaaggugu acccuagcag cgugaccgag uucaaguucc uguucgugag cggccacuuc	1140
agcggcaacu acaaggccca gcugaccagg cugaaccgca aaaccaacug caauggcgcc	1200

gugcugagcg	uggaggagcu	gcugaucggc	ggcgagaua	ucaaagccgg	cacccugaca	1260
cuggaggagg	ugcggcgcaa	guucaacaac	ggcgagaua	acuuucugaua	acucgagucu	1320
agaagcucgc	uuucuugcug	uccaauuuu	auuaaagguu	ccuuuguucc	cuaaguccaa	1380
cuacuaacu	ggggauuuu	augaaggcc	uugagcaucu	ggauucugcc	uaauaaaaaa	1440
cauuuauuuu	cauugcugcg	cuagaagcuc	gcuuucuugc	uguccaauuu	cuauuaagg	1500
uuccuuuguu	cccuuagucc	aacuacuaaa	cugggggaua	uuaugaaggg	ccuugagcau	1560
cuggauucug	ccuaauaaaaa	aacauuuuuu	uucauugcug	cgggacauuc	uuaauuaaaa	1620

aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaacuag	1680
------------	------------	------------	------------	------------	------------	------

<210> 31

<211> 381

<212> PRT

<213> Artificial Sequence

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

polypeptide

<400> 31

Met Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp His Asp Ile Asp

1 5 10 15

Tyr Lys Asp Asp Asp Lys Met Ala Pro Lys Lys Arg Lys Val

20 25 30

Gly Ile His Gly Val Pro Ala Ala Met Ala Glu Arg Pro Phe Gln Cys

35 40 45

Arg Ile Cys Met Gln Lys Phe Ala Arg Asn Asp His Arg Thr Thr His

50 55 60

Thr Lys Ile His Thr Gly Glu Lys Pro Phe Gln Cys Arg Ile Cys Met

65 70 75 80

Gln Asn Phe Ser Gln Lys Ala His Leu Ile Arg His Ile Arg Thr His

85 90 95

Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys Gly Arg Lys Phe Ala

100 105 110

Gln Lys Gly Thr Leu Gly Glu His Thr Lys Ile His Thr Gly Ser Gln

115 120 125

Lys Pro Phe Gln Cys Arg Ile Cys Met Gln Asn Phe Ser Arg Gly Arg

130	135	140
Asp Leu Ser Arg His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala		
145	150	155
Cys Asp Ile Cys Gly Arg Lys Phe Ala Arg Arg Asp Asn Leu His Ser		
165	170	175
His Thr Lys Ile His Leu Arg Gly Ser Gln Leu Val Lys Ser Glu Leu		
180	185	190
Glu Glu Lys Ser Glu Leu Arg His Lys Leu Lys Tyr Val Pro His		
195	200	205
Glu Tyr Ile Glu Leu Ile Glu Ile Ala Arg Asn Ser Thr Gln Asp Arg		
210	215	220
Ile Leu Glu Met Lys Val Met Glu Phe Phe Met Lys Val Tyr Gly Tyr		
225	230	235
Arg Gly Lys His Leu Gly Gly Ser Arg Lys Pro Asp Gly Ala Ile Tyr		
245	250	255
Thr Val Gly Ser Pro Ile Asp Tyr Gly Val Ile Val Asp Thr Lys Ala		
260	265	270
Tyr Ser Gly Gly Tyr Asn Leu Pro Ile Gly Gln Ala Asp Glu Met Gln		
275	280	285
Arg Tyr Val Lys Glu Asn Gln Thr Arg Asn Lys His Ile Asn Pro Asn		
290	295	300
Glu Trp Trp Lys Val Tyr Pro Ser Ser Val Thr Glu Phe Lys Phe Leu		
305	310	315
Phe Val Ser Gly His Phe Ser Gly Asn Tyr Lys Ala Gln Leu Thr Arg		
325	330	335
Leu Asn Arg Lys Thr Asn Cys Asn Gly Ala Val Leu Ser Val Glu Glu		
340	345	350
Leu Leu Ile Gly Gly Glu Met Ile Lys Ala Gly Thr Leu Thr Leu Glu		
355	360	365
Glu Val Arg Arg Lys Phe Asn Asn Gly Glu Ile Asn Phe		
370	375	380

<210> 32

<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> 32

acacgacgct ctccgatct nnnnagtct cttctacccc accca 45

<210> 33

<211> 46

<212> DNA

<213> Artificial Sequence

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

<400> 33

gacgtgtgct ctccgatct ctactcttag acataacaca ccaggg 46

<210> 34

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 34

Arg Gly Arg Asp Leu Ser Arg

1

5