



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0088911
(43) 공개일자 2018년08월07일

(51) 국제특허분류(Int. Cl.)
C12N 5/0783 (2010.01) A61K 35/17 (2014.01)
A61K 48/00 (2006.01) C07K 14/705 (2006.01)
C07K 14/725 (2006.01) C07K 14/74 (2006.01)
C07K 16/28 (2006.01)
(52) CPC특허분류
C12N 5/0636 (2013.01)
A61K 35/17 (2013.01)
(21) 출원번호 10-2018-7020453
(22) 출원일자(국제) 2016년12월15일
심사청구일자 없음
(85) 번역문제출일자 2018년07월17일
(86) 국제출원번호 PCT/US2016/066988
(87) 국제공개번호 WO 2017/106537
국제공개일자 2017년06월22일
(30) 우선권주장
62/269,410 2015년12월18일 미국(US)
(뒷면에 계속)

(71) 출원인
상가모 테라퓨틱스, 인코포레이티드
미국 캘리포니아 94804 리치몬드 스위트 에이100
커널 블러바드 501
(72) 발명자
리, 게리 케이.
미국 94804 캘리포니아주 리치몬드 스위트
에이100 커널 블러바드 501 상가모 테라퓨틱스,
인코포레이티드 내
파슨, 데이빗
미국 94804 캘리포니아주 리치몬드 스위트
에이100 커널 블러바드 501 상가모 테라퓨틱스,
인코포레이티드 내
장, 레이
미국 94804 캘리포니아주 리치몬드 스위트
에이100 커널 블러바드 501 상가모 테라퓨틱스,
인코포레이티드 내
(74) 대리인
양영준, 김영

전체 청구항 수 : 총 13 항

(54) 발명의 명칭 MHC 세포 수용체의 표적화된 붕괴

(57) 요약

본원에는 세포 생존 능력을 보존할 수 있는 조건 하에 징크 핑거 단백질 및 절단 도메인 또는 절단 절반-도메인을 포함하는 징크 핑거 뉴클레아제 (ZFN)를 사용하여, MHC 유전자를 불활성화시키는 방법 및 조성물이 개시된다. ZFN을 코딩하는 폴리뉴클레오티드, ZFN을 코딩하는 폴리뉴클레오티드를 포함하는 벡터 및 ZFN을 코딩하는 폴리뉴클레오티드를 포함하는 세포 및/또는 ZFN을 포함하는 세포가 또한 제공된다.

대표도

엑손 1

5' cgagATGTCTCGCTCCGT **A** GCTGTGCTCGCGCTACTCTCTCTTTCTGGCCTGGAGGCTATCCAGC

엑손 2

5' GTACTCCAAAGATTCAAGTTTACTCACGTCATCCAGCAGAGAATGGAAAGT **C** CAAATTTCCTGAATTGC **D** CAAATTCTGG
GTTTCATCCATCCGACATTGAAGTTGACTTACTGAAGAATGGAGAGAGAAATTGAAAAGTGAGCATTCACTTGTCT
E **G**
TTCAGCAAGGACTGGTCTTTCTATCTCTGTACTACACTGAATTCACCCCACTGAAAAAGATGAGTATGCCTGCCGTG
TGAACCATGTGACTTTGTACAGCCCAAGATAGTTAAGTGGG

(52) CPC특허분류

A61K 48/00 (2013.01)
C07K 14/7051 (2013.01)
C07K 14/70521 (2013.01)
C07K 14/70539 (2013.01)
C07K 14/70596 (2013.01)
C07K 16/2803 (2013.01)
C07K 2319/33 (2013.01)
C12N 2510/00 (2013.01)

(30) 우선권주장

62/305,097 2016년03월08일 미국(US)
62/329,439 2016년04월29일 미국(US)

명세서

청구범위

청구항 1

베타 2 마이크로글로불린 (B2M) 유전자의 발현이 B2M 유전자의 엑손 1 또는 엑손 2의 변형에 의해 조정되는 것인, 단리된 세포.

청구항 2

제1항에 있어서, 변형이 서열식별번호: 6-48 또는 137 내지 205 중 어느 것에 제시된 바와 같은 표적 부위에 결합하는 DNA-결합 도메인 및 기능적 도메인을 포함하는 외인성 융합 분자에 의해 이루어지는 것인 세포.

청구항 3

제1항에 있어서, 서열식별번호: 6-48 또는 137 내지 205 중 하나 이상 이내; 서열식별번호: 6-48 또는 137 내지 205의 어느 한 측 (플랭킹 게놈 서열) 상의 1-10개 염기 쌍 이내; 또는 GGCCTTA, TCAAAT, TCAAATT, TTACTGA 및/또는 AATTGAA 이내에서의 삽입 및/또는 결실을 포함하는 세포.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 불활성화된 T-세포 수용체 유전자, PD1 및/또는 CTLA4 유전자를 추가로 포함하는 세포.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 키메라 항원 수용체 (CAR)를 코딩하는 트랜스진, 항체 커플링된 T-세포 수용체 (ACTR)를 코딩하는 트랜스진 및/또는 조작된 TCR을 코딩하는 트랜스진을 추가로 포함하는 세포.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, 림프성 세포, 줄기 세포, 또는 전구 세포인 세포.

청구항 7

제6항에 있어서, T-세포, 유도된 다능성 줄기 세포 (iPSC), 배아 줄기 세포, 중간엽 줄기 세포 (MSC), 또는 조혈 줄기 세포 (HSC)인 세포.

청구항 8

제1항 내지 제7항 중 어느 한 항에 따른 세포를 포함하는 제약 조성물.

청구항 9

B2M 유전자의 엑손 1 또는 엑손 2에 결합하는 DNA-결합 도메인, 및 전사 조절성 도메인 또는 뉴클레아제 도메인을 포함하며, 여기서 DNA-결합 도메인은 표 1의 단일 행에 제시된 바와 같은 징크 핑거 단백질 (ZFP), 표 2B의 단일 행에 제시된 바와 같은 TALE-이펙터 단백질 또는 표 2A의 단일 행에 제시된 바와 같은 단일 가이드 RNA (sgRNA)를 포함하는 것인, 융합 분자.

청구항 10

제9항의 융합 분자를 코딩하는 폴리뉴클레오티드.

청구항 11

제10항에 있어서, 바이러스 벡터, 플라스미드 또는 mRNA인 폴리뉴클레오티드.

청구항 12

제1항 내지 제7항 중 어느 한 항에 따른 세포 또는 제8항에 따른 제약 조성물을 암에 걸린 대상체 내로 도입하는 것을 포함하는, 암을 치료 또는 예방하는 방법.

청구항 13

암에 걸린 대상체를 치료하기 위한, 제1항 내지 제7항 중 어느 한 항에 따른 세포, 제8항에 따른 제약 조성물 또는 제9항에 따른 융합 분자 또는 제10항에 따른 폴리뉴클레오티드의 용도.

발명의 설명

기술 분야

[0001] 관련 출원에 대한 상호 참조

[0002] 본 출원은 2015년 12월 18일에 출원된 미국 가출원 번호 62/269,410; 2016년 3월 8일에 출원된 미국 가출원 번호 62/305,097; 및 2016년 4월 29일에 출원된 미국 가출원 번호 62/329,439를 우선권 주장하며, 이들 가출원의 개시내용은 그 전문이 본원에 참조로 포함된다.

[0003] 본 개시내용은 림프구 및 줄기 세포를 포함한 인간 세포의 게놈 변형 분야에 관한 것이다.

배경 기술

[0004] 유전자 요법은 인간 치료제의 새로운 시대에 막대한 잠재력을 가지고 있다. 이러한 방법론은 표준 의료 행위에 의해서는 해결될 수 없었던 병태의 치료를 허용할 것이다. 유전자 요법은 트랜스진과 융합된 특이적 외인성 프로모터에 의해 제어될 수 있거나, 또는 게놈 내로의 삽입 부위에서 발견된 내인성 프로모터에 의해 제어될 수 있는 발현가능한 트랜스진의 삽입, 및 유전자 로커스의 붕괴 또는 교정과 같은 게놈 편집 기술의 많은 변이를 포함할 수 있다.

[0005] 트랜스진의 전달 및 삽입이 이러한 기술의 임의의 실제 구현을 위해 해결되어야만 하는 장애물의 예이다. 예를 들어, 각종 유전자 전달 방법이 치료용으로 잠재적으로 이용가능하지만, 모두 안전성, 내구성 및 발현 수준 간의 실질적인 상충 관계를 수반한다. 예피솜으로서 트랜스진을 제공하는 방법 (예를 들어 기본 아데노바이러스 (Ad), 아데노 관련 바이러스 (AAV) 및 플라스미드 기반 시스템)은 일반적으로 안전하고 높은 초기 발현 수준을 산출시킬 수 있지만, 이러한 방법은 유사분열로 활성인 조직에서의 발현의 지속 기간을 제한할 수 있는 강력한 예피솜 복제가 결여된다. 반대로, 목적하는 트랜스진 (예를 들어, 통합되는 렌티바이러스 (LV))의 무작위 통합을 초래하는 전달 방법은 보다 지속적인 발현을 제공하지만, 무작위 삽입의 비표적화된 성질로 인해 수용자 세포에서 비조절된 성장을 유발시켜, 무작위로 통합된 트랜스진 카세트의 부근에서 종양 유전자의 활성화를 통하여 잠재적으로 악성 종양이 야기된다. 더욱이, 트랜스진 통합이 복제 구동된 손실을 피하긴 하지만, 트랜스진과 융합된 외인성 프로모터의 궁극적인 침묵을 방지시키지는 못한다. 시간이 지남에 따라, 그러한 침묵은 대다수의 비-특이적 삽입 이벤트에 대한 트랜스진 발현을 저하시킨다. 또한, 트랜스진의 통합은 모든 표적 세포에서 거의 일어나지 않으며, 이는 목적하는 치료 효과를 달성하기 위해 관심 트랜스진의 충분히 높은 발현 수준을 달성하는 것을 어렵게 만들 수 있다.

[0006] 최근 몇 년 동안, 부위-특이적 뉴클레아제 (예를 들어, 징크 핑거 뉴클레아제 (ZFN), 전사 활성화제-유사 이펙터 도메인 뉴클레아제 (TALEN), 조작된 crRNA/tracrRNA ('단일 가이드 RNA')를 수반한 CRISPR/Cas 시스템, 및 특이적 절단을 가이드하는 Cpf1/CRISPR 시스템 등)로의 절단을 이용하여 선택된 게놈 로커스 내로의 삽입을 편향시키게 하는, 트랜스진 통합에 대한 새로운 전략이 개발된 바 있다. 예를 들어, 미국 특허 번호 9,255,250; 9,045,763; 9,005,973; 8,956,828; 8,945,868; 8,703,489; 8,586,526; 6,534,261; 6,599,692; 6,503,717; 6,689,558; 7,067,317; 7,262,054; 7,888,121; 7,972,854; 7,914,796; 7,951,925; 8,110,379; 8,409,861; 미국 특허 공개 20030232410; 20050208489; 20050026157; 20050064474; 20060063231; 20080159996; 201000218264; 20120017290; 20110265198; 20130137104; 20130122591; 20130177983 및 20130177960 및 20150056705 참조. 추가로, 아르곤노트(Argonaute) 시스템 (예를 들어, 'TtAgo'로서 공지된 티. 썬포필루스 (*T. thermophilus*)로부터 유래됨)에 근거하여 표적화된 뉴클레아제가 개발 중에 있으며 (문헌 [Swartz et al. (2014) *Nature* 507(7491): 258-261] 참조), 이는 또한, 게놈 편집과 유전자 요법에 사용하기 위한 잠재력을 가질 수 있다. 이러한 뉴클레아제-매개 트랜스진 통합에 대한 접근 방식은 기존의 통합 접근 방식과 비교해서, 개선된 트랜스진 발현, 증가된 안전성 및 발현 지속성의 가능성을 제공하며 이는 근처의 종양 유전자의 활성화

또는 유전자 침묵의 위험을 최소화하기 위해 정확한 트랜스진 위치 설정을 허용하기 때문이다.

- [0007] T 세포 수용체 (TCR)는 T 세포의 선택적 활성화의 필수적인 부분이다. TCR의 항원 인식 부분은 항체와 일부 유사성이 있지만, 이는 전형적으로, α 와 β 의 2개쇄로 만들어지며, 이들쇄는 공동 어셈블리되어 이중-이량체를 형성한다. 항체 유사성은 TCR 알파와 베타 복합체를 코딩하는 단일 유전자를 함께 모으는 방식에 있다. TCR 알파 (TCR α)쇄 및 베타 (TCR β)쇄는 각각 2개 영역, 즉 C-말단 불변 영역 및 N-말단 가변 영역으로 구성된다. TCR 알파 및 베�타쇄를 코딩하는 게놈 로커스는, TCR α 유전자가 V 및 J 절편을 포함하는 반면, β 쇄 로커스는 V 및 J 절편 외에 D 절편을 포함한다는 점에서 항체 코딩 로커스와 유사하다. TCR β 로커스의 경우에는, 선택 프로세스 동안 선택되는 부가적으로 2개의 상이한 불변 영역이 있다. T 세포 발생 동안에는, 다양한 절편들이 재조합되어, 각각의 T 세포가 알파 및 베타쇄 내에 상보성 결정 영역 (CDR)으로 지칭되는 독특한 TCR 가변 부분을 포함하도록 하고, 신체는 그들의 독특한 CDR로 인해 항원 제시 세포에 의해 표시된 독특한 항원과 상호작용할 수 있는 T 세포의 큰 레퍼토리를 갖는다. 일단 TCR α 또는 β 유전자 재배열이 발생하면, 제2의 상응하는 TCR α 또는 TCR β 의 발현은, 각각의 T 세포가 '항원 수용체 대립 유전자 배제'로 지칭되는 프로세스에서 하나의 독특한 TCR 구조만을 발현하도록 억제된다 (문헌 [Brady et al., (2010) *J Immunol* 185:3801-3808] 참조).
- [0008] T 세포 활성화 동안, TCR은 항원 제시 세포의 주요 조직 적합성 복합체 (MHC) 상의 펩티드로서 표시된 항원과 상호작용한다. TCR에 의한 항원-MHC 복합체의 인식은 T 세포 자극을 야기시키고, 결국에는 기억 및 이펙터 림프구 내에서 T 헬퍼 세포 (CD4+)와 세포독성 T 림프구 (CD8+) 둘 다의 분화가 초래된다. 그런 다음 이들 세포는 하나의 특정한 항원에 반응할 수 있는 전체 T 세포 집단 내에서 활성화된 부분 모집단을 제공하기 위해 클로날 방식으로 확장될 수 있다.
- [0009] MHC 단백질은 2가지 부류, 즉 I 및 II이다. 부류 I MHC 단백질은 2가지 단백질, 즉 MHC1 부류 I 유전자에 의해 코딩된 막횡단 단백질인 α 쇄와, MHC 유전자 클러스터 내에 있지 않는 특정 유전자에 의해 코딩되는 작은 세포외 단백질인 β 2 마이크로글로불린쇄 (종종 B2M으로서 지칭됨)의 이중-이량체이다. α 쇄는 3개의 구형 도메인으로 폴딩되고, β 2 마이크로글로불린쇄가 연합되면, 이러한 구형 구조 복합체는 항체 복합체와 유사하다. 외래 펩티드는 가장 가변적이기도 한 2개의 대부분의 N-말단 도메인 상에 존재한다. 부류 II MHC 단백질 또한 이중-이량체이지만, 이러한 이중-이량체는 MHC 복합체 내의 유전자에 의해 코딩된 2개의 막횡단 단백질을 포함한다. 부류 I MHC:항원 복합체는 세포독성 T 세포와 상호작용하는 반면, 부류 II MHC는 항원을 헬퍼 T 세포에 제시한다. 또한, 부류 I MHC 단백질은 거의 모든 핵형성된 세포 및 혈소판 (및 마우스 내의 적혈구)에서 발현되는 경향이 있는 반면, 부류 II MHC 단백질은 보다 선택적으로 발현된다. 전형적으로, 부류 II MHC 단백질은 B 세포, 일부 대식세포 및 단핵구, 랑게르한스 세포, 및 수지상 세포 상에서 발현된다.
- [0010] 인간에게서 부류 I HLA 유전자 클러스터는 3가지 주요 로커스, 즉 B, C 및 A 뿐만 아니라 몇 가지 소수 로커스 (E, G 및 F 포함, 모두 염색체 6 상의 HLA 영역에서 발견됨)를 포함한다. 부류 II HLA 클러스터는 또한, 3가지 주요 로커스, 즉 DP, DQ 및 DR을 포함하고, 부류 I 유전자 클러스터와 부류 II 유전자 클러스터 둘 다는, 집단 내에 부류 I 유전자와 II 유전자 둘 다의 몇 가지 상이한 대립 유전자가 있다는 점에서 다형태성이다. 또한 HLA 기능에 있어 일정 역할을 하는 몇 가지 부속 단백질이 있다. β -2 마이크로글로불린은 샤프롱 (염색체 15 상에 위치한, B2M에 의해 코딩됨)으로서 기능하고, 세포 표면 상에 발현된 HLA A, B 또는 C 단백질을 안정화시키며, 또한 부류 I 구조 상의 항원 표시 홈을 안정화시킨다. 이는 혈청 및 소변에서 정상적으로 적은 양으로 발견된다.
- [0011] HLA는 이식 거부에 있어 중요한 역할을 한다. 이식 거부의 급성기는 약 1-3주 이내에 발생할 수 있고, 통상적으로 공여자 부류 I 및 부류 II HLA 분자에 대한 숙주 시스템의 감각으로 인해 공여자 조직 상에서의 숙주 T 림프구의 작용을 포함한다. 대부분의 경우에, 촉발되는 항원이 부류 I HLA이다. 최선의 성공을 위해서는, 공여자를 HLA에 대한 유형으로 분류시키고 가능한 한 완벽하게 환자 수용자와 일치시킨다. 그러나 높은 백분율의 HLA 동일성을 공유할 수 있는 패밀리 구성원 간의 기증조차도 여전히 종종 성공적이지 못하다. 따라서, 수용자 내에서 이식편 조직을 보존하기 위해, 환자는 종종, 거부 반응을 방지하기 위해 엄청난 면역 억제 요법을 받아야만 한다. 이러한 요법은 기회 감염으로 인한 합병증과 중대한 발병률을 유발시킬 수 있어, 환자가 이를 극복하는데 어려움을 겪을 수 있다. 부류 I 또는 II 유전자의 조절은 일부 종양의 존재 하에 붕괴될 수 있고, 이러한 붕괴는 환자의 예후에 영향을 줄 수 있다. 예를 들어, B2M 발현의 저하가 전이성 결장직장암에서 발견되었다 (Shrout et al. (2008) *Br J Canc* 98:1999). B2M은 MHC 부류 I 복합체를 안정화시키는데 있어서 중요한 역할을 하기 때문에, 특정 고형암에서의 B2M의 손실이 T 세포 구동된 면역 감시로부터의 면역 탈출의 메카니즘인 것으로 가정된 바 있다. 하락한 B2M 발현은 유전자 녹-아웃을 초래하는 B2M 코딩 서열 상의 특이적 돌연변이

및/또는 정상적인 IFN 감마 B2M 발현 조절의 억제에 따른 결과인 것으로 밝혀진 바 있다 (Shrout et al., 상기 문헌). 혼동스럽게도, 증가된 B2M은 또한 일부 유형의 암과 연관 있다. 소변에서의 증가된 B2M 수준은 전립선암, 만성 림프구성 백혈병 (CLL) 및 비호지킨 림프종을 포함한 몇 가지 암의 예후 인자로서 작용한다.

[0012] 입양 세포 요법 (ACT)은 전달된 세포가 환자의 암을 공격하고 제거하기 위해 종양 특이적 면역 세포를 환자에게 전달하는 것에 근거하여 개발 중인 암 요법의 형태이다. ACT는 환자 자신의 종양 덩어리로부터 단리되고 환자에게 다시 주입하기 위해 생체외에서 확장되는 T 세포인 종양 침윤성 림프구 (TIL)의 사용을 포함할 수 있다. 이러한 접근 방식은 한 연구에서 >50%의 장기간 반응물이 관찰되는 전이성 흑색종을 치료하는데 유망해지고 있다 (예를 들어, 문헌 [Rosenberg et al. (2011) *Clin Canc Res* 17(13): 4550] 참조). TIL은 종양에 존재하는 종양 관련 항원 (TAA)에 대해 특이적인 T-세포 수용체 (TCR)를 갖는 환자 자신의 세포의 혼합된 세트이기 때문에 유망한 세포 공급원이다 (Wu et al. (2012) *Cancer J* 18(2):160). 다른 접근 방식은 환자의 혈액으로부터 단리된 T 세포를 편집하여, 이들 세포가 일부 방식으로 종양에 반응하도록 조작되는 것을 포함한다 (Kalos et al. (2011) *Sci Transl Med* 3(95):95ra73).

[0013] 키메라 항원 수용체 (CAR)는 세포 표면 상에 발현된 특이적 분자 표적을 면역 세포가 표적으로 하도록 디자인된 분자이다. 가장 기본적인 형태로, 이들은 세포 외부에 발현된 특이성 도메인을 세포 내부의 시그널링 경로에 커플링시키는 세포 내로 도입된 수용체이므로, 특이성 도메인이 그의 표적과 상호작용할 때 그 세포가 활성화된다. 종종 CAR은 항원 특이적 도메인, 예컨대 scFv 또는 일부 유형의 수용체가 시그널링 도메인, 예컨대 ITAM 및 다른 공동 자극성 도메인과 융합되는 T-세포 수용체 (TCR)의 기능적 도메인을 모방하는 것으로 만들어진다. 이어서, 이들 구조물을 생체외에서 T-세포 내로 도입하여, T-세포를 표적 항원을 발현하는 세포의 존재 하에서 활성화될 수 있도록 하여, T-세포가 환자 내로 재도입될 때 비-MHC 의존적 방식 (문헌 [Chicaybam et al. (2011) *Int Rev Immunol* 30:294-311, Kalos 상기 문헌] 참조)으로 활성화된 T-세포에 의한 표적화된 세포에 대한 공격이 초래된다. 따라서, 조작된 TCR 또는 CAR로 생체외에서 변경된 T 세포를 이용한 입양 세포 요법은 몇 가지 유형의 질환에 대한 매우 유망한 임상적 접근 방식이다. 예를 들어, 표적화되는 암 및 그의 항원은 여포성 림프종 (CD20 또는 GD2), 신경모세포종 (CD171), 비호지킨 림프종 (CD19 및 CD20), 림프종 (CD19), 교모세포종 (IL13R α2), 만성 림프구성 백혈병 또는 CLL 및 급성 림프구성 백혈병 또는 ALL (둘 다 CD19)을 포함한다. 바이러스 특이적 CAR은 또한, HIV와 같은 바이러스를 정착시킨 세포를 공격하기 위해 개발되었다. 예를 들어, HIV 치료를 위해 Gp100에 대해 특이적인 CAR을 사용하여 임상 시험을 개시하였다 (Chicaybam, 상기 문헌).

[0014] ACTR (항체 커플링된 T-세포 수용체)은 외인적으로 공급된 항체에 결합할 수 있는 조작된 T 세포 성분이다. ACTR 성분에 대한 항체의 결합은 T 세포가 항체에 의해 인식된 항원과 상호작용하도록 해주고, 그 항원과 직면할 때 T 세포를 포함하는 ACTR이 항원과 상호작용하도록 촉발된다 (미국 특허 공개 20150139943 참조).

[0015] 그러나 입양 세포 요법의 단점 중 하나는 이식된 세포의 잠재적 거부 반응을 피하기 위해 세포 생성물의 공급원이 환자 특이적 (자가)이어야만 한다는 것이다. 이로써, 연구원들은 이러한 거부 반응을 피하기 위해 환자 자신의 T 세포를 편집하는 방법을 개발하게 되었다. 예를 들어, 환자의 T 세포 또는 조혈 줄기 세포는 조작된 CAR, ACTR 및/또는 T 세포 수용체 (TCR)의 부가를 동반하여 생체외에서 조작된 다음, T 세포 체크 포인트 억제제, 예컨대 PD1 및/또는 CTLA4를 녹아웃시키기 위해 조작된 뉴클레아제로 추가로 처리될 수 있다 (PCT 공개 WO2014/059173 참조). 이러한 기술을 더 큰 환자 집단에 적용하기 위해서는, 보편적인 세포 집단 (동종이계)을 개발하는 것이 유리할 것이다. 또한, TCR의 녹아웃은, 일단 환자 내로 도입되면 이식편 대 숙주 질환 (GVHD) 반응을 일으킬 수 없는 세포가 생성될 것이다.

[0016] 따라서, MHC 유전자 발현을 변형 (예를 들어, B2M을 녹아웃)시키고/거나 T 세포에서의 TCR 발현을 녹아웃시키는 데 사용될 수 있는 방법 및 조성물에 대한 필요성이 대두된다.

발명의 내용

[0017] 본원에는 B2M 유전자의 부분적 또는 완전한 불활성화 또는 붕괴를 위한 조성물 및 방법, 및 내인성 TCR 및/또는 B2M의 붕괴 후에 또는 이러한 붕괴와 동시에, T 림프구 내에 목적하는 수준의 외인성 TCR, CAR 또는 ACTR 트랜스진으로 도입 및 발현시키기 위한 조성물 및 방법이 개시된다. 추가적으로, 본원에는 HLA 부류 I 널 T 세포, 줄기 세포, 조직 또는 전체 유기체, 예를 들어 그의 표면 상에 하나 이상의 HLA 수용체를 발현하지 않는 세포를 생성하기 위해 B2M 유전자를 결실 (불활성화) 또는 억제하기 위한 방법 및 조성물이 제공된다. 특정 실시양태에서, HLA 널 세포 또는 조직은 이식에 사용하는데 유리한 인간 세포 또는 조직이다. 바람직한 실시양태에서, HLC 널 T 세포는 입양 T 세포 요법에 사용하기 위해 제조된다.

[0018] 한 측면에서, 본원에는 베타 2 마이크로글로불린 (B2M) 유전자의 발현이 B2M 유전자의 엑손 1 또는 엑손 2의 변형에 의해 조정되는 것인 단리된 세포 (예를 들어, 진핵 세포, 예컨대 림프성 세포 포함 포유류 세포, 줄기 세포 (예를 들어, iPSC, 배아 줄기 세포, MSC 또는 HSC), 또는 전구 세포)가 기재된다. 특정 실시양태에서, 상기 변형은 서열식별번호(SEQ ID NO): 6-48 또는 137 내지 205 중 하나 이상에 제시된 바와 같은 서열에 대한 변형; 서열식별번호: 6-48 또는 137 내지 205의 어느 한 측 (플래킹 게놈 서열) 상의 1-5개 염기 쌍 이내, 1-10개 염기 쌍 이내 또는 1-20개 염기 쌍 이내에서의 변형; 또는 GGCCTTA, TCAAAT, TCAAATT, TTACTGA 및/또는 AATTGAA 이내에서의 변형이다. 상기 변형은 기능적 도메인 (예를 들어, 전사 조절성 도메인, 뉴클레아제 도메인) 및 DNA-결합 도메인을 포함하는 외인성 융합 분자에 의해서 일어날 수 있으며, 이는 하기를 포함하지만, 그에 제한되지 않는다: (i) 서열식별번호: 6-48 또는 137 내지 205 중 어느 것에 제시된 바와 같은 표적 부위에 결합하는 DNA-결합 도메인, 및 전사 인자가 B2M 유전자 발현을 변형시키는 전사 조절성 도메인을 포함하는 외인성 전사 인자를 포함하는 세포 및/또는 (ii) 서열식별번호: 6-48 또는 137 내지 205 중 하나 이상 이내에서; 서열식별번호: 6-48 또는 137 내지 205의 어느 한 측 (플래킹 게놈 서열) 상의 1-5개 염기 쌍 이내, 1-10개 염기 쌍 이내 또는 1-20개 염기 쌍 이내에서; 또는 GGCCTTA, TCAAAT, TCAAATT, TTACTGA 및/또는 AATTGAA 이내에서의 삽입 및/또는 결실을 포함하는 세포. 이러한 세포는 추가의 변형, 예를 들어 불활성화된 T-세포 수용체 유전자, PD1 및/또는 CTLA4 유전자 및/또는 키메라 항원 수용체 (CAR)를 코딩하는 트랜스진, 항체 커플링된 T-세포 수용체 (ACTR)를 코딩하는 트랜스진 및/또는 항체를 코딩하는 트랜스진을 포함할 수 있다. 본원에 기재된 바와 같은 임의의 세포를 포함하는 제약 조성물 뿐만 아니라 특정 대상체에게서 특정 장애 (예를 들어, 암)를 치료하기 위한 생체의 요법에서 상기 세포 및 제약 조성물을 사용하는 방법이 또한 제공된다.

[0019] 따라서, 한 측면에서, 본원에는 B2M 유전자의 발현이 조정되는 (예를 들어, 활성화되거나, 억제되거나 또는 불활성화되는) 것인 세포가 기재된다. 바람직한 실시양태에서, B2M의 엑손 1 또는 엑손 2가 조정된다. 이러한 조정은 B2M 유전자에 결합하고 B2M 발현을 조절하는 외인성 분자 (예를 들어, DNA-결합 도메인 및 전사 활성화 또는 억제 도메인을 포함하는 조작된 전사 인자)에 의해서 일어날 수 있고/거나 B2M 유전자의 서열 변형을 통해서 일어날 수 있다 (예를 들어, B2M 유전자를 절단시키고 삽입 및/또는 결실에 의해 상기 유전자 서열을 변형시키는 뉴클레아제를 사용함). 특정 실시양태에서, 하나 이상의 부가 유전자 (예를 들어, TCR 유전자, 예컨대 TCRA 유전자)의 발현이 또한 조정된다. 일부 실시양태에서, 부류 I HLA 복합체가 탈안정화되도록 B2M 유전자의 녹아웃을 유발시키는 조작된 뉴클레아제를 포함하는 세포가 기재된다. 바람직한 실시양태에서, 부류 I HLA의 탈안정화는 세포의 표면 상에서 부류 I HLA 복합체의 현저한 손실을 초래한다. 다른 실시양태에서, B2M 유전자의 발현이 조정되도록 외인성 분자, 예를 들어 조작된 전사 인자 (TF)를 포함하는 세포가 기재된다. 일부 실시양태에서, 상기 세포는 T 세포이다. 추가로, B2M 유전자의 발현이 조정되는 것인 세포가 기재되며, 여기서 이러한 세포는 적어도 하나의 외인성 트랜스진을 포함하거나 또는 적어도 하나의 내인성 유전자의 부가의 녹아웃을 포함하거나 또는 그의 조합을 포함하도록 추가로 조작된다. 외인성 트랜스진은 TCR 유전자 내로 통합될 수 있고/거나 (예를 들어, TCR 유전자가 녹아웃되는 경우), B2M 유전자 내로 통합될 수 있고/거나 비-TCR, 비-B2M 로커스, 예컨대 안전한 정착 로커스 내로 통합될 수 있다. 일부 경우에, 외인성 트랜스진은 ACTR, 조작된 TCR 및/또는 CAR을 코딩한다. 트랜스진 구조물은 HDR- 또는 NHEJ-구동된 프로세스에 의해 삽입될 수 있다. 일부 측면에서, 조정된 B2M 발현을 수반하는 세포는 그의 세포 표면 상에 상당한 부류 I HLA가 결여되고, 적어도 외인성 ACTR 및/또는 외인성 CAR을 추가로 포함한다. B2M 조정인자를 포함하는 일부 세포는 하나 이상의 체크 포인트 억제제 유전자의 녹아웃을 추가로 포함한다. 일부 실시양태에서, 이러한 체크 포인트 억제제는 PD1이다. 다른 실시양태에서, 체크 포인트 억제제는 CTLA4이다. 추가 측면에서, B2M 조정된 세포는 PD1 녹아웃 및 CTLA4 녹아웃을 포함한다. 일부 실시양태에서, 상기 세포는 TCR 유전자에서 추가로 변형된다. 일부 실시양태에서, 조정된 TCR 유전자는 TCR β 를 코딩하는 유전자 (TCRB)이다. 일부 실시양태에서 이는 상기 유전자의 불변 영역 (TCR β 불변 영역, 또는 TRBC)의 표적화된 절단을 통해서 달성된다. 특정 실시양태에서, 조정된 TCR 유전자는 TCR α 를 코딩하는 유전자 (TCRA)이다. 추가 실시양태에서, 삽입은 TCR α 유전자의 불변 영역 (본원에서 "TRAC" 서열로서 지칭됨)의 표적화된 절단을 포함한, TCR 유전자의 불변 영역의 표적화된 절단을 통해서 달성된다. 일부 실시양태에서, B2M-유전자 변형된 세포는 TCR 유전자, HLA-A, -B, -C 유전자, 또는 TAP 유전자, 또는 그의 임의의 조합에서 추가로 변형된다. 다른 실시양태에서, HLA 부류 II에 대한 조절인자인 CTIIA가 또한 변형된다.

[0020] 특정 실시양태에서, 본원에 기재된 세포는 B2M 유전자에 대한 변형 (예를 들어, B2M을 억제하기 위한 조작된 TF의 결합, 결실 및/또는 삽입) (예를 들어, 엑손 1 또는 엑손 2의 변형)을 포함한다. 특정 실시양태에서, 상기 변형은 서열식별번호: 6-48 및/또는 137-205의 변형이며, 이는 이들 서열 중 어느 것 이내에서 및/또는 B2M 유전자 내의 이들 서열을 플래킹하는 유전자 (게놈) 서열의 1-50개 염기 쌍 (이들 사이의 임의의 값, 예컨대 1-

5개, 1-10개 또는 1-20개 염기 쌍 포함) 이내에서 하나 이상의 뉴클레오티드에 결합하고/거나, 이를 절단하고/거나, 삽입하고/거나 결실시키는 것에 의한 변형을 포함한다. 특정 실시양태에서, 상기 세포는 B2M 유전자 (예를 들어, 엑손 1 및/또는 2, 도 1 참조) 이내의 하기 서열 중 하나 이상 이내에서의 변형 (결합, 절단, 삽입 및/또는 결실)을 포함한다: GGCCTTA, TCAAATT, TCAAAT, TTACTGA 및/또는 AATTGAA. 특정 실시양태에서, 상기 변형은 B2M 발현이 조정되도록, 예를 들어 억제 또는 활성화되도록 본원에 기재된 바와 같이 조작된 TF의 결합을 포함한다. 다른 실시양태에서, 상기 변형은 뉴클레아제(들) 결합 (표적) 및/또는 절단 부위(들)에서 또는 그 근처에서의 유전적 변형 (뉴클레오티드 서열의 변경)이며, 이는 절단 부위(들) 및/또는 결합 부위의 상류, 하류 및/또는 1개 이상의 염기 쌍을 포함함 1-300개 염기 쌍 (또는 그 사이의 임의의 수의 염기 쌍) 이내에서의 서열에 대한 변형; 결합 및/또는 절단 부위(들)를 포함하고/거나 그의 어느 한 측 상에서 1-100개 염기 쌍 (또는 그 사이의 임의의 수의 염기 쌍) 이내에서의 변형; 결합 및/또는 절단 부위(들)를 포함하고/거나 그의 어느 한 측 상에서 (예를 들어, 1 내지 5개, 1 내지 10개, 1-20개 또는 그 초과)의 염기 쌍) 1 내지 50개 염기 쌍 (또는 그 사이의 임의의 수의 염기 쌍) 이내에서의 변형; 및/또는 뉴클레아제 결합 부위 및/또는 절단 부위 이내의 하나 이상의 염기 쌍에 대한 변형을 포함하지만, 이에 제한되지 않는다. 특정 실시양태에서, 상기 변형은 서열식별 번호: 6-48 또는 137-205 중 어느 것을 둘러싸고 있는 B2M 유전자 서열에서 또는 그 근처에서이다 (예를 들어, 1-300개, 1-50개, 1-20개, 1-10개 또는 1-5개 또는 그 초과)의 염기 쌍 또는 그 사이의 임의의 수의 염기 쌍). 특정 실시양태에서, 상기 변형은 서열식별번호: 6-48 또는 137-205에 제시된 서열 중 하나 이상 이내에서 또는 B2M 유전자 (예를 들어, 엑손 1 및/또는 엑손 2)의 GGCCTTA, TCAAATT, TCAAAT, TTACTGA 및/또는 AATTGAA 이내에서의 B2M 유전자의 변형, 예를 들어, 이들 서열 중 하나 이상에 대한 1개 이상의 염기 쌍의 변형을 포함한다. 특정 실시양태에서, 뉴클레아제-매개된 유전적 변형은 쌍을 이룬 표적 부위 사이에서이다 (표적을 절단하기 위해 이량체가 사용되는 경우). 뉴클레아제-매개된 유전적 변형은 임의의 수의 염기 쌍의 삽입 및/또는 결실을 포함할 수 있으며, 이는 임의의 길이의 비-코딩 서열 및/또는 임의의 길이의 트랜스젠의 삽입 및/또는 1개 염기 쌍 내지 1000 kb 초과 (또는 1-100개 염기 쌍, 1-50개 염기 쌍, 1-30개 염기 쌍, 1-20개, 1-10개, 또는 1-5개 염기 쌍을 포함하지만, 이에 제한되지 않는 그 사이의 임의의 값)의 결실을 포함한다.

[0021] 본 발명의 변형된 세포는 림프성 세포 (예를 들어, T-세포), 줄기/전구 세포 (예를 들어, 유도된 다능성 줄기 세포 (iPSC)), 배아 줄기 세포 (예를 들어, 인간 ES), 중간엽 줄기 세포 (MSC), 또는 조혈 줄기 세포 (HSC)일 수 있다. 줄기 세포는 전능성 또는 다능성 (예를 들어, 부분적으로 분화된, 예컨대 다능성 골수성 또는 림프성 줄기 세포인 HSC)일 수 있다. 다른 실시양태에서, 본 발명은 HLA 발현에 대한 널 표현형을 갖는 세포를 생성하는 방법을 제공한다. 이어서, 본원에 기재된 변형된 (B2M 로커스에서 변형된) 줄기 세포 중 어느 것을 분화시켜, 본원에 기재된 바와 같은 줄기 세포로부터 유래된 분화된 (생체내 또는 시험관내) 세포를 생성시킬 수 있다. 본원에 기재된 변형된 줄기 세포 중 어느 것은 다른 관심 유전자 (예를 들어 TCRA, TCRB, PD1, CTLA4 등)에서의 추가의 변형을 포함할 수 있다.

[0022] 또 다른 측면에서, 본원에 기재된 조성물 (변형된 세포) 및 방법은, 예를 들어, 특정 장애의 치료 또는 예방 또는 완화에 사용될 수 있다. 상기 방법은 전형적으로, (a) 뉴클레아제 (예를 들어, ZFN 또는 TALEN) 또는 뉴클레아제 시스템, 예컨대 조작된 crRNA/tracrRNA를 수반한 CRISPR/Cas 또는 Cpf1/CRISPR을 사용하거나, 또는 조작된 전사 인자 (예를 들어 ZFN-TF, TALE-TF, Cpf1-TF 또는 Cas9-TF)를 사용하여 단리된 세포 (예를 들어, T-세포 또는 림프구)에서 내인성 B2M 유전자를 절단하거나 또는 하향 조절하여, 이러한 B2M 유전자가 불활성화되거나 또는 하향 조정되도록 하는 단계; 및 (b) 상기 세포를 대상체 내로 도입함으로써, 상기 장애를 치료 또는 예방하는 단계를 포함한다.

[0023] 일부 실시양태에서, TCR α 를 코딩하는 유전자 (TCRA) 및/또는 TCR β 를 코딩하는 유전자 (TCRB)는 또한, 불활성화되거나 또는 하향 조정된다. 일부 실시양태에서 불활성화는 이러한 유전자의 불변 영역 (TCR β 불변 영역, 또는 TRBC)의 표적화된 절단을 통하여 달성된다. 바람직한 실시양태에서, TCR α 를 코딩하는 유전자 (TCRA)는 불활성화되거나 또는 하향 조정된다. 추가의 바람직한 실시양태에서, 상기 장애는 암 또는 감염성 질환이다. 추가의 바람직한 실시양태에서 불활성화는 이러한 유전자의 불변 영역 (TCR α 불변 영역, TRAC로 약칭됨)의 표적화된 절단을 통하여 달성된다.

[0024] 상기 전사 인자 및/또는 뉴클레아제(들)는 mRNA로서, 단백질 형태로 및/또는 뉴클레아제(들)를 코딩하는 DNA 서열로서 세포 내로 도입될 수 있다. 특정 실시양태에서, 대상체 내로 도입된 상기 단리된 세포는 부가의 게놈 변형, 예를 들어, (절단된 B2M, TCR 유전자 또는 다른 유전자, 예를 들어 안전한 정착 유전자 또는 로커스 내로) 통합된 외인성 서열 및/또는 부가의 유전자, 예를 들어 하나 이상의 HLA 및/또는 TAP 유전자의 불활성화 (예를 들어, 뉴클레아제-매개됨)를 추가로 포함한다. 상기 외인성 서열은 벡터 (예를 들어 Ad, AAV, LV)를 통

하여 도입되거나 또는 전기천공과 같은 기술을 사용함으로써 도입될 수 있다. 일부 실시양태에서, 단백질은 세포 압착에 의해 세포 내로 도입된다 (문헌 [Kollmannsperger et al. (2016) *Nat Comm* 7, 10372 doi:10.1038/ncomms10372] 참조). 일부 측면에서, 상기 조성물은 단리된 세포 단편 및/또는 분화된 (부분적 또는 완전히) 세포를 포함할 수 있다.

[0025] 일부 측면에서, 성숙한 세포가 세포 요법, 예를 들어, 입양 세포 전달을 위해 사용될 수 있다. 다른 실시양태에서, T 세포 이식에 사용하기 위한 세포는 또 다른 관심 유전자 변형을 함유한다. 한 측면에서, T 세포는 암 마커에 대해 특이적인 삽입된 키메라 항원 수용체 (CAR)를 함유한다. 추가 측면에서, 이러한 삽입된 CAR은 B 세포 악성 종양에 특징적인 CD19 마커에 대해 특이적이다. 이러한 세포는 HLA를 일치시킬 필요없이 환자를 치료하기 위한 치료 조성물에 유용할 것이므로, 치료를 필요로 하는 임의의 환자를 위한 "쉽게 사용할 수 있는" 치료제로서 사용될 수 있을 것이다.

[0026] 또 다른 측면에서, B2M-조정된 (변형된) T 세포는 삽입된 항체 커플링된 T-세포 수용체 (ACTR) 공여자 서열을 함유한다. 일부 실시양태에서, ACTR 공여자 서열은 B2M 또는 TCR 유전자 내로 삽입되어 뉴클레아제-유도된 절단 후에 그러한 유전자의 발현을 붕괴시킨다. 다른 실시양태에서, 상기 공여자 서열은 "안전한 정착" 로커스, 예컨대 AAVS1, HPRT, 알부민 및 CCR5 유전자 내로 삽입된다. 일부 실시양태에서, ACTR 서열은 이러한 ACTR 공여자 서열이 조작된 뉴클레아제의 절단 부위를 플랭킹하는 서열과의 상동성을 갖는 플랭킹 상동성 아암(arm)을 포함하는 표적화된 통합을 통하여 삽입된다. 일부 실시양태에서 ACTR 공여자 서열은 프로모터 및/또는 다른 전사 조절성 서열을 추가로 포함한다. 다른 실시양태에서, ACTR 공여자 서열에는 프로모터가 결여된다. 일부 실시양태에서, ACTR 공여자는 TCR β 코딩 유전자 (TCRB) 내로 삽입된다. 일부 실시양태에서 삽입은 이러한 유전자의 불변 영역 (TCR β 불변 영역, 또는 TRBC)의 표적화된 절단을 통하여 달성된다. 바람직한 실시양태에서, ACTR 공여자는 TCR α 코딩 유전자 (TCRA) 내로 삽입된다. 추가의 바람직한 실시양태에서 삽입은 이러한 유전자의 불변 영역 (TCR α 불변 영역, TRAC로 약칭됨)의 표적화된 절단을 통하여 달성된다. 일부 실시양태에서, 상기 공여자는 TCRA 내의 엑손 서열 내로 삽입되는 반면, 다른 실시양태에서는 상기 공여자가 TCRA 내의 인트론 서열 내로 삽입된다. 일부 실시양태에서, TCR-조정된 세포는 CAR을 추가로 포함한다. 또한 추가 실시양태에서, B2M-조정된 세포는 부가적으로, HLA 유전자 또는 체크포인트 억제제 유전자에서 조정된다.

[0027] 또한, 본원에 기재된 바와 같은 변형된 세포 (예를 들어, 불활성화된 B2M 유전자를 수반하는 T 세포 또는 줄기 세포)를 포함하는 제약 조성물, 또는 본원에 기재된 바와 같은 B2M-결합 분자 (예를 들어, 조작된 전사 인자 및/또는 뉴클레아제) 중 하나 이상을 포함하는 제약 조성물이 제공된다. 특정 실시양태에서, 제약 조성물은 하나 이상의 제약상 허용되는 부형제를 추가로 포함한다. 상기 변형된 세포, B2M-결합 분자 (또는 이들 분자를 코딩하는 폴리뉴클레오티드) 및/또는 이들 세포 또는 분자를 포함하는 제약 조성물은 관련 기술분야에 공지된 방법, 예를 들어 정맥내 주입, 간동맥과 같은 특이적 혈관 내로의 주입, 또는 직접적인 조직 주사 (예를 들어, 근육)를 통하여 대상체 내로 도입된다. 일부 실시양태에서, 대상체는 상기 조성물로 치료 또는 완화될 수 있는 질환 또는 병태가 있는 성인 인간이다. 다른 실시양태에서, 대상체는 상기 조성물을 투여하여 상기 질환 또는 병태 (예를 들어, 암, 이식편 대 숙주 질환 등)를 예방, 치료 또는 완화시키는 소아 대상체이다.

[0028] 일부 측면에서, 상기 조성물 (ACTR을 포함하는 B2M 조정된 세포)은 외인성 항체를 추가로 포함한다. 또한, 미국 출원 번호 15/357,772 참조. 일부 측면에서, 이러한 항체는 특정 병태를 예방 또는 치료하도록 ACTR-포함 T 세포를 무장시키는데 유용하다. 일부 실시양태에서, 상기 항체는 종양 세포와 연관되거나 또는 암 관련 프로세스와 연관된 항원, 예컨대 EpCAM, CEA, gpA33, 뮤신, TAG-72, CAIX, PSMA, 폴레이트-결합 항체, CD19, EGFR, ERBB2, ERBB3, MET, IGF1R, EPHA3, TRAILR1, TRAILR2, RANKL, FAP, VEGF, VEGFR, $\alpha V\beta 3$ 및 $\alpha 5\beta 1$ 인테그린, CD20, CD30, CD33, CD52, CTLA4, 및 테나신을 인식한다 (Scott et al. (2012) *Nat Rev Cancer* 12:278). 다른 실시양태에서, 항체는 감염성 질환과 연관된 항원, 예컨대 HIV, HCV 등을 인식한다.

[0029] 또 다른 측면에서, 본원에는 B2M 유전자 내의 표적 부위에 결합하는 B2M DNA-결합 도메인 (예를 들어, ZFP, TALE 및 sgRNA)이 제공된다. 특정 실시양태에서, 상기 DNA 결합 도메인은 표 1의 단일 행에 제시된 바와 같은 순서로 인식 나선 영역을 수반한 ZFP; 표 2B의 단일 행에 제시된 바와 같은 RVD를 수반한 TAL-이펙터 도메인 DNA-결합 단백질; 및/또는 표 2A의 단일 행에 제시된 바와 같은 sgRNA를 포함한다. 이들 DNA-결합 단백질은 전사 조절성 도메인과 연합되어, B2M 발현을 조정하는 조작된 전사 인자를 형성시킬 수 있다. 대안적으로, 이들 DNA-결합 단백질은 하나 이상의 뉴클레아제 도메인과 연합되어, B2M 유전자에 결합하고 이를 절단시키는 조작된 징크 핑거 뉴클레아제 (ZFN), TALEN 및/또는 CRISPR/Cas 시스템을 형성시킬 수 있다. 특정 실시양태에서, CRISPR/Cas 시스템의 ZFN, TALEN 또는 단일 가이드 RNA (sgRNA)는 인간 B2M 유전자 내의 표적 부위에 결합한다. 전사 인자 또는 뉴클레아제 (예를 들어, ZFP, TALE, sgRNA)의 DNA-결합 도메인은 서열식별번호: 6

내지 48 또는 137-205 중 어느 것의 9, 10, 11, 12개 또는 그 초과 (예를 들어, 13, 14, 15, 16, 17, 18, 19, 20개 또는 그 초과)의 뉴클레오티드를 포함하는 B2M 유전자 내의 표적 부위에 결합할 수 있다. 징크 핑거 단백질은 1, 2, 3, 4, 5, 6개 또는 그 초과를 포함할 수 있으며, 각각의 징크 핑거는 표적 유전자 내의 표적 하위 부위와 특이적으로 접촉하는 인식 나선을 갖는다. 특정 실시양태에서, 징크 핑거 단백질은, 예를 들어 표 1에 제시된 바와 같은, 4개 또는 5개 또는 6개의 핑거 (F1, F2, F3, F4, F5 및 F6으로 지명되고, N-말단에서부터 C-말단으로 F1 내지 F4 또는 F5 또는 F6으로 정돈됨)를 포함한다. 다른 실시양태에서, 단일 가이드 RNA 또는 TAL-이펙터 DNA-결합 도메인은 서열식별번호: 6-48 또는 137-205 중 어느 것 (또는 서열식별번호: 6-48 중 어느 것 이내의 12개 이상의 염기 쌍)에 제시된 표적 부위에 결합할 수 있다. 예시적인 sgRNA 표적 부위는 서열식별번호: 16-48에 제시되고, 예시적인 TALEN 결합 부위는 표 2B에 제시된다 (서열식별번호: 137-205). 부가의 TALEN은 미국 특허 번호 8,586,526 및 9,458,205에 기재된 바와 같은 정준 또는 비-정준 RVD를 사용하여 본원에 기재된 바와 같은 부위를 표적화하도록 디자인될 수 있다. 본원에 기재된 뉴클레아제 (ZFP, TALE 또는 sgRNA DNA-결합 도메인을 포함함)는 서열식별번호: 6-48 또는 137-205 중 어느 것을 포함하는 B2M 유전자 내에서의 유전적 변형을 만들 수 있으며, 이는 이들 서열 (서열식별번호: 6-48 또는 137-205) 중 어느 것 이내에서의 변형 (삽입 및/또는 결실) 및/또는 서열식별번호: 6-48 또는 137-205에 제시된 표적 부위 서열을 플랭킹하는 B2M 유전자 서열에 대한 변형, 예를 들어 하기 서열: GGCCTTA, TCAAATT, TCAAAT, TTACTGA 및/또는 AATTGAA 중 하나 이상 이내의 B2M 유전자의 엑손 1 또는 엑손 2 이내에서의 변형을 포함한다.

[0030] 또한, B2M 유전자의 엑손 1 또는 엑손 2에 결합하는 DNA-결합 도메인, 및 전사 조절성 도메인 또는 뉴클레아제 도메인을 포함하는 융합 분자가 제공되며, 여기서 상기 DNA-결합 도메인은 표 1의 단일 행에 제시된 바와 같은 징크 핑거 단백질 (ZFP), 표 2B의 단일 행에 제시된 바와 같은 TALE-이펙터 단백질 또는 표 2A의 단일 행에 제시된 바와 같은 단일 가이드 RNA (sgRNA)를 포함한다.

[0031] 본원에 기재된 단백질 중 어느 것은 절단 도메인 및/또는 절단 절반-도메인 (예를 들어, 야생형 또는 조작된 *FokI* 절단 절반-도메인)을 추가로 포함할 수 있다. 따라서, 본원에 기재된 뉴클레아제 (예를 들어, ZFN, TALEN, CRISPR/Cas 시스템) 중 어느 것에서, 뉴클레아제 도메인은 야생형 뉴클레아제 도메인 또는 뉴클레아제 절반-도메인 (예를 들어, *FokI* 절단 절반 도메인)을 포함할 수 있다. 다른 실시양태에서, 뉴클레아제 (예를 들어, ZFN, TALEN, CRISPR/Cas 뉴클레아제)는 조작된 뉴클레아제 도메인 또는 절반-도메인, 예를 들어 절단적인 이중-이량체를 형성하는 조작된 *FokI* 절단 절반 도메인을 포함한다. 예를 들어, 미국 특허 공개 번호 20080131962 참조.

[0032] 또 다른 측면에서, 본 개시내용은 본원에 기재된 단백질, 융합 분자 및/또는 그의 성분 (예를 들어, sgRNA 또는 다른 DNA-결합 도메인) 중 어느 것을 코딩하는 폴리뉴클레오티드를 제공한다. 이러한 폴리뉴클레오티드는 바이러스 벡터, 비-바이러스 벡터 (예를 들어, 플라스미드)의 일부일 수 있거나 또는 mRNA 형태일 수 있다. 본원에 기재된 폴리뉴클레오티드 중 어느 것은 또한, B2M, TCR α 및/또는 TCR β 유전자 내로의 표적화된 삽입을 위한 서열 (공여자, 상동성 아암 또는 패치 서열)을 포함할 수 있다. 또한 또 다른 측면에서, 본원에 기재된 폴리뉴클레오티드 중 어느 것을 포함하는 유전자 전달 벡터가 제공된다. 특정 실시양태에서, 이러한 벡터는 아데노바이러스 벡터 (예를 들어, Ad5/F35 벡터) 또는 렌티바이러스 벡터 (LV)이며, 이는 통합 적격하거나 또는 통합-결합있는 렌티바이러스 벡터 또는 아데노 관련 벡터 (AAV)를 포함한다. 따라서, 또한 본원에는 뉴클레아제 (예를 들어 ZFN 또는 TALEN) 및/또는 뉴클레아제 시스템 (CRISPR/Cas 또는 Ttgo)을 코딩하는 서열 및/또는 표적 유전자 내로의 표적화된 통합을 위한 공여자 서열을 포함하는 바이러스 벡터가 제공된다. 일부 실시양태에서, 상기 공여자 서열 및 뉴클레아제를 코딩하는 상기 서열은 상이한 벡터 상에 있다. 다른 실시양태에서, 뉴클레아제는 폴리캡티드로서 공급된다. 바람직한 실시양태에서, 폴리뉴클레오티드는 mRNA이다. 일부 측면에서, mRNA는 화학적으로 변형될 수 있다 (예를 들어, 문헌 [Kormann et al., (2011) *Nature Biotechnology* 29(2):154-157] 참조). 다른 측면에서, mRNA는 ARCA 캡을 포함할 수 있다 (미국 특허 7,074,596 및 8,153,773 참조). 일부 측면에서, mRNA는 효소적 변형에 의해 도입된 캡을 포함할 수 있다. 효소적으로 변형된 캡은 캡0, 캡1 또는 캡2를 포함할 수 있다 (예를 들어, 문헌 [Smietanski et al., (2014) *Nature Communications* 5:3004] 참조). 추가 측면에서, mRNA는 화학적 변형에 의해 캡핑될 수 있다. 추가 실시양태에서, mRNA는 비변형된 뉴클레오티드와 변형된 뉴클레오티드의 혼합물을 포함할 수 있다 (미국 특허 공개 2012-0195936 참조). 또한 추가 실시양태에서, mRNA는 WPRE 요소를 포함할 수 있다 (미국 특허 출원 15/141,333 참조). 일부 실시양태에서, mRNA는 이중 가닥이다 (예를 들어, 문헌 [Kariko et al. (2011) *Nucl Acid Res* 39:e142] 참조).

[0033] 또한 또 다른 측면에서, 본 개시내용은 본원에 기재된 단백질, 폴리뉴클레오티드 및/또는 벡터 중 어느 것을 포함하는 단리된 세포를 제공한다. 특정 실시양태에서, 이러한 세포는 줄기/전구 세포, 또는 T-세포 (예를 들어,

CD4⁺ T-세포)로 이루어진 군으로부터 선택된다. 또한 추가 측면에서, 본 개시내용은 본원에 기재된 단백질, 폴리뉴클레오티드 및/또는 벡터 중 어느 것을 포함하는 세포 또는 세포주로부터 유래되는 세포 또는 세포주, 즉 B2M이 하나 이상의 ZFN에 의해 불활성화되었고/거나 공여자 폴리뉴클레오티드 (예를 들어 ACTR, 조작된 TCR 및/또는 CAR)가 세포의 게놈 내로 안정적으로 통합된 세포로부터 유래된 (예를 들어, 배양 중인) 세포 또는 세포주를 제공한다. 따라서, 본원에 기재된 바와 같은 세포의 자손은 그 자체가 본원에 기재된 단백질, 폴리뉴클레오티드 및/또는 벡터를 포함할 수 없지만, 이들 세포에서는, 적어도 B2M 유전자가 불활성화되고/거나 공여자 폴리뉴클레오티드가 게놈 내로 통합되고/거나 발현된다.

[0034] 또 다른 측면에서, 본원에는 하나 이상의 단백질, 폴리뉴클레오티드 및/또는 벡터를 본원에 기재된 바와 같은 세포 내로 도입함으로써 특정 세포 내의 B2M 유전자를 불활성화시키는 방법이 기재된다. 본원에 기재된 방법 중 어느 것에서 뉴클레아제는 표적화된 돌연변이 유발, 세포성 DNA 서열의 결실을 유도시킬 수 있고/거나 미리 결정된 염색체 로커스에서 표적화된 재조합을 촉진시킬 수 있다. 따라서, 특정 실시양태에서, 상기 뉴클레아제는 하나 이상의 뉴클레오티드를 표적 유전자로부터 결실시키거나 또는 표적 유전자 내로 삽입시킨다. 일부 실시양태에서 B2M 유전자는 뉴클레아제 절단에 이어 비-상동 말단 결합에 의해 불활성화된다. 다른 실시양태에서, 표적 유전자 내의 게놈 서열은, 예를 들어 본원에 기재된 바와 같은 뉴클레아제 (또는 이러한 뉴클레아제를 코딩하는 벡터) 및 표적화된 절단 후에 상기 유전자 내로 삽입되는 "공여자" 서열을 사용하여 상기 뉴클레아제로 대체시킨다. 공여자 서열은 뉴클레아제 벡터에 존재하거나, 별도의 벡터 (예를 들어, AAV, Ad 또는 LV 벡터)에 존재하거나, 또는 대안적으로, 상이한 핵산 전달 메커니즘을 사용하여 세포 내로 도입될 수 있다.

[0035] 더욱이, 본원에 기재된 방법 중 어느 것은 시험관내, 생체내 및/또는 생체외에서 실시될 수 있다. 특정 실시양태에서, 상기 방법은, 예를 들어 T-세포를 변형시켜 이것이 대상체 (예를 들어, 암에 걸린 대상체)를 치료하기 위한 동종이체 환경에서 치료제로서 유용하도록 만들기 위해 생체외에서 실시된다. 치료 및/또는 예방될 수 있는 암의 비제한적 예는 폐 암종, 췌장암, 간암, 골암, 유방암, 결장직장암, 백혈병, 난소암, 림프종, 뇌암 등을 포함한다.

도면의 간단한 설명

[0036] 도 1은 뉴클레아제에 의해 표적화된 B2M 유전자의 엑손 1 및 엑손 2의 서열 (서열식별번호: 1 및 2)을 도시한 것이다. 박스는 ZFN 결합 (표적) 부위에 의해 플랭킹되는 5개의 상이한 절단 영역 (A (GGCCTTA), C (TCAAATT), D (TCAAAT), E (TTACTGA) 및 G (AATTGAA))을 표시한다.

도 2a 및 도 2b는 뉴클레아제 활성을 도시한다. 도 2a는 2 또는 6 μ g의 용량에서 도 1에 제시된 바와 같은 B2M 부위 A, C, D, E 및 G에 대해 특이적인 ZFN으로 처리된 T 세포 내의 각각의 부위에서 유전자 변형의 퍼센트를 도시하는 막대 그래프이다. 도 2b는 K562 세포에서 B2M 유전자에 대한 TALEN 활성을 도시한 것이다.

도 3은 FACS 분석에 의해 분석된 바와 같이 B2M-특이적 ZFN 쌍으로의 처리 후 HLA 음성 T 세포의 퍼센트를 도시한 것이다.

도 4a 내지 4e는 세포를 B2M-특이적 ZFN 및 TCRA-특이적 ZFN 둘 다로 처리한 것으로부터의 FACS 결과를 도시한 것이다. 도 4a는 ZFN 처리 없는 경우의 결과를 도시한 것이고, 도 4b는 TCRA-특이적 ZFN 단독 후의 결과를 도시한 것이며 (CD3 시그널의 96% 녹아웃), 도 4c는 B2M-특이적 ZFN 단독 후의 결과를 도시한 것이다 (HLA 시그널의 92% 녹아웃). 도 4d는 이중 녹아웃을 가진 (이로써 HLA 마킹과 CD3 마킹 둘 다 손실됨) 세포의 위치를 보여주는 예시도이다. 도 4e는 세포를 TCRA-특이적 ZFN 및 CD3-특이적 ZFN 둘 다로 처리한 후의 결과를 도시하며, 이는 세포의 82%에서 이중 녹아웃이 발생하였다는 것을 입증해 준다.

도 5는 TRAC (TCRA) 및 B2M 이중 녹아웃 및 TRAC (TCRA) 또는 B2M 로커스 중 하나 내로의 공여자의 표적화된 통합의 결과를 도시한 것이다.

발명을 실시하기 위한 구체적인 내용

[0037] 본원에는 세포가 그의 세포 표면 상에 HLA 부류 I을 더 이상 포함하지 않도록 B2M 유전자의 발현을 조정하는, 상기 세포를 생성하기 위한 조성물 및 방법이 개시된다. 이러한 방식으로 변형된 세포는 치료제, 예를 들어, 이식 조직으로서 사용될 수 있으며, 이는 B2M 발현의 결여가 HLA 기반 면역 반응을 방지 또는 저하시키기 때문이다. 부가적으로, 다른 관심 유전자가, B2M 유전자를 조작시킨 세포 내로 삽입될 수 있고/거나 다른 관심 유

전자가 녹아아웃될 수 있다.

[0038] **일반사항**

[0039] 본원에 개시된 방법의 실시 뿐만 아니라 조성물의 제조 및 용도는 달리 표시되지 않는 한, 분자 생물학, 생화학, 염색질 구조 및 분석, 전산 화학, 세포 배양, 재조합 DNA 및 관련 기술분야의 기술 내에 있는 바와 같은 관련된 분야에서의 통상적인 기술을 이용한다. 이들 기술은 하기 문헌에서 완전히 설명된다. 예를 들어, 문헌 [Sambrook et al. MOLECULAR CLONING: A LABORATORY MANUAL, Second edition, Cold Spring Harbor Laboratory Press, 1989 and Third edition, 2001; Ausubel et al., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, New York, 1987 and periodic updates; the series 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. Wassarman and A. P. Wolffe, eds.), Academic Press, San Diego, 1999; and METHODS IN MOLECULAR BIOLOGY, Vol. 119, "Chromatin Protocols" (P.B. Becker, ed.) Humana Press, Totowa, 1999] 참조.

[0040] **정의**

[0041] 용어 "핵산", "폴리뉴클레오티드" 및 "올리고뉴클레오티드"는 상호 교환적으로 사용되고, 선형 또는 환상 입체 형태, 및 단일 가닥 또는 이중 가닥 형태의 데옥시리보뉴클레오티드 또는 리보뉴클레오티드 중합체를 지칭한다. 본 개시내용의 목적상, 이들 용어는 중합체의 길이에 대해 제한하는 것으로 해석되서는 안된다. 상기 용어는 자연 뉴클레오티드 뿐만 아니라 염기, 당 및/또는 포스페이트 모이어티 (예를 들어, 포스포로티오에이트 주쇄)에 있어서 변형되는 뉴클레오티드의 공지된 유사체를 포괄할 수 있다. 일반적으로, 특정한 뉴클레오티드의 유사체는 동일한 염기-쌍형성 특이성을 가지며; 즉 A의 유사체는 T와 염기-쌍형성할 것이다.

[0042] 용어 "폴리펩티드", "펩티드" 및 "단백질"은 아미노산 잔기의 중합체를 지칭하기 위해 상호 교환적으로 사용된다. 상기 용어는 하나 이상의 아미노산이 상응하는 자연적으로 발생하는 아미노산의 화학적 유사체 또는 변형된 유도체인 아미노산 중합체에도 적용된다.

[0043] "결합"은 거대분자 간의 (예를 들어, 단백질과 핵산 간의) 서열-특이적, 비-공유적 상호작용을 지칭한다. 결합 상호작용의 모든 성분이 반드시 서열-특이적 (예를 들어, DNA 주쇄 내의 포스페이트 잔기와 접촉)일 필요는 없지만, 이러한 상호작용은 총괄적으로 서열-특이적이어야 한다. 이러한 상호작용은 일반적으로, 10^{-6} M^{-1} 이하의 해리 상수 (K_d)로써 특징지을 수 있다. "친화도"는 결합 강도를 지칭하는데: 증가된 결합 친화도는 더 낮은 K_d 와 상관이 있다. "비-특이적 결합"은 임의의 관심 분자 (예를 들어, 조작된 뉴클레아제)와, 표적 서열에 의존적이지 않는 거대분자 (예를 들어 DNA) 간에 일어나는 비-공유적 상호작용을 지칭한다.

[0044] "DNA 결합 분자"는 DNA에 결합할 수 있는 분자이다. 이러한 DNA 결합 분자는 폴리펩티드, 단백질의 도메인, 더 큰 단백질 내의 도메인 또는 폴리뉴클레오티드일 수 있다. 일부 실시양태에서, 폴리뉴클레오티드는 DNA인 반면, 다른 실시양태에서는, 폴리뉴클레오티드가 RNA이다. 일부 실시양태에서, DNA 결합 분자는 뉴클레아제의 단백질 도메인 (예를 들어, FokI 도메인)인 반면, 다른 실시양태에서는, DNA 결합 분자가 RNA-가이드된 뉴클레아제 (예를 들어 Cas9 또는 Cfp1)의 가이드 RNA 성분이다.

[0045] "결합 단백질"은 또 다른 분자에 비-공유적으로 결합할 수 있는 단백질이다. 결합 단백질은, 예를 들어, DNA 분자 (DNA-결합 단백질), RNA 분자 (RNA-결합 단백질) 및/또는 단백질 분자 (단백질-결합 단백질)에 결합할 수 있다. 단백질-결합 단백질의 경우, 이는 그 자체에 결합할 수 있고/거나 (동종이량체, 동종삼량체 등을 형성하기 위함) 상이한 단백질(들)의 하나 이상의 분자에 결합할 수 있다. 결합 단백질은 2가지 이상 유형의 결합 활성을 가질 수 있다. 예를 들어, 징크 핑거 단백질은 DNA-결합, RNA-결합 및 단백질-결합 활성을 갖는다.

[0046] "징크 핑거 DNA 결합 단백질" (또는 결합 도메인)은 그의 구조가 아연 이온의 협조를 통하여 안정화되는 결합 도메인 내의 아미노산 서열의 영역인, 하나 이상의 징크 핑거를 통하여 서열-특이적 방식으로 DNA에 결합하는 단백질, 또는 더 큰 단백질 내의 도메인이다. 용어 징크 핑거 DNA 결합 단백질은 종종, 징크 핑거 단백질 또는 ZFP로서 약칭된다.

[0047] "TALE DNA 결합 도메인" 또는 "TALE"는 하나 이상의 TALE 반복 도메인/단위를 포함하는 폴리펩티드이다. 각각 반복 가변 이잔기 (RVD)를 포함하는 반복 도메인은 TALE의 그의 동족 표적 DNA 서열에의 결합에 관여한다. 단일 "반복 단위" ("반복"으로서 지칭되기도 함)는 전형적으로, 그 길이가 33-35개 아미노산이고 자연적으로 발생하는 TALE 단백질 내의 다른 TALE 반복 서열과 적어도 일부 서열 상동성을 나타낸다. TALE 단백질은 반복 단위

내의 정준 또는 비-정준 RVD를 사용하여 표적 부위에 결합하도록 디자인될 수 있다. 예를 들어, 미국 특허 번호 8,586,526 및 9,458,205 참조, 그 전문이 본원에 참조로 포함된다.

[0048] 징크 핑거 및 TALE DNA-결합 도메인은, 예를 들어 자연적으로 발생하는 징크 핑거 단백질의 인식 나선 영역의 조작 (하나 이상의 아미노산의 변경)을 통하여 또는 DNA 결합에 관여하는 아미노산 (반복 가변 이잔기 또는 RVD 영역)의 조작에 의해, 미리 결정된 뉴클레오타이드 서열에 결합하도록 "조작될" 수 있다. 따라서, 조작된 징크 핑거 단백질 또는 TALE 단백질은 비-자연적으로 발생하는 단백질이다. 징크 핑거 단백질 및 TALE를 조작하는 방법의 비-제한적 예는 디자인 및 선택이다. 디자인된 단백질은 자연적으로 발생하지 않는 단백질로서 그의 디자인/조성이 주로 합리적인 기준으로부터 비롯된다. 합리적인 디자인 기준은 기존의 ZFP 또는 TALE 디자인 (정준 및 비-정준 RVD) 및 결합 데이터에 관한 정보를 저장하는 데이터베이스에서 정보를 처리하기 위한 대체 규칙 및 컴퓨터 알고리즘의 적용을 포함한다. 예를 들어, 미국 특허 번호 9,458,205; 8,586,526; 6,140,081; 6,453,242; 및 6,534,261 참조; 또한 WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 및 WO 03/016496 참조.

[0049] "선택된" 징크 핑거 단백질, TALE 단백질 또는 CRISPR/Cas 시스템은 자연에서 발견되지 않고, 그의 생성이 주로, 실험 과정, 예컨대 파지 디스플레이, 상호작용 트랩 또는 하이브리드 선택으로부터 비롯된다. 예를 들어, U.S. 5,789,538; U.S. 5,925,523; U.S. 6,007,988; U.S. 6,013,453; U.S. 6,200,759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197 및 WO 02/099084 참조.

[0050] "TtAgo"는 유전자 침묵에 관여하는 것으로 여겨지는 원핵 아르코노트 단백질이다. TtAgo는 박테리아 썬무스 썬 모필루스(*Thermus thermophilus*)로부터 유래된다. 예를 들어, 문헌 [Swarts et al., 상기 문헌, 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 포함한, 요구되는 모든 성분이다.

[0051] "재조합"은 두 폴리뉴클레오타이드 간의 유전적 정보의 교환 프로세스를 지칭한다. 본 개시내용의 목적상, "상동 재조합 (HR)"은, 예를 들어, 상동성 유도 복구 메커니즘을 통하여 세포 내에서 이중 가닥 파단의 복구 동안 일어나는 전문화된 형태의 상기 교환을 지칭한다. 이러한 프로세스는 뉴클레오타이드 서열 상동성을 필요로 하고, "표적" 분자 (즉, 이중 가닥 파단을 경험한 것)의 주형 복구에 대한 "공여자" 분자를 이용하며, "비-교차 유전자 전환" 또는 "짧은 트랙 유전자 전환"으로서 다양하게 공지되어 있으며, 이는 유전적 정보를 공여자에서 표적으로 전달해 주기 때문이다. 어떠한 특정한 이론에도 얽매하고자 한 것은 아니지만, 이러한 전달은 파괴된 표적과 공여자 간에 형성되는 이중 이중 나선 DNA의 불일치 교정, 및/또는 공여자가 표적의 일부가 될 유전적 정보를 재합성하는데 사용되는 "합성-의존적 가닥 어닐링", 및/또는 관련 프로세스를 포함할 수 있다. 이러한 전문화된 HR은 종종, 표적 분자의 서열의 변경을 초래하여, 공여자 폴리뉴클레오타이드의 서열의 일부 또는 전부가 표적 폴리뉴클레오타이드 내로 혼입되도록 한다.

[0052] 본 개시내용의 방법에서, 본원에 기재된 바와 같은 하나 이상의 표적화된 뉴클레아제는 미리 결정된 부위 (예를 들어, 관심 유전자 또는 로커스)에서 표적 서열 (예를 들어, 세포성 염색질) 내에 이중 가닥 파단 (DSB)을 창출시키고, 이러한 파단의 영역 내의 뉴클레오타이드 서열에 대한 상동성을 갖는 "공여자" 폴리뉴클레오타이드가 세포 내로 도입될 수 있다. DSB의 존재가 공여자 서열의 통합을 촉진시키는 것으로 밝혀진 바 있다. 임의로, 상기 구조물은 파단의 영역 내의 뉴클레오타이드 서열에 대한 상동성을 갖는다. 공여자 서열은 물리적으로 통합될 수 있거나, 또는 대안적으로, 공여자 폴리뉴클레오타이드는 상동 재조합을 통하여 파단의 복구를 위한 주형으로서 사용되어, 공여자에서와 같은 뉴클레오타이드 서열의 전부 또는 일부가 세포성 염색질 내로 도입된다. 따라서, 세포성 염색질 내의 제1 서열이 변경될 수 있고, 특정 실시양태에서는, 공여자 폴리뉴클레오타이드에 존재하는 서열로 전환될 수 있다. 따라서, 용어 "대체하는" 또는 "대체"의 사용은 하나의 뉴클레오타이드 서열을 또 다른 뉴클레오타이드 서열로 대체 (즉, 정보를 제공하는 의미에서 특정 서열의 대체)하는 것을 나타내는 것으로 이해될 수 있고, 하나의 폴리뉴클레오타이드를 또 다른 폴리뉴클레오타이드로 물리적 또는 화학적으로 대체하는 것을 반드시 필요로 하진 않는다.

[0053] 본원에 기재된 방법 중 어느 것에서, 부가 쌍의 징크 핑거 단백질이 상기 세포 내에서의 부가 표적 부위의 부가 이중 가닥 절단을 위해 사용될 수 있다.

[0054] 세포성 염색질 내의 관심 영역에서의 특정 서열의 표적화된 재조합 및/또는 대체 및/또는 변경을 위한 방법의 특정 실시양태에서, 염색체 서열은 외인성 "공여자" 뉴클레오타이드 서열과의 상동 재조합에 의해 변경된다. 이러한 상동 재조합은 이중 가닥 파단의 영역과 상동인 서열이 존재하는 경우에, 세포성 염색질에서의 상기 이중 가닥 파단의 존재에 의해 자극된다.

- [0055] 본원에 기재된 방법 중 어느 것에서, 제1 뉴클레오타이드 서열 ("공여자 서열")은 관심 영역 내의 게놈 서열과 상동이지만, 동일하지 않음으로써, 관심 영역 내에 동일하지 않는 서열을 삽입하기 위한 상동 재조합을 자극하는 서열을 함유할 수 있다. 따라서, 특정 실시양태에서, 관심 영역 내의 서열과 상동인 공여자 서열의 부분은, 대체되는 게놈 서열과 약 80 내지 99% (또는 그 사이의 임의의 정수) 서열 동일성을 나타낸다. 다른 실시양태에서, 공여자와 게놈 서열 간의 상동성은, 예를 들어, 100개의 연속되는 염기 쌍 전반에 걸쳐 공여자와 게놈 서열 간에 1개의 뉴클레오타이드만이 상이한 경우에는, 99% 보다 더 높다. 특정의 경우에, 공여자 서열의 비-상동 부분은 관심 영역에 존재하지 않는 서열을 함유할 수 있으므로, 새로운 서열이 관심 영역 내로 도입된다. 이러한 경우, 비-상동 서열은 일반적으로, 관심 영역 내의 서열과 상동이거나 또는 동일한 50-1,000개 염기 쌍 (또는 그 사이의 임의의 정수 값) 또는 1,000개 초과 임의의 수의 염기 쌍의 서열에 의해 플랭킹된다. 다른 실시양태에서, 공여자 서열은 제1 서열과 비-상동이고, 비-상동 재조합 메커니즘에 의해 게놈 내로 삽입된다.
- [0056] 본원에 기재된 방법 중 어느 것은 관심 유전자(들)의 발현을 붕괴시키는 공여자 서열의 표적화된 통합에 의해 특정 세포에서 하나 이상의 표적 서열의 부분적 또는 완전한 불활성화를 위해 사용될 수 있다. 부분적으로 또는 완전하게 불활성화된 유전자를 수반한 세포주가 또한 제공된다.
- [0057] 더욱이, 본원에 기재된 바와 같은 표적화된 통합 방법은 또한, 하나 이상의 외인성 서열을 통합시키기 위해 사용될 수 있다. 외인성 핵산 서열은, 예를 들어, 하나 이상의 유전자 또는 cDNA 분자, 또는 임의의 유형의 코딩 또는 비코딩 서열 뿐만 아니라 하나 이상의 제어 요소 (예를 들어, 프로모터)를 포함할 수 있다. 또한, 외인성 핵산 서열은 하나 이상의 RNA 분자 (예를 들어, 작은 헤어핀 RNA (shRNA), 억제성 RNA (RNAi), 마이크로RNA (miRNA) 등)를 생성시킬 수 있다.
- [0058] "절단"은 DNA 분자의 공유 주쇄의 파손을 지칭한다. 절단은 포스포디에스테르 결합의 효소적 또는 화학적 가수분해를 포함하지만, 이에 제한되지 않는 각종 방법에 의해 개시될 수 있다. 단일 가닥 절단과 이중 가닥 절단 둘 다 가능하고, 이중 가닥 절단은 2개의 별개의 단일 가닥 절단 이벤트에 따른 결과로서 발생할 수 있다. DNA 절단은 평활 말단 또는 엇갈린 말단 중 하나의 생성을 초래할 수 있다. 특정 실시양태에서, 융합 폴리펩티드가 표적화된 이중 가닥 DNA 절단을 위해 사용된다.
- [0059] "절단 절반-도메인"은 제2 폴리펩티드 (동일하거나 상이함)와 연계해서, 절단 활성 (바람직하게 이중 가닥 절단 활성)을 갖는 복합체를 형성하는 폴리펩티드 서열이다. 용어 "제1 및 제2 절단 절반-도메인"; "+ 및 - 절단 절반-도메인" 및 "오른쪽 및 왼쪽 절단 절반-도메인"은 이량체화되는 절단 절반-도메인의 쌍을 지칭하기 위해 상호 교환적으로 사용된다.
- [0060] "조작된 절단 절반-도메인"은 또 다른 절단 절반-도메인 (예를 들어, 또 다른 조작된 절단 절반-도메인)과 절대적인 이중-이량체를 형성하도록 변형시켰던 절단 절반-도메인이다. 또한, 미국 특허 번호 7,888,121; 7,914,796; 8,034,598; 8,623,618 및 미국 특허 공개 번호 2011/0201055 참조, 그 전문이 본원에 참조로 포함된다.
- [0061] 용어 "서열"은 DNA 또는 RNA일 수 있고; 선형, 환상 또는 분지될 수 있으며 단일 가닥 또는 이중 가닥일 수 있는, 임의의 길이의 뉴클레오타이드 서열을 지칭한다. 용어 "공여자 서열"은 게놈 내로 삽입되는 뉴클레오타이드 서열을 지칭한다. 공여자 서열은 임의의 길이, 예를 들어 2 내지 10,000개 뉴클레오타이드 길이 (또는 그 사이 또는 그 위의 임의의 정수 값), 바람직하게 약 100 내지 1,000개 뉴클레오타이드 길이 (또는 그 사이의 임의의 정수), 보다 바람직하게 약 200 내지 500개 뉴클레오타이드 길이의 것일 수 있다.
- [0062] "염색질"은 세포성 게놈을 포함하는 핵단백질 구조이다. 세포성 염색질은 핵산, 주로 DNA, 및 단백질 (히스톤 및 비-히스톤 염색체 단백질 포함)을 포함한다. 대다수의 진핵 세포성 염색질은 뉴클레오솜의 형태로 존재하며, 여기서 뉴클레오솜 코어는 히스톤 H2A, H2B, H3 및 H4 각각 2개를 포함하는 8량체와 관련된 대략 150개 염기 쌍의 DNA를 포함하고; 링커 DNA (유기체에 따라서 가변적인 길이를 가짐)가 뉴클레오솜 코어 사이에서 연장된다. 히스톤 H1의 분자는 일반적으로, 링커 DNA와 연합된다. 본 개시내용의 목적상, 용어 "염색질"은 원핵과 진핵 둘 다의 모든 유형의 세포성 핵단백질을 포괄하는 것을 의미한다. 세포성 염색질은 염색체성과 에피솜성 염색질 둘 다를 포함한다.
- [0063] "염색체"는 세포의 게놈의 전부 또는 특정 부분을 포함하는 염색질 복합체이다. 세포의 게놈은 종종, 이러한 세포의 게놈을 포함하는 모든 염색체의 컬렉션인 그의 핵형으로써 특징지을 수 있다. 세포의 게놈은 하나 이상의 염색체를 포함할 수 있다.
- [0064] "에피솜"은 복제성 핵산, 핵단백질 복합체, 또는 세포의 염색체 핵형의 일부가 아닌 핵산을 포함하는 다른 구조

이다. 에피솜의 예는 플라스미드 및 특정의 바이러스 게놈을 포함한다.

[0065] "표적 부위" 또는 "표적 서열"은 결합을 위한 충분한 조건이 존재한다면, 결합 분자에 결합할 핵산의 특정 부분을 규정하는 핵산 서열이다. 예를 들어, 서열 5' GAATTC 3'는 Eco RI 제한 엔도뉴클레아제에 대한 표적 부위이다.

[0066] "외인성" 분자는 정상적으로는 세포에 존재하지 않지만, 하나 이상의 유전적, 생화학적 또는 다른 방법에 의해 세포 내로 도입될 수 있는 분자이다. "세포에서의 정상적인 존재"는 세포의 특정한 발달 단계 및 환경 조건과 관련하여 결정된다. 따라서, 예를 들어, 근육의 배아 발달 동안에만 존재하는 분자는 성체 근육 세포에 대한 외인성 분자이다. 유사하게, 열 충격에 의해 유도된 분자는 비-열 충격받은 세포에 대한 외인성 분자이다. 외인성 분자는, 예를 들어, 오작동하는 내인성 분자의 기능적 버전 또는 정상적으로 기능하는 내인성 분자의 오작동 버전을 포함할 수 있다.

[0067] 외인성 분자는 다른 것들 중에서도, 조합 화학 프로세스에 의해 생성되는 것과 같은 작은 분자, 또는 거대분자, 예컨대 단백질, 핵산, 탄수화물, 지질, 당단백질, 지단백질, 폴리사카라이드, 상기 분자의 임의의 변형된 유도체, 또는 상기 분자 중 하나 이상을 포함하는 임의의 복합체일 수 있다. 핵산은 DNA 및 RNA를 포함하고, 단일 가닥 또는 또는 이중 가닥일 수 있으며; 선형, 분지된 또는 환상일 수 있고; 임의의 길이일 수 있다. 예를 들어, 미국 특허 번호 8,703,489 및 9,255,259 참조. 핵산은 이중 나선을 형성할 수 있는 것 뿐만 아니라 삼중체 형성 핵산을 포함한다. 예를 들어, 미국 특허 번호 5,176,996 및 5,422,251 참조. 단백질은 DNA-결합 단백질, 전사 인자, 염색질 재형성 인자, 메틸화 DNA 결합 단백질, 폴리머라제, 메틸라제, 데메틸라제, 아세틸라제, 데아세틸라제, 키나제, 포스파타제, 인테그라제, 리컴비나제, 라이가제, 토포이소머라제, 자이라제 및 헬리카제를 포함하지만, 이에 제한되지 않는다.

[0068] 외인성 분자는 내인성 분자와 동일한 유형의 분자, 예를 들어, 외인성 단백질 또는 핵산일 수 있다. 예를 들어, 외인성 핵산은 감염성 바이러스 게놈, 세포 내로 도입된 플라스미드 또는 에피솜, 또는 세포 내에 정상적으로 존재하지 않는 염색체를 포함할 수 있다. 외인성 분자를 세포 내로 도입하는 방법은 관련 기술분야의 통상의 기술자에게 공지되어 있고, 지질 매개된 전이 (즉, 중성 및 양이온성 지질을 포함한 리포솜), 전기천공, 직접적인 주사, 세포 융합, 입자 충격, 인산칼슘 공동 침전, DEAE-텍스트란 매개된 전이 및 바이러스 벡터-매개된 전이를 포함하지만, 이에 제한되지 않는다. 외인성 분자는 또한, 내인성 분자와 동일한 유형의 분자일 수 있지만, 세포가 유래되는 것과 상이한 종으로부터 유래될 수 있다. 예를 들어, 인간 핵산 서열은 원래 마우스 또는 햄스터로부터 유래된 세포주 내로 도입될 수 있다.

[0069] 반대로, "내인성" 분자는 특정한 환경 조건 하에서 특정한 발달 단계 하의 특정한 세포에 정상적으로 존재하는 분자이다. 예를 들어, 내인성 핵산은 염색체, 미토콘드리아의 게놈, 엽록체 또는 다른 세포 소기관, 또는 자연적으로 발생하는 에피솜 핵산을 포함할 수 있다. 부가의 내인성 분자는 단백질, 예를 들어, 전사 인자 및 효소를 포함할 수 있다.

[0070] "융합" 분자는 2개 이상의 서브유닛 분자가, 바람직하게 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 간의 융합) 및 융합 핵산 (예를 들어, 상기 기재된 융합 단백질을 코딩하는 핵산)을 포함하지만, 이에 제한되지 않는다. 제2 유형의 융합 분자의 예는 삼중체 형성 핵산과 폴리펩티드 간의 융합, 및 작은 홈 결합체와 핵산 간의 융합을 포함하지만, 이에 제한되지 않는다. 상기 용어는 또한, 폴리뉴클레오티드 성분이 폴리펩티드 성분과 연합하여 기능적 분자를 형성하는 시스템 (예를 들어, 단일 가이드 RNA가 기능적 도메인과 연합하여 유전자 발현을 조정하는 CRISPR/Cas 시스템)을 포함한다.

[0071] 특정 세포 내에서의 융합 단백질의 발현은 이러한 세포로의 융합 단백질의 전달로부터 비롯되거나 또는 융합 단백질을 코딩하는 폴리뉴클레오티드를 세포로 전달함으로써 비롯될 수 있으며, 여기서 상기 폴리뉴클레오티드가 전사되고, 전사체가 번역되어 융합 단백질이 생성된다. 트랜스-스플라이싱, 폴리펩티드 절단 및 폴리펩티드 라이게이션이 또한, 특정 세포에서의 단백질의 발현에 관여할 수 있다. 세포에 대한 폴리뉴클레오티드 및 폴리펩티드 전달 방법은 본 개시내용의 다른 곳에 제시된다.

[0072] "유전자"는 본 개시내용의 목적상, 유전자 생성물 (하기 참조)을 코딩하는 DNA 영역 뿐만 아니라 조절성 서열이 코딩 및/또는 전사된 서열에 인접하는 지의 여부와 상관없이, 유전자 생성물의 생성을 조절하는 모든 DNA 영역을 포함한다. 따라서, 유전자는 프로모터 서열, 종결인자, 번역 조절성 서열, 예컨대 리보솜 결합 부위 및 내

부 리보솜 진입 부위, 증강인자, 침묵인자, 절연체, 경계 요소, 복제 기점, 매트릭스 부착 부위 및 로커스 제어 영역을 포함하지만, 이에 반드시 제한되지 않는다.

- [0073] "안전한 정착" 로커스는 숙주 세포에 대한 어떠한 해로운 효과도 없이 유전자가 삽입될 수 있는 게놈 내의 로커스이다. 삽입된 유전자 서열의 발현이 이웃하는 유전자로부터의 임의의 연속 관독에 의해 동요되지 않는 안전한 정착 로커스가 가장 유익하다. 뉴클레아제(들)에 의해 표적화되는 안전한 정착 로커스의 비-제한적 예는 CCR5, CCR5, HPRT, AAVS1, *Rosa* 및 알부민을 포함한다. 예를 들어, 미국 특허 번호 8,771,985; 8,110,379; 7,951,925; 미국 공개 번호 20100218264; 20110265198; 20130137104; 20130122591; 20130177983; 20130177960; 20150056705 및 20150159172 참조.
- [0074] "유전자 발현"은 유전자 내에 함유된 정보를 유전자 생성물로 전환시키는 것을 지칭한다. 유전자 생성물은 유전자의 직접 전사 생성물 (예를 들어, mRNA, tRNA, rRNA, 안티센스 RNA, 리보자임, 구조적 RNA 또는 임의의 다른 유형의 RNA) 또는 mRNA의 번역에 의해 생성된 단백질일 수 있다. 유전자 생성물은 또한, 캡핑, 폴리아데닐화, 메틸화 및 편집과 같은 프로세스에 의해 변형되는 RNA, 및 예를 들어, 메틸화, 아세틸화, 인산화, 유비퀴틴화, ADP-리보실화, 미리스틸화, 및 글리코실화에 의해 변형된 단백질을 포함한다.
- [0075] 유전자 발현의 "조정" 또는 "변형"은 유전자의 활성 상의 변화를 지칭한다. 발현의 조정은 외인성 분자 (예를 들어, 조작된 전사 인자)의 결합을 통한 유전자의 변형에 의한 것을 포함한, 유전자 활성화 및 유전자 억제를 포함할 수 있지만, 이에 제한되지 않는다. 조정은 또한, 게놈 편집 (예를 들어, 절단, 변경, 불활성화, 무작위 돌연변이)를 통한 유전자 서열의 변형에 의해 달성될 수 있다. 유전자 불활성화는 본원에 기재된 바와 같이 변형되지 않았던 세포와 비교해서 유전자 발현 상의 임의의 저하를 지칭한다. 따라서, 유전자 불활성화는 부분적이거나 또는 완전할 수 있다.
- [0076] "관심 영역"은 세포성 염색질의 임의의 영역, 예컨대, 예를 들어, 외인성 분자에 결합하는 것이 바람직한 특정 유전자, 또는 유전자 내의 또는 이와 인접한 비-코딩 서열을 지칭한다. 결합은 표적화된 DNA 절단 및/또는 표적화된 재조합의 목적을 위한 것일 수 있다. 관심 영역은, 예를 들어 염색체, 예피솜, 세포 소기관 게놈 (예를 들어, 미토콘드리아, 엽록체), 또는 감염 바이러스 게놈에 존재할 수 있다. 관심 영역은 유전자의 코딩 영역 내에, 전사된 비-코딩 영역, 예컨대, 예를 들어, 리더 서열, 트레일러 서열 또는 인트론 내에, 또는 코딩 영역의 상류 또는 하류의 비-전사된 영역 내에 존재할 수 있다. 관심 영역은 단일 뉴클레오타이드 쌍 정도로 작을 수 있거나 또는 2,000개 이하의 뉴클레오타이드 쌍 길이, 또는 임의의 정수 값의 뉴클레오타이드 쌍일 수 있다.
- [0077] "진핵" 세포는 진균 세포 (예컨대 효모), 식물 세포, 동물 세포, 포유류 세포 및 인간 세포 (예를 들어, T-세포)를 포함하지만, 이에 제한되지 않는다.
- [0078] 용어 "작동적 연결" 및 "작동적으로 연결된" (또는 "작동적으로 연결된")은 2가지 이상의 성분 (예컨대, 서열 요소)의 병렬과 관련하여 상호 교환적으로 사용되며, 여기서는 양쪽 성분이 정상적으로 기능하고 이러한 성분 중 적어도 하나가, 다른 성분 중 적어도 하나 상에서 발휘되는 기능을 매개할 수 있는 가능성을 허용하도록 상기 성분들이 배열된다. 한 예시로서, 전사 조절성 서열이 하나 이상의 전사 조절성 인자의 존재 또는 부재에 반응하여 코딩 서열의 전사 수준을 제어하는 경우에는, 이러한 전사 조절성 서열, 예컨대 프로모터가 상기 코딩 서열과 작동적으로 연결된다. 전사 조절성 서열은 일반적으로, 코딩 서열과 시스로 작동적으로 연결되지만, 이에 직접적으로 인접할 필요는 없다. 예를 들어, 증강인자는 코딩 서열과 비록 연속되지 않더라도 이러한 코딩 서열과 작동적으로 연결되는 전사 조절성 서열이다.
- [0079] 융합 폴리펩티드와 관련하여, 용어 "작동적으로 연결된"이란 각각의 성분이 그렇게 연결되어 있지 않은 것처럼 다른 성분과 연동하여 동일한 기능을 수행한다는 사실을 지칭할 수 있다. 예를 들어, DNA-결합 도메인 (예를 들어, ZFP, TALE)이 활성화 도메인과 융합되는 융합 폴리펩티드와 관련해서는, 이러한 융합 폴리펩티드 내에서 DNA-결합 도메인 부분이 그의 표적 부위 및/또는 그의 결합 부위에 결합할 수 있는 반면, 활성화 도메인은 유전자 발현을 상향 조절할 수 있는 경우에, 상기 DNA-결합 도메인과 활성화 도메인은 작동적으로 연결되어 있다. DNA-결합 도메인이 절단 도메인과 융합되는 융합 폴리펩티드와 관련해서는, 이러한 융합 폴리펩티드 내에서 DNA-결합 도메인 부분이 그의 표적 부위 및/또는 그의 결합 부위에 결합할 수 있는 반면, 절단 도메인은 표적 부위의 부근에서 DNA를 절단할 수 있는 경우에, 상기 DNA-결합 도메인과 절단 도메인은 작동적으로 연결되어 있다. 유사하게, DNA-결합 도메인이 활성화 또는 억제 도메인과 융합되는 융합 폴리펩티드와 관련해서는, 이러한 융합 폴리펩티드 내에서 DNA-결합 도메인 부분이 그의 표적 부위 및/또는 그의 결합 부위에 결합할 수 있는 반면, 활성화 도메인은 유전자 발현을 상향 조절할 수 있거나 또는 억제 도메인은 유전자 발현을 하향 조절할 수 있는 경우에, 상기 DNA-결합 도메인과 활성화 또는 억제 도메인은 작동적으로 연결되어 있다.

- [0080] 단백질, 폴리펩티드 또는 핵산의 "기능적 단편"은 그의 서열이 완전한 길이의 단백질, 폴리펩티드 또는 핵산과 동일하지는 않지만, 완전한 길이의 단백질, 폴리펩티드 또는 핵산과 동일한 기능을 보유하고 있는 단백질, 폴리펩티드 또는 핵산이다. 기능적 단편은 상응하는 천연 분자보다 더 많거나, 더 적거나, 또는 천연 분자와 동일한 수의 잔기를 보유할 수 있고/거나 하나 이상의 아미노산 또는 뉴클레오티드 치환을 함유할 수 있다. 핵산의 기능 (예를 들어, 또 다른 핵산과 혼성화할 수 있는 능력인 코딩 기능)을 결정하는 방법은 관련 기술분야에 널리 공지되어 있다. 유사하게, 단백질 기능을 결정하는 방법은 널리 공지되어 있다. 예를 들어, 폴리펩티드의 DNA-결합 기능은, 예를 들어, 필터-결합, 전기영동 이동-시프트, 또는 면역침전 검정에 의해 결정될 수 있다. DNA 절단은 겔 전기영동에 의해 검정될 수 있다. 상기 문헌 [Ausubel et al.] 참조. 또 다른 단백질과 상호작용할 수 있는 특정 단백질의 능력은, 예를 들어, 공동 면역침전, 2-하이브리드 검정 또는 상보성 (유전적 및 생화학적 둘 다)에 의해 결정될 수 있다. 예를 들어, 문헌 [Fields et al. (1989) *Nature* 340:245-246]; 미국 특허 번호 5,585,245 및 PCT WO 98/44350 참조.
- [0081] "벡터"는 유전자 서열을 표적 세포로 전이시킬 수 있다. 전형적으로, "벡터 구조물", "발현 벡터", "발현 구조물", "발현 카세트" 및 "유전자 전이 벡터"는 관심 유전자의 발현을 지시할 수 있고 유전자 서열을 표적 세포로 전이시킬 수 있는 임의의 핵산 구조물을 의미한다. 따라서, 상기 용어는 클로닝, 및 발현 비히클 뿐만 아니라 통합 벡터를 포함한다.
- [0082] "리포터 유전자" 또는 "리포터 서열"은 반드시 아니지만 바람직하게 일상적인 검정에서, 용이하게 측정되는 단백질 생성물을 생성시키는 임의의 서열을 지칭한다. 적합한 리포터 유전자는 항생제 내성 (예를 들어, 앰피실린 내성, 네오마이신 내성, G418 내성, 푸로마이신 내성)을 매개하는 단백질을 코딩하는 서열, 착색되거나 또는 형광성 또는 발광성 단백질 (예를 들어, 그린 형광성 단백질, 증강된 그린 형광성 단백질, 레드 형광성 단백질, 루시페라제)을 코딩하는 서열, 및 증강된 세포 성장 및/또는 유전자 증폭을 매개하는 단백질 (예를 들어, 디히드로폴레이트 리덕타제)을 코딩하는 서열을 포함하지만, 이에 제한되지 않는다. 에피토프 태그는, 예를 들어, FLAG, His, myc, Tap, HA 또는 임의의 검출가능한 아미노산 서열의 하나 이상의 카피를 포함한다. "발현 태그"는 관심 유전자의 발현을 모니터링하기 위해 목적하는 유전자 서열과 작동적으로 연결될 수 있는 리포터를 코딩하는 서열을 포함한다.
- [0083] 용어 "대상체" 및 "환자"는 상호 교환적으로 사용되고, 포유류, 예컨대 인간 환자 및 비-인간 영장류 뿐만 아니라 실험 동물, 예컨대 토끼, 개, 고양이, 래트, 마우스, 및 다른 동물을 지칭한다. 따라서, 본원에 사용된 바와 같은 용어 "대상체" 및 "환자"는 본 발명의 발현 카세트가 투여될 수 있는 임의의 포유류 환자 또는 대상체를 의미한다. 본 발명의 대상체는 특정 장애가 있는 대상체 또는 특정 장애가 발생할 위험이 있는 대상체를 포함한다.
- [0084] 본원에 사용된 바와 같은 용어 "치료하는" 및 "치료"는 증상의 증증도 및/또는 빈도 상의 감소, 증상 및/또는 근원적인 원인의 제거, 증상 및/또는 그의 근원적인 원인의 발생의 방지, 및 손상의 호전 또는 개선을 지칭한다. 암 및 이식편 대 숙주 질환은 본원에 기재된 조성물 및 방법을 이용하여 치료될 수 있는 병태의 비-제한적 예이다. 따라서, "치료하는" 및 "치료"는 하기를 포함한다:
- [0085] (i) 특히, 특정 포유류가 특정 병태에 걸리기는 쉽지만 아직 이러한 병태에 걸린 것으로 진단되지 않는 경우에, 이러한 포유류에게서 질환 또는 병태가 발생하지 못하게 함;
- [0086] (ii) 질환 또는 병태를 억제하거나, 즉 그의 발생을 저지함;
- [0087] (iii) 질환 또는 병태를 경감시키거나, 즉 질환 또는 병태의 퇴행을 유발시킴; 또는
- [0088] (iv) 질환 또는 병태로부터 비롯되는 증상을 경감시키는 것, 즉 근원적인 질환 또는 병태를 해결하지 않고서도 통증을 경감시킴.
- [0089] 본원에 사용된 바와 같은 용어 "질환" 및 "병태"는 상호 교환적으로 사용될 수 있거나, 또는 (병인이 아직 해결되지 않았기 때문에) 특정한 병 또는 병태가 공지된 원인 병원체를 가지지 않을 수 있어서 아직까지 질환으로서 인정되지는 않지만 바람직하지 않은 병태 또는 증후군으로서만 인정될 수 있다는 점 (여기서, 다소 구체적인 증상 세트가 임상적에 의해 확인되었음)에서 상이할 수 있다.
- [0090] "제약 조성물"은 본 발명의 화합물과, 생물학적으로 활성인 화합물을 포유류, 예를 들어, 인간에게 전달하기 위하여 관련 기술분야에서 일반적으로 허용된 매질의 제형을 지칭한다. 이러한 매질은 그에 대한 모든 제약상 허용되는 담체, 희석제 또는 부형제를 포함한다.

- [0091] "유효량" 또는 "치료상 유효량"은 포유류, 바람직하게 인간에게 투여될 때, 이러한 포유류, 바람직하게 인간에게서 치료를 가져오기에 충분한, 본 발명의 화합물의 그 양을 지칭한다. "치료상 유효량"을 구성하는 본 발명의 조성물의 양은 상기 화합물, 병태 및 그의 중증도, 투여 방식, 및 치료하고자 하는 포유류의 연령에 따라 다양할 것이지만, 관련 기술분야의 통상의 기술자 자신의 지식과 본 개시내용과 관련하여 상기 기술자에 의해 일상적으로 결정될 수 있다.
- [0092] **DNA-결합 도메인**
- [0093] 본원에는 HLA 유전자 또는 HLA 조절인자를 포함하는 임의의 유전자 (B2M 유전자 포함) 내의 표적 부위에 특이적으로 결합하는 DNA-결합 도메인을 포함하는 조성물이 기재된다. 임의의 DNA-결합 도메인이 본원에 개시된 조성물 및 방법에 사용될 수 있으며, 이는 징크 핑거 DNA-결합 도메인, TALE DNA 결합 도메인, CRISPR/Cas 뉴클레아제의 DNA-결합 부분 (sgRNA), 또는 메가뉴클레아제로부터의 DNA-결합 도메인을 포함하지만, 이에 제한되지 않는다. DNA-결합 도메인은 유전자 내의 임의의 표적 서열에 결합할 수 있으며, 이러한 서열은 서열식별번호: 6-48 중 어느 것에 제시된 바와 같은 12개 이상의 뉴클레오타이드의 표적 서열을 포함하지만, 이에 제한되지 않는다.
- [0094] 특정 실시양태에서, DNA 결합 도메인은 징크 핑거 단백질을 포함한다. 바람직하게, 징크 핑거 단백질은 그것이 선택되는 표적 부위에 결합하도록 조작된다는 점에서 비-자연적으로 발생한다. 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 번호 6,453,242; 6,534,261; 6,599,692; 6,503,717; 6,689,558; 7,030,215; 6,794,136; 7,067,317; 7,262,054; 7,070,934; 7,361,635; 7,253,273; 및 미국 특허 공개 번호 2005/0064474; 2007/0218528; 2005/0267061 참조; 모두 그 전문이 본원에 참조로 포함된다. 특정 실시양태에서, DNA-결합 도메인은 미국 특허 공개 번호 2012/0060230 (예를 들어, 표 1) (그 전문이 본원에 참조로 포함됨)에 개시된 징크 핑거 단백질을 포함한다.
- [0095] 조작된 징크 핑거 결합 도메인은 자연적으로 발생하는 징크 핑거 단백질과 비교해서 신규의 결합 특이성을 가질 수 있다. 조작 방법은 합리적 디자인 및 다양한 유형의 선택을 포함하지만, 이에 제한되지 않는다. 합리적 디자인은, 예를 들어, 삼중자 (또는 사중자) 뉴클레오타이드 서열 및 개별 징크 핑거 아미노산 서열을 포함하는 데이터베이스를 사용하는 것을 포함하며, 여기서 각각의 삼중자 또는 사중자 뉴클레오타이드 서열은 특정한 삼중자 또는 사중자 서열에 결합하는 징크 핑거의 하나 이상의 아미노산 서열과 연관된다. 예를 들어, 미국 특허 6,453,242 및 6,534,261 참조, 그 전문이 본원에 참조로 포함된다.
- [0096] 파지 디스플레이 및 2-하이브리드 시스템을 포함한, 예시적인 선택 방법이 미국 특허 5,789,538; 5,925,523; 6,007,988; 6,013,453; 6,410,248; 6,140,466; 6,200,759; 및 6,242,568; 뿐만 아니라 WO 98/37186; WO 98/53057; WO 00/27878; WO 01/88197 및 GB 2,338,237에 개시된다. 또한, 징크 핑거 결합 도메인에 대한 결합 특이성의 증강은, 예를 들어, 미국 특허 번호 6,794,136에 기재되어 있다.
- [0097] 또한, 이들 및 다른 참고문헌에 개시된 바와 같이, 징크 핑거 도메인 및/또는 다중 핑거된 징크 핑거 단백질은 임의의 적합한 링커 서열, 예를 들어, 5개 이상의 아미노산 길이의 링커를 사용하여 함께 연결시킬 수 있다. 또한, 6개 이상의 아미노산 길이의 예시적인 링커 서열에 관해서는 미국 특허 번호 6,479,626; 6,903,185; 및 7,153,949를 참조할 수 있다. 본원에 기재된 단백질은 이러한 단백질의 개별 징크 핑거 간의 적합한 링커의 임의의 조합을 포함할 수 있다. 또한, 징크 핑거 결합 도메인에 대한 결합 특이성의 증강이, 예를 들어, 미국 특허 번호 6,794,136에 기재되어 있다.
- [0098] 표적 부위의 선택; ZFP, 및 융합 단백질 (및 이를 코딩하는 폴리뉴클레오타이드)의 디자인 및 구축 방법은 관련 기술분야의 통상의 기술자에게 공지되어 있고, 미국 특허 번호 6,140,081; 5,789,538; 6,453,242; 6,534,261; 5,925,523; 6,007,988; 6,013,453; 6,200,759; WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197; WO 02/099084; WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536 및 WO 03/016496에 상세히 기재되어 있다.
- [0099] 또한, 이들 및 다른 참고문헌에 개시된 바와 같이, 징크 핑거 도메인 및/또는 다중 핑거된 징크 핑거 단백질은 임의의 적합한 링커 서열, 예를 들어, 5개 이상의 아미노산 길이의 링커를 사용하여 함께 연결시킬 수 있다. 또한, 6개 이상의 아미노산 길이의 예시적인 링커 서열에 관해서는 미국 특허 번호 6,479,626; 6,903,185; 및 7,153,949를 참조할 수 있다. 본원에 기재된 단백질은 이러한 단백질의 개별 징크 핑거 간의 적합한 링커의 임의의 조합을 포함할 수 있다.

- [0100] 특정 실시양태에서, DNA 결합 도메인은 B2M 유전자 또는 B2M 조절성 유전자 내의 표적 부위와 (서열-특이적 방식으로) 결합하고 B2M 유전자의 발현을 조정하는 조작된 징크 핑거 단백질이다. 일부 실시양태에서, 징크 핑거 단백질은 B2M 내의 표적 부위에 결합하는 반면, 다른 실시양태에서는, 징크 핑거가 B2M 내의 표적 부위에 결합한다.
- [0101] 통상적으로, ZFP는 적어도 3개의 핑거를 포함한다. 특정의 ZFP는 4개, 5개 또는 6개의 핑거를 포함한다. 3개의 핑거를 포함하는 ZFP는 전형적으로, 9개 또는 10개의 뉴클레오티드를 포함하는 표적 부위를 인식하고; 4개의 핑거를 포함하는 ZFP는 전형적으로, 12 내지 14개의 뉴클레오티드를 포함하는 표적 부위를 인식하는 반면; 6개의 핑거를 갖는 ZFP는 18 내지 21개의 뉴클레오티드를 포함하는 표적 부위를 인식할 수 있다. ZFP는 또한, 하나 이상의 조절성 도메인을 포함하는 융합 단백질일 수 있으며, 이러한 도메인은 전사 활성화 또는 억제 도메인일 수 있다.
- [0102] 일부 실시양태에서, DNA-결합 도메인은 뉴클레아제로부터 유래될 수 있다. 예를 들어, 귀소 엔도뉴클레아제 및 메가뉴클레아제, 예컨대 I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-TevI, I-TevII 및 I-TevIII의 인식 서열은 공지되어 있다. 또한, 미국 특허 번호 5,420,032; 미국 특허 번호 6,833,252; 문헌 [Belfort et al. (1997) *Nucleic Acids Res.* 25:3379-3388; Dujon et al. (1989) *Gene* 82:115-118; Perler et al. (1994) *Nucleic Acids Res.* 22, 1125-1127; Jasin (1996) *Trends Genet.* 12:224-228; Gimble et al. (1996) *J. Mol. Biol.* 263:163-180; Argast et al. (1998) *J. Mol. Biol.* 280:345-353 and the New England Biolabs catalogue] 참조. 또한, 귀소 엔도뉴클레아제 및 메가뉴클레아제의 DNA-결합 특이성은 비-자연 표적 부위에 결합하도록 조작될 수 있다. 예를 들어, 문헌 [Chevalier et al. (2002) *Molec. Cell* 10:895-905; Epinat et al. (2003) *Nucleic Acids Res.* 31:2952-2962; Ashworth et al. (2006) *Nature* 441:656-659; Paques et al. (2007) *Current Gene Therapy* 7:49-66]; 미국 특허 공개 번호 20070117128 참조.
- [0103] 다른 실시양태에서, DNA 결합 도메인은 식물 병원체 크산토모나스(*Xanthomonas*) (문헌 [Boch et al., (2009) *Science* 326: 1509-1512 and Moscou and Bogdanove, (2009) *Science* 326: 1501] 참조) 및 랄스토니아(*Ralstonia*) (문헌 [Heuer et al. (2007) *Applied and Environmental Microbiology* 73(13): 4379-4384]; 미국 특허 출원 번호 20110301073 및 20110145940 참조)로부터 유래된 것과 유사한 TAL 이펙터로부터 조작된 도메인을 포함한다. 크산토모나스 속의 식물 병원성 박테리아는 중요한 농작물에서 많은 질병을 유발시키는 것으로 공지되어 있다. 크산토모나스의 병원성은 식물 세포 내로 25종 초과 상이한 이펙터 단백질을 주입하는 보존된 유형 III 분비 (T3S) 시스템에 좌우된다. 이들 중에서 주입된 단백질은 식물 전사 활성화를 모방하고 식물 전사체를 조작하는 전사 활성화제-유사 이펙터 (TALE)이다 (문헌 [Kay et al. (2007) *Science* 318:648-651] 참조). 이들 단백질은 DNA 결합 도메인 및 전사 활성화 도메인을 함유한다. 가장 널리 특징규명된 TALE 중 하나가 크산토모나스 캄페스트그리스 pv. 베시카토리아(*Xanthomonas campestris* pv. *Vesicatoria*)로부터의 AvrBs3이다. 문헌 [Bonas et al. (1989) *Mol Gen Genet* 218: 127-136] 및 W02010079430 참조. TALE는 탠덤 반복 서열의 중앙 집중화된 도메인을 포함하며, 각각의 반복 서열은 이들 단백질의 DNA 결합 특이성의 핵심인 대략 34개의 아미노산을 함유한다. 또한, 이들은 핵 국제화 서열과 산성 전사 활성화 도메인을 함유한다 (고찰을 위해, 문헌 [Schornack S, et al. (2006) *J Plant Physiol* 163(3): 256-272] 참조). 또한, 식물 병원성 박테리아 랄스토니아 솔라나세아룸(*Ralstonia solanacearum*)에서는, brg11 및 hpx17로 지명된 2개의 유전자가 알. 솔라나세아룸(*R. solanacearum*) 생물형 변이체 1 균주 GMI1000 및 생물형 변이체 4 균주 RS1000에서 크산토모나스의 AvrBs3 패밀리에 상동인 것으로 밝혀졌다 (문헌 [Heuer et al. (2007) *Appl and Envir Micro* 73(13): 4379-4384] 참조). 이들 유전자는 뉴클레오티드 서열에 있어서 서로 98.9% 동일하지만, hpx17의 반복 도메인에 있어서 1,575 bp의 결실 정도 상이하다. 그러나, 양쪽 유전자 생성물은 크산토모나스의 AvrBs3 패밀리 단백질과 40% 미만의 서열 동일성을 갖는다.
- [0104] 이들 TAL 이펙터의 특이성은 탠덤 반복 서열에서 발견된 서열에 좌우된다. 이러한 반복된 서열은 대략 102개의 염기 쌍을 포함하고, 반복 서열은 전형적으로, 서로 91-100% 상동이다 (Bonas et al., 상기 문헌). 반복 서열의 다형성은 통상적으로, 위치 12 및 13에 위치하고, 위치 12 및 13에서의 초가변 이란기 (반복 가변 이란기 또는 RVD 영역)의 동일성과 TAL-이펙터의 표적 서열 내의 연속되는 뉴클레오티드의 동일성 간에는 1 대 1 상응도가 있는 것으로 여겨진다 (문헌 [Moscou and Bogdanove, (2009) *Science* 326:1501 and Boch et al. (2009) *Science* 326:1509-1512] 참조). 실험적으로, 이들 TAL-이펙터의 DNA 인식에 대한 자연 코드는, 위치 12 및 13에서의 HD 서열 (반복 가변 이란기 또는 RVD)이 시토신 (C)에 대한 결합을 초래하고, NG가 T에 결합하며, NI가 A, C, G 또는 T에 결합하고, NN이 A 또는 G에 결합하며, ING가 T에 결합하도록 결정되어 있다. 이들 DNA 결합

반복 서열은 새로운 조합과 수의 반복 서열을 수반한 단백질로 어셈블리되어, 새로운 서열과 상호작용할 수 있고 식물 세포에서 비-내인성 리포터 유전자의 발현을 활성화시킬 수 있는 인공 전사 인자를 만들었다 (Boch et al., 상기 문헌). 조작된 TAL 단백질을 *FokI* 절단 절반 도메인과 연결시켜 TAL 이펙터 도메인 뉴클레아제 융합 (TALEN) (비정형 RVD를 수반한 TALEN 포함)을 산출시켰다. 예를 들어, 미국 특허 번호 8,586,526 참조.

[0105] 일부 실시양태에서, TALEN은 엔도뉴클레아제 (예를 들어, *FokI*) 절단 도메인 또는 절단 절반-도메인을 포함한다. 다른 실시양태에서, TALE-뉴클레아제는 메가 TAL이다. 이들 메가 TAL 뉴클레아제는 TALE DNA 결합 도메인과 메가뉴클레아제 절단 도메인을 포함하는 융합 단백질이다. 메가뉴클레아제 절단 도메인은 단량체로서 활성이고, 활성을 위해 이량체화를 필요로 하지 않는다. (문헌 [Boissel et al., (2013) *Nucl Acid Res*: 1-13, doi: 10.1093/nar/gkt1224] 참조).

[0106] 또한 추가 실시양태에서, 뉴클레아제는 치밀한 TALEN을 포함한다. 이들은 TALE DNA 결합 도메인을 *TevI* 뉴클레아제 도메인과 연결시키는 단일 채 융합 단백질이다. 이러한 융합 단백질은 TALE 영역에 의해 국제된 니카제로서 작용할 수 있거나, 또는 TALE DNA 결합 도메인이 *TevI* 뉴클레아제 도메인에 대해서 위치되는 곳에 따라서 이중 가닥 파단을 창출시킬 수 있다 (문헌 [Beurdeley et al. (2013) *Nat Comm*: 1-8 DOI: 10.1038/ncomms2782] 참조). 또한, 뉴클레아제 도메인은 또한, DNA-결합 기능성을 나타낼 수 있다. 임의의 TALEN이 하나 이상의 메가-TALE를 수반한 부가의 TALEN (예를 들어, 하나 이상의 TALEN (cTALEN 또는 *FokI*-TALEN))과 조합하여 사용될 수 있다.

[0107] 또한, 이들 및 다른 참고문헌에 개시된 바와 같이, 징크 핑거 도메인 및/또는 다중 핑거된 징크 핑거 단백질 또는 TALE는 임의의 적합한 링커 서열, 예를 들어, 5개 이상의 아미노산 길이의 링커를 사용하여 함께 연결시킬 수 있다. 또한, 6개 이상의 아미노산 길이의 예시적인 링커 서열에 관해서는 미국 특허 번호 6,479,626; 6,903,185; 및 7,153,949를 참조할 수 있다. 본원에 기재된 단백질은 이러한 단백질의 개별 징크 핑거 간의 적합한 링커의 임의의 조합을 포함할 수 있다. 또한, 징크 핑거 결합 도메인에 대한 결합 특이성의 증강이, 예를 들어, 미국 특허 번호 6,794,136에 기재되어 있다.

[0108] 특정 실시양태에서, DNA-결합 도메인은 CRISPR/Cas 뉴클레아제 시스템의 일부이며, 이는 DNA에 결합하는 단일 가이드 RNA (sgRNA)를 포함한다. 예를 들어, 미국 특허 번호 8,697,359 및 미국 특허 공개 번호 20150056705 및 20150159172 참조. 시스템의 RNA 성분을 코딩하는 CRISPR (클러스터된 규칙적으로 공간을 둔 짧은 팔린드롬성 반복 서열) 로커스, 및 단백질을 코딩하는 cas (CRISPR-관련) 로커스 (문헌 [Jansen et al., 2002. *Mol. Microbiol.* 43: 1565-1575; Makarova et al., 2002. *Nucleic Acids Res.* 30: 482-496; Makarova et al., 2006. *Biol. Direct* 1: 7; Haft et al., 2005. *PLoS Comput. Biol.* 1: e60])가 CRISPR/Cas 뉴클레아제 시스템의 유전자 서열을 구성한다. 미생물 숙주 내에서의 CRISPR 로커스는 CRISPR-관련 (Cas) 유전자 뿐만 아니라 CRISPR-매개된 핵산 절단의 특이성을 프로그래밍할 수 있는 비-코딩 RNA 요소의 조합을 함유한다.

[0109] 유형 II CRISPR은 가장 널리 특징규명된 시스템 중 하나이고, 4가지 일련의 단계에서 표적화된 DNA 이중 가닥 파단을 수행한다. 첫째, 2개의 비-코딩 RNA인 프리-crRNA 어레이 및 tracrRNA가 상기 CRISPR 로커스로부터 전사된다. 둘째, tracrRNA는 프리-crRNA의 반복 영역과 혼성화되고, 프리-crRNA가 개별 스페이서 서열을 함유하는 성숙한 crRNA로 프로세싱되는 것을 매개한다. 셋째, 성숙한 crRNA:tracrRNA 복합체는 crRNA 상의 스페이서와, 표적 인식을 위해 부가적으로 요구되는 프로토스페이서 인접 모티프 (PAM) 다음에 있는 표적 DNA 상의 프로토스페이서 간의 왓슨-크릭(Watson-Crick) 염기-쌍형성을 통하여 기능적 도메인 (예를 들어, 뉴클레아제, 예컨대 Cas)이 표적 DNA를 향하게 한다. 최종적으로, Cas9는 표적 DNA의 절단을 매개하여 프로토스페이서 내에 이중 가닥 파단을 창출시킨다. CRISPR/Cas 시스템의 활성은 3가지 단계를 포함한다: (i) '적응'으로 지칭되는 프로세스에서, 향후 공격을 방지하기 위해 CRISPR 어레이 내로의 외계 DNA 서열의 삽입, (ii) 관련 단백질의 발현 뿐만 아니라 상기 어레이의 발현 및 프로세싱에 이어, (iii) 상기 외계 핵산으로의 RNA-매개된 간섭. 따라서, 박테리아 세포에서는, 소위 'Cas' 단백질 중 몇 가지가 CRISPR/Cas 시스템의 자연적 기능과 관련이 있고, 외계 DNA의 삽입 등과 같은 기능에 있어서 일정 역할을 한다.

[0110] 특정 실시양태에서, Cas 단백질은 자연적으로 발생하는 Cas 단백질의 "기능적 유도체"일 수 있다. 천연 서열 폴리펩티드의 "기능적 유도체"는 천연 서열 폴리펩티드와 공통으로 정성적 생물학적 특성을 갖는 화합물이다. "기능적 유도체"는 천연 서열의 단편, 및 천연 서열 폴리펩티드 및 그의 단편의 유도체를 포함하지만, 이에 제한되지 않으며, 단 이들은 상응하는 천연 서열 폴리펩티드와 공통으로 생물학적 활성을 가져야 한다. 본원에서 고려된 생물학적 활성은 DNA 기질을 단편으로 가수분해시킬 수 있는 기능적 유도체의 능력이다. 용어 "유도체"는 폴리펩티드의 아미노산 서열 변이체, 공유적 변형물, 및 그의 융합물, 예컨대 유도체 Cas 단백질 모두를

포괄한다. Cas 폴리펩티드 또는 그의 단편의 적합한 유도체는 Cas 단백질 또는 그의 단편의 돌연변이체, 융합물, 공유적 변형물을 포함하지만, 이에 제한되지 않는다. Cas 단백질 또는 그의 단편을 포함하는 Cas 단백질 뿐만 아니라 Cas 단백질 또는 그의 단편의 유도체는 세포로부터 수득가능할 수 있거나 또는 화학적으로 합성될 수 있거나 또는 이들 두 절차의 조합에 의해 수득가능할 수 있다. 상기 세포는 Cas 단백질을 자연적으로 생성하는 세포, 또는 Cas 단백질을 자연적으로 생성하고 내인성 Cas 단백질을 더 높은 발현 수준에서 생성하도록 유전적으로 조작되거나 또는 외인적으로 도입된 핵산으로부터 Cas 단백질을 생성하도록 유전적으로 조작되는 세포일 수 있으며, 핵산은 내인성 Cas와 동일하거나 또는 상이한 Cas를 코딩한다. 일부 경우에, 상기 세포는 Cas 단백질을 자연적으로 생성하지 않고 Cas 단백질을 생성하도록 유전적으로 조작된다. 일부 실시양태에서, Cas 단백질은 AAV 벡터를 통하여 전달하기 위한 작은 Cas9 오르토로그이다 (Ran et al. (2015) *Nature* 510, p. 186).

[0111] 일부 실시양태에서, DNA 결합 도메인은 TtAgo 시스템의 일부이다 (문헌 [Swarts et al., 상기 문헌; Sheng et al., 상기 문헌] 참조). 진핵 생물에서는, 유전자 침묵이 아르코노트 (Ago) 패밀리의 단백질에 의해 매개된다. 이러한 패러다임에서는, Ago가 작은 (19-31개 nt) RNA에 결합된다. 이러한 단백질-RNA 침묵 복합체는 작은 RNA와 표적 간의 왓슨-크릭 염기 쌍형성을 통하여 표적 RNA를 인식하고, 표적 RNA를 엔도뉴클레오 분해적으로 절단한다 (Vogel (2014) *Science* 344:972-973). 반대로, 원핵 Ago 단백질은 작은 단일 가닥 DNA 단편에 결합하고, 외래 (종종 바이러스) DNA를 검출 및 제거하는 기능을 하는 것으로 예상된다 (Yuan et al., (2005) *Mol. Cell* 19, 405; Olovnikov, et al. (2013) *Mol. Cell* 51, 594; Swarts et al., 상기 문헌). 예시적인 원핵 Ago 단백질은 아키펙스 아에올리쿠스 (*Aquifex aeolicus*), 로도박터 스파에로이데스 (*Rhodobacter sphaeroides*), 및 씨무스 씨모필루스로부터의 것을 포함한다.

[0112] 가장 널리 특징규명된 원핵 Ago 단백질 중 하나는 티. 씨모필루스로부터의 것이다 (TtAgo; Swarts et al. 상기 문헌). TtAgo는 5' 포스페이트 기를 수반한 15개 nt 또는 13-25개 nt 단일 가닥 DNA 단편과 연합된다. 이와 같이 TtAgo에 의해 결합된 "가이드 DNA"는 단백질-DNA 복합체가 제3자 분자의 DNA에서 왓슨-크릭 상보적 DNA 서열에 결합하도록 지시하는 작용을 한다. 일단 이들 가이드 DNA에서의 서열 정보가 표적 DNA의 확인을 허용하면, TtAgo-가이드 DNA 복합체는 표적 DNA를 절단한다. 이러한 메카니즘은 또한, 그의 표적 DNA에 결합되는 동안 TtAgo-가이드 DNA 복합체의 구조에 의해 뒷받침된다 (G. Sheng et al., 상기 문헌). 로도박터 스파에로이데스로부터의 Ago (RsAgo)가 유사한 특성을 갖고 있다 (Olovnikov et al. 상기 문헌).

[0113] 임의의 DNA 서열의 외인성 가이드 DNA가 TtAgo 단백질 상으로 부하될 수 있다 (Swarts et al. 상기 문헌). TtAgo 절단의 특이성이 가이드 DNA에 의해 지시되기 때문에, 연구자에 의해 명시된 외인성 가이드 DNA와 함께 형성된 TtAgo-DNA 복합체는 TtAgo 표적 DNA 절단이 연구자에 의해 명시된 상보적 표적 DNA를 향하도록 할 것이다. 이러한 방식으로, DNA에서 표적화된 이중 가닥 파단을 창출시킬 수 있다. TtAgo-가이드 DNA 시스템 (또는 다른 유기체로부터의 오르토로그의 Ago-가이드 DNA 시스템)을 사용하면, 세포 내에서의 게놈 DNA의 표적화된 절단이 허용된다. 이러한 절단은 단일 가닥 또는 이중 가닥일 수 있다. 포유류 게놈 DNA의 절단을 위해서는, 포유류 세포에서의 발현을 위해 최적화된 TtAgo 코돈의 특정 버전을 사용하는 것이 바람직할 것이다. 추가로, 세포를 시험관 내에서 형성된 TtAgo-DNA 복합체로 처리하는 것이 바람직할 수도 있으며, 여기서는 TtAgo 단백질이 세포 침투성 펩티드와 융합된다. 추가로, 37°C에서 개선된 활성을 갖도록 돌연변이 유발을 통하여 변경시킨 TtAgo 단백질의 특정 버전을 사용하는 것이 바람직할 수도 있다. Ago-RNA-매개된 DNA 절단을 사용하여, DNA 파단의 이용을 위하여 관련 기술분야에서의 표준 기술을 사용하는 유전자 녹아웃, 표적화된 유전자 부가, 유전자 교정, 표적화된 유전자 결실을 포함한, 성과의 파노폴리(panopoly)에 영향을 미칠 수 있었다.

[0114] 따라서, 임의의 DNA-결합 도메인을 사용할 수 있다

[0115] 융합 분자

[0116] 이중 조절성 (기능적) 도메인 (또는 그의 기능적 단편)과 연합된 본원에 기재된 바와 같은 DNA-결합 도메인 (예를 들어, ZFP 또는 TALE, CRISPR/Cas 성분, 예컨대 단일 가이드 RNA)을 포함하는 융합 분자가 또한 제공된다. 공통의 도메인은, 예를 들어, 전사 인자 도메인 (활성화제, 억제인자, 공동 활성화제, 공동 억제인자), 침묵인자, 종양 유전자 (예를 들어, myc, jun, fos, myb, max, mad, rel, ets, bcl, myb, mos 패밀리 구성원 등); DNA 복구 효소 및 그의 관련 인자 및 변형인자; DNA 재배열 효소 및 그의 관련 인자 및 변형인자; 염색질 관련 단백질 및 그의 변형인자 (예를 들어, 키나제, 아세틸라제 및 데아세틸라제); 및 DNA 변형 효소 (예를 들어, 메틸트랜스퍼라제, 토포이소머라제, 헬리카제, 라이가제, 키나제, 포스파타제, 폴리머라제, 엔도뉴클레아제) 및 그의 관련 인자 및 변형인자를 포함한다. 이러한 융합 분자는 본원에 기재된 DNA-결합 도메인 및 전사 조절성

도메인을 포함하는 전사 인자 뿐만 아니라 DNA-결합 도메인 및 하나 이상의 뉴클레아제 도메인을 포함하는 뉴클레아제를 포함한다.

- [0117] 활성화를 달성하는데 적합한 도메인 (전사 활성화 도메인)은 HSV VP16 활성화 도메인 (예를 들어, 문헌 [Hagmann et al., *J. Virol.* 71, 5952-5962 (1997)] 참조) 핵 호르몬 수용체 (예를 들어, 문헌 [Torchia et al., *Curr. Opin. Cell. Biol.* 10:373-383 (1998)] 참조); 핵 인자 카파 B의 p65 서브유닛 (문헌 [Bitko & Barik, *J. Virol.* 72:5610-5618 (1998) and Doyle & Hunt, *Neuroreport* 8:2937-2942 (1997); Liu et al., *Cancer Gene Ther.* 5:3-28 (1998)]), 또는 인공 키메라 기능적 도메인, 예컨대 VP64 (문헌 [Beerli et al., (1998) *Proc. Natl. Acad. Sci. USA* 95:14623-33]), 및 데그론(degron) (문헌 [Molinari et al., (1999) *EMBO J.* 18, 6439-6447])을 포함한다. 부가의 예시적인 활성화 도메인은 Oct 1, Oct-2A, Sp1, AP-2, 및 CTF1 (문헌 [Seipel et al., *EMBO J.* 11, 4961-4968 (1992)]) 뿐만 아니라 p300, CBP, PCAF, SRC1 PVALF, AtHD2A 및 ERF-2를 포함한다. 예를 들어, 문헌 [Robyr et al. (2000) *Mol. Endocrinol.* 14:329-347; Collingwood et al. (1999) *J. Mol. Endocrinol.* 23:255-275; Leo et al. (2000) *Gene* 245:1-11; Manteuffel-Cymborowska (1999) *Acta Biochim. Pol.* 46:77-89; McKenna et al. (1999) *J. Steroid Biochem. Mol. Biol.* 69:3-12; Malik et al. (2000) *Trends Biochem. Sci.* 25:277-283; and Lemon et al. (1999) *Curr. Opin. Genet. Dev.* 9:499-504] 참조. 부가의 예시적인 활성화 도메인은 OsGAI, HALF-1, C1, AP1, ARF-5, -6, -7, 및 -8, CPRF1, CPRF4, MYC-RP/GP, 및 TRAB1을 포함하지만, 이에 제한되지 않는다. 예를 들어, 문헌 [Ogawa et al. (2000) *Gene* 245:21-29; Okanami et al. (1996) *Genes Cells* 1:87-99; Goff et al. (1991) *Genes Dev.* 5:298-309; Cho et al. (1999) *Plant Mol. Biol.* 40:419-429; Ulmason et al. (1999) *Proc. Natl. Acad. Sci. USA* 96:5844-5849; Sprenger-Haussels et al. (2000) *Plant J.* 22:1-8; Gong et al. (1999) *Plant Mol. Biol.* 41:33-44; and Hobo et al. (1999) *Proc. Natl. Acad. Sci. USA* 96:15,348-15,353] 참조.
- [0118] DNA-결합 도메인과 기능적 도메인 간에 융합 단백질 (또는 이를 코딩하는 핵산)을 형성하는데 있어서, 활성화 도메인이나 또는 활성화 도메인과 상호작용하는 분자가 기능적 도메인으로서 적합하다는 것이 관련 기술분야의 통상의 기술자에게 명백할 것이다. 활성화 복합체 및/또는 활성화 활성 (예컨대, 예를 들어, 히스톤 아세틸화)을 표적 유전자로 동원할 수 있는 본질적으로 어떠한 분자도 융합 단백질의 활성화 도메인으로서 유용하다. 융합 분자에서 기능적 도메인으로서 사용하기 적합한 절연체 도메인, 국제화 도메인, 및 염색질 재형성 단백질, 예컨대 ISWI-함유 도메인 및/또는 메틸 결합 도메인 단백질이, 예를 들어, 미국 특허 번호 7,053,264에 기재되어 있다.
- [0119] 예시적인 억제 도메인은 KRAB A/B, KOX, TGF-베타-유도성 초기 유전자 (TIEG), v-erbA, SID, MBD2, MBD3, DNMT 패밀리 구성원 (예를 들어, DNMT1, DNMT3A, DNMT3B), Rb, 및 MeCP2를 포함하지만, 이에 제한되지 않는다. 예를 들어, 문헌 [Bird et al. (1999) *Cell* 99:451-454; Tyler et al. (1999) *Cell* 99:443-446; Knoepfler et al. (1999) *Cell* 99:447-450; and Robertson et al. (2000) *Nature Genet.* 25:338-342] 참조. 부가의 예시적인 억제 도메인은 ROM2 및 AtHD2A를 포함하지만, 이에 제한되지 않는다. 예를 들어, 문헌 [Chem et al. (1996) *Plant Cell* 8:305-321; and Wu et al. (2000) *Plant J.* 22:19-27] 참조.
- [0120] 융합 분자는 관련 기술분야의 통상의 기술자에게 널리 공지되어 있는 클로닝 및 생화학적 접합 방법에 의해 구축된다. 융합 분자는 DNA-결합 도메인 (예를 들어, ZFP, TALE, sgRNA) 및 기능적 도메인 (예를 들어, 전사 활성화 또는 억제 도메인)을 포함한다. 융합 분자 또한 임의로, 핵 국제화 시그널 (예컨대, 예를 들어, SV40 매질 T-항원으로부터의 것) 및 에피토프 태그 (예컨대, 예를 들어, FLAG 및 적혈구 응집소)를 포함한다. 융합 단백질 (및 이를 코딩하는 핵산)은 번역 리딩 프레임이 융합 성분들 간에 보유되도록 디자인된다.
- [0121] 한편으로는 기능적 도메인 (또는 그의 기능적 단편)의 폴리펩티드 성분과, 다른 한편으로는 비-단백질 DNA-결합 도메인 (예를 들어, 항생제, 삽입제, 작은 홈 결합제, 핵산) 간의 융합물은 관련 기술분야의 통상의 기술자에게 공지된 생화학적 접합 방법에 의해 구축된다. 예를 들어, 문헌 [the Pierce Chemical Company (Rockford, IL) Catalogue] 참조. 작은 홈 결합제와 폴리펩티드 간에 융합물을 제조하는 방법 및 조성물이 기재되어 있다 (Mapp et al. (2000) *Proc. Natl. Acad. Sci. USA* 97:3930-3935). 더욱이, CRISPR/Cas 시스템의 단일 가이드 RNA가 기능적 도메인과 연합하여 활성의 전사 조절인자 및 뉴클레아제를 형성한다.
- [0122] 특정 실시양태에서, 표적 부위는 세포성 염색질의 접근가능한 영역에 존재한다. 접근가능한 영역은, 예를 들어, 미국 특허 번호 7,217,509 및 7,923,542에 기재된 바와 같이 결정될 수 있다. 표적 부위가 세포성 염색질의 접근가능한 영역에 존재하지 않는 경우, 하나 이상의 접근가능한 영역이 미국 특허 번호 7,785,792 및 8,071,370에 기재된 바와 같이 생성될 수 있다. 부가의 실시양태에서, 융합 분자의 DNA-결합 도메인은 그의 표

적 부위가 접근가능한 영역 내에 있는 지의 여부와 상관없이 세포성 염색질에 결합할 수 있다. 예를 들어, 이러한 DNA-결합 도메인은 링커 DNA 및/또는 뉴클레오솜 DNA에 결합할 수 있다. 이러한 유형의 "선구자" DNA 결합 도메인의 예는 특정의 스테로이드 수용체 및 간세포 핵 인자 3 (HNF3)에서 발견된다 (Cordingley et al. (1987) *Cell* 48:261-270; Pina et al. (1990) *Cell* 60:719-731; and Cirillo et al. (1998) *EMBO J.* 17:244-254).

[0123] 융합 분자는 관련 기술분야의 통상의 기술자에 공지된 바와 같은 제약상 허용되는 담체와 함께 제형화될 수 있다. 예를 들어, 문헌 [Remington's Pharmaceutical Sciences, 17th ed., 1985]; 및 미국 특허 번호 6,453,242 및 6,534,261 참조.

[0124] 융합 분자의 기능적 성분/도메인은, 일단 이러한 융합 분자가 그의 DNA 결합 도메인을 통하여 표적 서열에 결합하면, 유전자의 전사에 영향을 미칠 수 있는 각종 상이한 성분 중 어느 것으로부터 선택될 수 있다. 따라서, 기능적 성분은 다양한 전사 인자 도메인, 예컨대 활성화제, 억제인자, 공동 활성화제, 공동 억제인자, 및 침묵 인자를 포함할 수 있지만, 이에 제한되지 않는다.

[0125] 부가의 예시적인 기능적 도메인이 예를 들어, 미국 특허 번호 6,534,261 및 6,933,113에 개시된다.

[0126] 외인성 작은 분자 또는 리간드에 의해 조절되는 기능적 도메인이 또한 선택될 수 있다. 예를 들어, 기능적 도메인이 외부 레오켄(RheoChem)TM 리간드의 존재 하에 그의 활성 입체 형태만을 추정하는 레오스위치(RheoSwitch®) 기술이 이용될 수 있다 (예를 들어, US 20090136465 참조). 따라서, ZFP가 조절가능한 기능적 도메인과 작동적으로 연결될 수 있으며, 여기서 이로써 생성된 ZFP-TF의 활성은 외부 리간드에 의해 제어된다.

[0127] 뉴클레아제

[0128] 특정 실시양태에서, 융합 분자는 절단 (뉴클레아제) 도메인과 연합된 DNA-결합 도메인을 포함한다. 따라서, 유전자 변형은 뉴클레아제, 예를 들어 조작된 뉴클레아제를 사용하여 달성될 수 있다. 조작된 뉴클레아제 기술은 자연적으로 발생하는 DNA-결합 단백질을 조작하는 것에 근거한다. 예를 들어, 맞춘 DNA-결합 특이성을 수반한 귀소 엔도뉴클레아제의 조작이 기재되어 있다 (Chames et al. (2005) *Nucleic Acids Res* 33(20):e178; Arnould et al. (2006) *J. Mol. Biol.* 355:443-458). 또한, ZFP의 조작이 또한 기재되어 있다. 예를 들어, 미국 특허 번호 6,534,261; 6,607,882; 6,824,978; 6,979,539; 6,933,113; 7,163,824; 및 7,013,219 참조.

[0129] 또한, ZFP 및/또는 TALE를 뉴클레아제 도메인과 융합시켜 ZFN 및 TALEN을 창출시킬 수 있으며, 이는 그의 조작된 (ZFP 또는 TALE) DNA 결합 도메인을 통하여 그의 의도된 핵산 표적을 인식할 수 있고 뉴클레아제 활성을 통하여 DNA가 DNA 결합 부위 근처에서 커팅될 수 있게 해주는 기능적 실체이다.

[0130] 따라서, 본원에 기재된 방법 및 조성물은 광범위하게 적용가능하고, 임의의 관심 뉴클레아제를 포함할 수 있다. 뉴클레아제의 비-제한적 예는 메가뉴클레아제, TALEN 및 징크 핑거 뉴클레아제를 포함한다. 뉴클레아제는 이중 DNA-결합 및 절단 도메인 (예를 들어, 징크 핑거 뉴클레아제; 이중 절단 도메인을 수반한 메가뉴클레아제 DNA-결합 도메인)을 포함할 수 있거나, 또는 대안적으로, 자연적으로 발생하는 뉴클레아제의 DNA-결합 도메인이 특정의 선택된 표적 부위에 결합하기 위해 변경될 수 있다 (예를 들어, 동족 결합 부위와 상이한 부위에 결합하도록 조작된 메가뉴클레아제).

[0131] 본원에 기재된 뉴클레아제 중 어느 것에서, 뉴클레아제는 조작된 TALE DNA-결합 도메인 및 뉴클레아제 도메인 (예를 들어, 엔도뉴클레아제 및/또는 메가뉴클레아제 도메인) (TALEN으로서 지칭되기도 함)을 포함할 수 있다. 사용자가 선택한 표적 서열과 강력한 부위 특이적 상호작용하기 위하여 이들 TALEN 단백질을 조작하는 방법 및 조성물이 공개되었다 (미국 특허 번호 8,586,526 참조). 일부 실시양태에서, TALEN은 엔도뉴클레아제 (예를 들어, FokI) 절단 도메인 또는 절단 절반-도메인을 포함한다. 다른 실시양태에서, TALE-뉴클레아제는 메가 TALE이다. 이들 메가 TALE 뉴클레아제는 TALE DNA 결합 도메인 및 메가뉴클레아제 절단 도메인을 포함하는 융합 단백질이다. 메가뉴클레아제 절단 도메인은 단량체로서 활성이고, 활성을 위하여 이량체화를 필요로 하지 않는다. (문헌 [Boissel et al., (2013) *Nucl Acid Res*: 1-13, doi: 10.1093/nar/gkt1224] 참조). 또한, 뉴클레아제 도메인은 또한, DNA-결합 기능성을 나타낼 수 있다.

[0132] 또한 추가 실시양태에서, 뉴클레아제는 치밀한 TALEN (cTALEN)을 포함한다. 이들은 TALE DNA 결합 도메인을 *TevI* 뉴클레아제 도메인과 연결시키는 단일쇄 융합 단백질이다. 이러한 융합 단백질은 TALE 영역에 의해 국제된 니카제로서 작용할 수 있거나, 또는 TALE DNA 결합 도메인이 *TevI* 뉴클레아제 도메인에 대해서 위치되는 곳에 따라서 이중 가닥 파단을 창출시킬 수 있다 (문헌 [Beurdeley et al. (2013) *Nat Comm*: 1-8 DOI: 10.1038/ncomms2782] 참조). 임의의 TALEN이 부가의 TALEN (예를 들어, 하나 이상의 메가-TALE를 수반한 하나

이상의 TALEN (cTALEN 또는 FokI-TALEN)) 또는 다른 DNA 절단 효소와 조합하여 사용될 수 있다.

- [0133] 특정 실시양태에서, 뉴클레아제는 메가뉴클레아제 (귀소 엔도뉴클레아제) 또는 절단 활성을 나타내는 그의 특정 부분을 포함한다. 자연적으로 발생하는 메가뉴클레아제는 15-40개 염기-쌍 절단 부위를 인식하고, 통상적으로 하기 4가지 패밀리로 분류된다: LAGLIDADG 패밀리, GIY-YIG 패밀리, His-Cyst 박스 패밀리 및 HNH 패밀리. 예시적인 귀소 엔도뉴클레아제는 I-SceI, I-CeuI, PI-PspI, PI-Sce, I-SceIV, I-CsmI, I-PanI, I-SceII, I-PpoI, I-SceIII, I-CreI, I-TevI, I-TevII 및 I-TevIII를 포함한다. 그의 인식 서열이 공지되어 있다. 또한, 미국 특허 번호 5,420,032; 미국 특허 번호 6,833,252; 문헌 [Belfort et al. (1997) *Nucleic Acids Res.* 25:3379-3388; Dujon et al. (1989) *Gene* 82:115-118; Perler et al. (1994) *Nucleic Acids Res.* 22, 1125-1127; Jasin (1996) *Trends Genet.* 12:224-228; Gimble et al. (1996) *J. Mol. Biol.* 263:163-180; Argast et al. (1998) *J. Mol. Biol.* 280:345-353 and the New England Biolabs catalogue] 참조.
- [0134] 자연적으로 발생하는 메가뉴클레아제, 주로 LAGLIDADG 패밀리로부터의 DNA-결합 도메인이 식물, 효모, 드로소필라(*Drosophila*), 포유류 세포 및 마우스에서 부위-특이적 게놈 변형을 증진시키기 위해 사용되어 왔지만, 이러한 접근 방식은 메가뉴클레아제 인식 서열을 보존하는 상동 유전자의 변형 (문헌 [Monet et al. (1999), *Biochem. Biophysics. Res. Common.* 255: 88-93]) 또는 인식 서열이 도입된 미리-조작된 게놈에 대한 변형으로 제한되었다 (Route et al. (1994), *Mol. Cell. Biol.* 14: 8096-106; Chilton et al. (2003), *Plant Physiology.* 133: 956-65; Puchta et al. (1996), *Proc. Natl. Acad. Sci. USA* 93: 5055-60; Rong et al. (2002), *Genes Dev.* 16: 1568-81; Gouble et al. (2006), *J. Gene Med.* 8(5):616-622). 따라서, 의학적으로 또는 생물 공학적으로 관련된 부위에서 신규 결합 특이성을 나타내도록 메가뉴클레아제를 조작하는 시도가 있어 왔다 (Porteus et al. (2005), *Nat. Biotechnol.* 23: 967-73; Sussman et al. (2004), *J. Mol. Biol.* 342: 31-41; Epinat et al. (2003), *Nucleic Acids Res.* 31: 2952-62; Chevalier et al. (2002) *Molec. Cell* 10:895-905; Epinat et al. (2003) *Nucleic Acids Res.* 31:2952-2962; Ashworth et al. (2006) *Nature* 441:656-659; Paques et al. (2007) *Current Gene Therapy* 7:49-66; 미국 특허 공개 번호 20070117128; 20060206949; 20060153826; 20060078552; 및 20040002092). 또한, 메가뉴클레아제로부터의 자연적으로 발생하는 또는 조작된 DNA-결합 도메인은 이중 뉴클레아제 (예를 들어, *FokI*)로부터의 절단 도메인과 작동적으로 연결될 수 있고/거나 메가뉴클레아제로부터의 절단 도메인은 이중 DNA-결합 도메인 (예를 들어, ZFP 또는 TALE)과 작동적으로 연결될 수 있다.
- [0135] 다른 실시양태에서, 뉴클레아제는 징크 핑거 뉴클레아제 (ZFN) 또는 TALE DNA 결합 도메인-뉴클레아제 융합 (TALEN)이다. ZFN 및 TALEN은 선택되는 유전자 내의 표적 부위에 결합하도록 조작된 DNA 결합 도메인 (징크 핑거 단백질 또는 TALE DNA 결합 도메인) 및 절단 도메인 또는 절단 절반-도메인 (예를 들어, 본원에 기재된 바와 같은 제한 및/또는 메가뉴클레아제로부터 유래됨)을 포함한다.
- [0136] 상기에 상세히 기재된 바와 같이, 징크 핑거 결합 도메인 및 TALE DNA 결합 도메인은 선택되는 서열에 결합하도록 조작될 수 있다. 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] 참조. 조작된 징크 핑거 결합 도메인 또는 TALE 단백질은 자연적으로 발생하는 단백질과 비교해서 신규의 결합 특이성을 가질 수 있다. 조작 방법은 합리적 디자인 및 다양한 유형의 선택을 포함하지만, 이에 제한되지 않는다. 합리적 디자인은, 예를 들어, 삼중자 (또는 사중자) 뉴클레오티드 서열 및 개별 징크 핑거 또는 TALE 아미노산 서열을 포함하는 데이터베이스를 사용하는 것을 포함하며, 여기서 각각의 삼중자 또는 사중자 뉴클레오티드 서열은 특정한 삼중자 또는 사중자 서열에 결합하는 TALE 반복 단위 또는 징크 핑거의 하나 이상의 아미노산 서열과 연합된다. 예를 들어, 미국 특허 6,453,242 및 6,534,261 참조, 그 전문이 본원에 참조로 포함된다.
- [0137] 표적 부위의 선택; 및 융합 단백질 (및 이를 코딩하는 폴리뉴클레오티드)의 디자인 및 구축 방법은 관련 기술분야의 통상의 기술자에게 공지되어 있고, 미국 특허 번호 7,888,121 및 8,409,861 (그 전문이 본원에 참조로 포함됨)에 상세히 기재되어 있다.
- [0138] 또한, 이들 및 다른 참고문헌에 개시된 바와 같이, 징크 핑거 도메인, TALE 및/또는 다중 핑거된 징크 핑거 단백질은 임의의 적합한 링커 서열, 예를 들어, 5개 이상의 아미노산 길이의 링커를 사용하여 함께 연결시킬 수 있다. 6개 이상의 아미노산 길이의 예시적인 링커 서열에 관해서는, 예를 들어, 미국 특허 번호 6,479,626; 6,903,185; 및 7,153,949를 참조할 수 있다. 본원에 기재된 단백질은 이러한 단백질의 개별 징크 핑거 간의 적합한 링커의 임의의 조합을 포함할 수 있다. 또한, 미국 특허 번호 8,772,453 참조.

- [0139] 따라서, 뉴클레아제, 예컨대 ZFN, TALEN 및/또는 메가뉴클레아제는 임의의 DNA-결합 도메인 및 임의의 뉴클레아제 (절단) 도메인 (절단 도메인, 절단 절반-도메인)을 포함할 수 있다. 상기 기재된 바와 같이, 절단 도메인은 DNA-결합 도메인, 예를 들어 징크 핑거 또는 TAL-이펙터 DNA-결합 도메인, 뉴클레아제 또는 메가뉴클레아제 DNA-결합 도메인으로부터의 절단 도메인 및 상이한 뉴클레아제로부터의 절단 도메인에 대해 이중될 수 있다. 이중 절단 도메인은 임의의 엔도뉴클레아제 또는 엑소뉴클레아제로부터 수득될 수 있다. 절단 도메인이 유래될 수 있는 예시적인 엔도뉴클레아제는 제한 엔도뉴클레아제 및 귀소 엔도뉴클레아제를 포함하지만, 이에 제한되지 않는다. 예를 들어, 문헌 [2002-2003 Catalogue, New England Biolabs, Beverly, MA; and Belfort et al. (1997) *Nucleic Acids Res.* 25:3379-3388] 참조. DNA를 절단하는 부가의 효소가 공지되어 있다 (예를 들어, S1 뉴클레아제; 녹두 뉴클레아제; 췌장 DNase I; 마이크로코쿠스 뉴클레아제; 효모 HO 엔도뉴클레아제; 또한 문헌 [Linn et al. (eds.) *Nucleases*, Cold Spring Harbor Laboratory Press, 1993] 참조). 이들 효소 (또는 그의 기능적 단편) 중 하나 이상이 절단 도메인 및 절단 절반-도메인의 공급원으로서 사용될 수 있다.
- [0140] 유사하게, 절단 절반-도메인은 상기 제시된 바와 같이, 절단 활성을 위하여 이량체화를 필요로 하는 임의의 뉴클레아제 또는 그의 부분으로부터 유래될 수 있다. 일반적으로, 2개의 융합 단백질이 절단 절반-도메인을 포함하는 경우, 이러한 융합 단백질은 절단할 필요가 있다. 대안적으로, 2개의 절단 절반-도메인을 포함하는 단일 단백질이 사용될 수 있다. 2개의 절단 절반-도메인은 동일한 엔도뉴클레아제 (또는 그의 기능적 단편)로부터 유래될 수 있거나, 또는 각각의 절단 절반-도메인은 상이한 엔도뉴클레아제 (또는 그의 기능적 단편)로부터 유래될 수 있다. 또한, 2개의 융합 단백질에 대한 표적 부위는 바람직하게 서로에 대해 배치되어, 이러한 2개의 융합 단백질을 그들 각각의 표적 부위에 결합시킴으로써 절단 절반-도메인을 서로에 대해 공간적 배향으로 위치시켜 절단 절반-도메인이, 예를 들어 이량체화에 의해 기능적 절단 도메인을 형성할 수 있도록 한다. 따라서, 특정 실시양태에서, 표적 부위의 근처 모서리는 5-8개의 뉴클레오티드 또는 15-18개의 뉴클레오티드에 의해 분리된다. 그러나, 임의의 정수의 뉴클레오티드 또는 뉴클레오티드 쌍이 두 표적 부위 사이에 개입될 수 있다 (예를 들어, 2 내지 50개의 뉴클레오티드 쌍 또는 그 초과). 일반적으로, 절단 부위는 표적 부위 사이에 있지만, 절단 부위로부터 1 킬로염기 이상 떨어져 있을 수 있으며, 이는 절단 부위로부터 1-50개 염기 쌍 (또는 1-5개, 1-10개, 및 1-20개 염기 쌍을 포함한, 그 사이의 임의의 값), 1-100개 염기 쌍 (또는 그 사이의 임의의 값), 100-500개 염기 쌍 (또는 그 사이의 임의의 값), 500 내지 1000개 염기 쌍 (또는 그 사이의 임의의 값) 또는 심지어 1 kb 초과를 포함한다.
- [0141] 제한 엔도뉴클레아제 (제한 효소)는 많은 종에 존재하고, (인식 부위에서) DNA와 서열-특이적으로 결합할 수 있으며 결합 부위에서 또는 그 근처에서 DNA를 절단할 수 있다. 특정의 제한 효소 (예를 들어, 유형 IIS)는 인식 부위로부터 제거된 부위에서 DNA를 절단하고, 분리가 가능한 결합 및 절단 도메인을 갖는다. 예를 들어, 유형 IIS 효소 Fok I은 한 가닥 상에서 그의 인식 부위로부터 9개 뉴클레오티드에서 DNA의 이중 가닥 절단을 촉매하고, 다른 가닥 상에서는 그의 인식 부위로부터 13개 뉴클레오티드에서 DNA의 이중 가닥 절단을 촉매한다. 예를 들어, 미국 특허 5,356,802; 5,436,150 및 5,487,994; 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, 한 실시양태에서, 융합 단백질은 적어도 하나의 유형 IIS 제한 효소로부터의 절단 도메인 (또는 절단 절반-도메인) 및 조작될 수 있거나 또는 조작되지 않을 수 있는 하나 이상의 징크 핑거 결합 도메인을 포함한다.
- [0142] 그의 절단 도메인이 결합 도메인으로부터 분리가 가능한 예시적인 유형 IIS 제한 효소는 Fok I이다. 이러한 특정한 효소는 이량체로서 활성이다 (Bitinaite et al. (1998) *Proc. Natl. Acad. Sci. USA* 95: 10,570-10,575). 따라서, 본 개시내용의 목적상, 본원에 개시된 융합 단백질에 사용된 Fok I 효소의 일정 부분이 절단 절반-도메인으로서 간주된다. 따라서, 징크 핑거-Fok I 융합물을 사용하여 세포성 서열의 표적화된 이중 가닥 절단 및/또는 표적화된 대체를 위해서는, 각각 Fok I 절단 절반-도메인을 포함하는 2개의 융합 단백질을 사용하여, 촉매적으로 활성인 절단 도메인을 재구성할 수 있다. 대안적으로, 징크 핑거 결합 도메인 및 2개의 Fok I 절단 절반-도메인을 함유하는 단일 폴리펩티드 분자를 사용할 수도 있다. 징크 핑거-Fok I 융합물을 사용하여 표적화된 절단 및 표적화된 서열 변경하기 위한 파라미터가 본 개시내용의 다른 곳에 제공된다.
- [0143] 절단 도메인 또는 절단 절반-도메인은 절단 활성을 보유하고 있거나 또는 기능적 절단 도메인을 형성하기 위해 다량체화 (예를 들어, 이량체화)할 수 있는 능력을 보유하고 있는 단백질의 임의의 부분일 수 있다.
- [0144] 예시적인 유형 IIS 제한 효소가 국제 공개 WO 07/014275 (그 전문이 본원에 참조로 포함됨)에 기재되어 있다. 부가의 제한 효소는 또한, 분리가 가능한 결합 및 절단 도메인을 함유하고, 이들은 본 개시내용에 의해 고려된다.

예를 들어, 문헌 [Roberts et al. (2003) *Nucleic Acids Res.* 31:418-420] 참조.

- [0145] 특정 실시양태에서, 절단 도메인은, 예를 들어, 미국 특허 번호 7,914,796; 8,034,598 및 8,623,618; 및 미국 특허 공개 번호 20110201055 (이들 모두의 개시내용은 그 전문이 본원에 참조로 포함됨)에 기재된 바와 같이, 동종이량체화를 최소화 또는 방지시키는 하나 이상의 조작된 절단 절반-도메인 (이량체화 도메인 돌연변이체로서 지칭되기도 함)을 포함한다. *FokI*의 위치 446, 447, 479, 483, 484, 486, 487, 490, 491, 496, 498, 499, 500, 531, 534, 537, 및 538에서의 아미노산 잔기는 *Fok I* 절단 절반-도메인의 이량체화에 영향을 미치기 위한 모든 표적이다.
- [0146] 절대적인 이중-이량체를 형성하는 *FokI*의 예시적인 조작된 절단 절반-도메인은 제1 절단 절반-도메인이 *FokI*의 위치 490 및 538에서의 아미노산 잔기에서의 돌연변이를 포함하고, 제2 절단 절반-도메인이 아미노산 잔기 486 및 499에서의 돌연변이를 포함하는 쌍을 포함한다.
- [0147] 따라서, 한 실시양태에서, 490에서의 돌연변이는 Glu (E)를 Lys (K)로 대체시키고; 538에서의 돌연변이는 Iso (I)를 Lys (K)로 대체시키며; 486에서의 돌연변이는 Gln (Q)을 Glu (E)로 대체시키고; 위치 499에서의 돌연변이는 Iso (I)를 Lys (K)로 대체시킨다. 구체적으로, 본원에 기재된 조작된 절단 절반-도메인은 하나의 절단 절반-도메인 내에서 위치 490 (E→K) 및 538 (I→K)을 돌연변이시켜 "E490K:I538K"로 지명된 조작된 절단 절반-도메인을 생성시키고, 또 다른 절단 절반-도메인 내에서 위치 486 (Q→E) 및 499 (I→L)를 돌연변이시켜 "Q486E:I499L"로 지명된 조작된 절단 절반-도메인을 생성시킴으로써 제조되었다. 본원에 기재된 조작된 절단 절반-도메인은 이상한 절단을 최소화하거나 또는 폐지시킨 절대적인 이중-이량체 돌연변이체이다. 예를 들어, 미국 특허 번호 7,914,796 및 8,034,598 참조; 그의 개시내용은 모든 목적을 위해 그 전문이 참조로 포함된다. 특정 실시양태에서, 조작된 절단 절반-도메인은 위치 486, 499 및 496 (야생형 *FokI*와 비교해서 넘버링됨)에서의 돌연변이, 예를 들어 위치 486에서의 야생형 Gln (Q) 잔기를 Glu (E) 잔기로 대체시키고, 위치 499에서의 야생형 Iso (I) 잔기를 Leu (L) 잔기로 대체시키며, 위치 496에서의 야생형 Asn (N) 잔기를 Asp (D) 또는 Glu (E) 잔기로 대체시키는 돌연변이 (각각 "ELD" 및 "ELE" 도메인으로서 지칭되기도 함)를 포함한다. 다른 실시양태에서, 조작된 절단 절반-도메인은 위치 490, 538 및 537 (야생형 *FokI*와 비교해서 넘버링됨)에서의 돌연변이, 예를 들어 위치 490에서의 야생형 Glu (E) 잔기를 Lys (K) 잔기로 대체시키고, 위치 538에서의 야생형 Iso (I) 잔기를 Lys (K) 잔기로 대체시키며, 위치 537에서의 야생형 His (H) 잔기를 Lys (K) 잔기 또는 Arg (R) 잔기로 대체시키는 돌연변이 (각각 "KKK" 및 "KKR" 도메인으로서 지칭되기도 함)를 포함한다. 다른 실시양태에서, 조작된 절단 절반-도메인은 위치 490 및 537 (야생형 *FokI*와 비교해서 넘버링됨)에서의 돌연변이, 예를 들어 위치 490에서의 야생형 Glu (E) 잔기를 Lys (K) 잔기로 대체시키고, 위치 537에서의 야생형 His (H) 잔기를 Lys (K) 잔기 또는 Arg (R) 잔기로 대체시키는 돌연변이 (각각 "KIK" 및 "KIR" 도메인으로서 지칭되기도 함)를 포함한다. 예를 들어, 미국 특허 번호 7,914,796; 8,034,598 및 8,623,618 참조; 그의 개시내용은 모든 목적을 위해 그 전문이 참조로 포함된다. 다른 실시양태에서, 조작된 절단 절반 도메인은 "샤키(Sharkey)" 및/또는 "샤키" 돌연변이를 포함한다 (문헌 [Guo et al., (2010) *J. Mol. Biol.* 400(1):96-107] 참조).
- [0148] 대안적으로, 뉴클레아제는 소위 "분할-효소" 기술을 이용하여 핵산 표적 부위에서 생체내에서 어셈블리될 수 있다 (예를 들어, 미국 특허 공개 번호 20090068164 참조). 이러한 분할 효소의 성분은 별도의 발현 구조물 상에서 발현될 수 있거나, 또는 개별 성분이, 예를 들어, 자기 절단성 2A 펩티드 또는 IRES 서열에 의해 분리되는 하나의 오픈 리딩 프레임 내에서 연결될 수 있다. 성분은 개별 징크 핑거 결합 도메인 또는 메가뉴클레아제 핵산 결합 도메인일 수 있다.
- [0149] 뉴클레아제 (예를 들어, ZFN 및/또는 TALEN)를 대상으로 하여, 사용하기 이전에, 예를 들어 미국 특허 번호 8,563,314에 기재된 바와 같은 효모 기반 염색체 시스템에서 활성화에 관하여 스크리닝할 수 있다.
- [0150] 특정 실시양태에서, 뉴클레아제는 CRISPR/Cas 시스템을 포함한다. 이러한 시스템의 RNA 성분을 코딩하는 CRISPR (클러스터된 규칙적으로 공간을 둔 짧은 팔린드롬성 반복 서열) 로커스, 및 단백질을 코딩하는 Cas (CRISPR-관련) 로커스 (문헌 [Jansen et al., 2002. *Mol. Microbiol.* 43: 1565-1575; Makarova et al., 2002. *Nucleic Acids Res.* 30: 482-496; Makarova et al., 2006. *Biol. Direct* 1: 7; Haft et al., 2005. *PLoS Comput. Biol.* 1: e60])가 CRISPR/Cas 뉴클레아제 시스템의 유전자 서열을 구성한다. 미생물 숙주 내에서의 CRISPR 로커스는 CRISPR-관련 (Cas) 유전자 뿐만 아니라 CRISPR-매개된 핵산 절단의 특이성을 프로그래밍할 수 있는 비-코딩 RNA 요소의 조합을 함유한다.
- [0151] 유형 II CRISPR은 가장 널리 특징규명된 시스템 중 하나이고, 4가지 일련의 단계에서 표적화된 DNA 이중 가닥 파단을 수행한다. 첫째, 2개의 비-코딩 RNA인 프리-crRNA 어레이 및 tracrRNA가 상기 CRISPR 로커스로부터 전

사된다. 둘째, tracrRNA는 프리-crRNA의 반복 영역과 혼성화되고, 프리-crRNA가 개별 스페이서 서열을 함유하는 성숙한 crRNA로 프로세싱되는 것을 매개한다. 셋째, 성숙한 crRNA:tracrRNA 복합체는 crRNA 상의 스페이서와, 표적 인식을 위해 부가적으로 요구되는 프로토스페이서 인접 모티프 (PAM) 다음에 있는 표적 DNA 상의 프로토스페이서 간의 왓슨-크릭 염기-쌍형성을 통하여 Cas9가 표적 DNA를 향하게 한다. 최종적으로, Cas9는 표적 DNA의 절단을 매개하여 프로토스페이서 내에 이중 가닥 파단을 창출시킨다. CRISPR/Cas 시스템의 활성화는 3가지 단계를 포함한다: (i) '적응'으로 지칭되는 프로세스에서, 향후의 공격을 방지하기 위해 CRISPR 어레이 내로의 외계 DNA 서열의 삽입, (ii) 관련 단백질의 발현 뿐만 아니라 상기 어레이의 발현 및 프로세싱에 이어, (iii) 상기 외계 핵산으로의 RNA-매개된 간섭. 따라서, 박테리아 세포에서는, 소위 'Cas' 단백질 중 몇 가지가 CRISPR/Cas 시스템의 자연적 기능과 관련이 있고, 외계 DNA의 삽입 등과 같은 기능에 있어서 일정 역할을 한다.

[0152] 일부 실시양태에서, CRISPR-Cpf1 시스템이 사용된다. 프란시셀라 종(*Francisella* spp)에서 확인된 CRISPR-Cpf1 시스템은 인간 세포에서의 강력한 DNA 간섭을 매개하는 부류 2 CRISPR-Cas 시스템이다. 기능적으로 보존되긴 하지만, Cpf1과 Cas9는 많은 측면에 있어서 상이하며, 예를 들어 그들의 가이드 RNA와 기질 특이성 측면에서 상이하다 (문헌 [Fagerlund et al., (2015) *Genom Bio* 16:251] 참조). Cas9 단백질과 Cpf1 단백질 간의 주요 차이는 Cpf1이 tracrRNA를 활용하지 않으므로, crRNA 만을 필요로 한다는 것이다. FnCpf1 crRNA는 42-44개 뉴클레오타이드 길이 (19개 뉴클레오타이드 반복 서열 및 23-25개 뉴클레오타이드 스페이서)이고, 2차 구조를 유지시키는 서열 변화를 관용하는 단일 줄기-루프를 함유한다. 또한, Cpf1 crRNA는 Cas9에 의해 요구되는 ~100개 뉴클레오타이드 조작된 sgRNA보다 상당히 더 짧고, FnCpf1에 대한 PAM 요구 사항은 변위된 가닥 상에서의 5'-TTN-3' 및 5'-CTA-3'이다. Cas9와 Cpf1 둘 다 표적 DNA 내에 이중 가닥 파단을 만들긴 하지만, Cas9는 그의 RuvC-및 HNH-유사 도메인을 사용하여, 가이드 RNA의 시드 서열 내에 평행 말단 커트를 만드는 반면, Cpf1은 RuvC-유사 도메인을 사용하여, 시드의 외부에 엇갈린 커트를 생성시킨다. Cpf1은 중대한 시드 영역으로부터 멀리 떨어져 엇갈린 커트를 만들기 때문에, NHEJ는 표적 부위를 붕괴시키지 않을 것이므로, 목적하는 HDR 재조합 이벤트가 일어날 때까지 Cpf1이 동일한 부위를 지속적으로 커팅할 수 있게 해준다. 따라서, 본원에 기재된 방법 및 조성물에서, 용어 "Cas"는 Cas9 단백질과 Cpf1 단백질 둘 다를 포함하는 것으로 이해된다. 따라서, 본원에 사용된 바와 같이, "CRISPR/Cas 시스템"은 CRISPR/Cas 및/또는 CRISPR/Cpf1 시스템 둘 다를 지칭하며, 이는 뉴클레아제 및/또는 전사 인자 시스템 둘 다를 포함한다.

[0153] 특정 실시양태에서, Cas 단백질은 자연적으로 발생하는 Cas 단백질의 "기능적 유도체"일 수 있다. 천연 서열 폴리펩티드의 "기능적 유도체"는 천연 서열 폴리펩티드와 공통으로 정성적 생물학적 특성을 갖는 화합물이다. "기능적 유도체"는 천연 서열의 단편, 및 천연 서열 폴리펩티드 및 그의 단편의 유도체를 포함하지만, 이에 제한되지 않으며, 단 이들은 상응하는 천연 서열 폴리펩티드와 공통으로 생물학적 활성을 가져야 한다. 본원에서 고려된 생물학적 활성은 DNA 기질을 단편으로 가수분해시킬 수 있는 기능적 유도체의 능력이다. 용어 "유도체"는 폴리펩티드의 아미노산 서열 변이체, 공유적 변형물, 및 그의 융합물 모두를 포괄한다. Cas 폴리펩티드 또는 그의 단편의 적합한 유도체는 Cas 단백질 또는 그의 단편의 돌연변이체, 융합물, 공유적 변형물을 포함하지만, 이에 제한되지 않는다. Cas 단백질 또는 그의 단편을 포함하는 Cas 단백질 뿐만 아니라 Cas 단백질 또는 그의 단편의 유도체는 세포로부터 수득가능할 수 있거나 또는 화학적으로 합성될 수 있거나 또는 이들 두 절차의 조합에 의해 수득가능할 수 있다. 상기 세포는 Cas 단백질을 자연적으로 생성하는 세포, 또는 Cas 단백질을 자연적으로 생성하고 내인성 Cas 단백질을 더 높은 발현 수준에서 생성하도록 유전적으로 조작되거나 또는 외인적으로 도입된 핵산으로부터 Cas 단백질을 생성하도록 유전적으로 조작되는 세포일 수 있으며, 핵산은 내인성 Cas와 동일하거나 또는 상이한 Cas를 코딩한다. 일부 경우에, 상기 세포는 Cas 단백질을 자연적으로 생성하지 않고 Cas 단백질을 생성하도록 유전적으로 조작된다.

[0154] TCR 유전자 및 다른 유전자에 대해 표적화된 예시적인 CRISPR/Cas 뉴클레아제 시스템이, 예를 들어, 미국 공개 번호 20150056705에 개시된다.

[0155] 뉴클레아제(들)는 표적 부위 내에 하나 이상의 이중 가닥 및/또는 단일 가닥 커트를 만들 수 있다. 특정 실시양태에서, 뉴클레아제는 촉매적으로 불활성인 절단 도메인 (예를 들어, FokI 및/또는 Cas 단백질)을 포함한다. 예를 들어, 미국 특허 번호 9,200,266; 8,703,489 및 문헌 [Guillinger et al. (2014) *Nature Biotech.* 32(6):577-582] 참조. 촉매적으로 불활성인 절단 도메인은 촉매적으로 활성인 도메인과 조합하여, 니카제로서 작용하여 단일 가닥 커트를 만들 수 있다. 따라서, 2개의 니카제를 조합하여 사용하여 특이적 영역에서 이중 가닥 커트를 만들 수 있다. 부가의 니카제가 또한 관련 기술분야, 예를 들어, 문헌 [McCaffery et al. (2016) *Nucleic Acids Res.* 44(2):e11. doi: 10.1093/nar/gkv878. Epub 2015 Oct 19]에 공지되어 있다.

- [0156] 전달
- [0157] 단백질 (예를 들어, 전사 인자, 뉴클레아제, TCR 및 CAR 분자), 폴리뉴클레오티드, 및/또는 본원에 기재된 단백질 및/또는 폴리뉴클레오티드를 포함하는 조성물은, 예를 들어, 단백질 및/또는 mRNA 성분의 주입에 의한 것을 포함한 임의의 적합한 수단에 의해 표적 세포로 전달될 수 있다.
- [0158] 적합한 세포는 진핵 및 원핵 세포 및/또는 세포주를 포함하지만, 이에 제한되지 않는다. 이러한 세포 또는 이러한 세포로부터 생성된 세포주의 비-제한적 예는 T-세포, COS, CHO (예를 들어, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11, CHO-DUKX, CHOK1SV), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NSO, SP2/0-Ag14, HeLa, HEK293 (예를 들어, HEK293-F, HEK293-H, HEK293-T), 및 perC6 세포 뿐만 아니라 곤충 세포, 예컨대 스포둑테라 푸기페르다(*Spodoptera fugiperda*) (Sf), 또는 진균 세포, 예컨대 사카로미세스(*Saccharomyces*), 피키아(Pichia) 및 슈조사카로미세스(*Schizosaccharomyces*)를 포함한다. 특정 실시양태에서, 세포주는 CHO-K1, MDCK 또는 HEK293 세포주이다. 적합한 세포는 또한, 줄기 세포, 예컨대, 한 예로서, 배아 줄기 세포, 유도된 다능성 줄기 세포 (iPS 세포), 조혈 줄기 세포, 뉴런의 줄기 세포 및 중간엽 줄기 세포를 포함한다.
- [0159] 본원에 기재된 바와 같은 DNA-결합 도메인을 포함하는 단백질의 전달 방법은, 예를 들어, 미국 특허 번호 6,453,242; 6,503,717; 6,534,261; 6,599,692; 6,607,882; 6,689,558; 6,824,978; 6,933,113; 6,979,539; 7,013,219; 및 7,163,824 (이들 모두의 개시내용은 그 전문이 본원에 참조로 포함됨)에 기재되어 있다.
- [0160] DNA 결합 도메인, 및 본원에 기재된 바와 같은 이들 DNA 결합 도메인을 포함하는 융합 단백질은 또한, DNA-결합 단백질(들) 중 하나 이상을 코딩하는 서열을 함유하는 벡터를 사용하여 전달될 수 있다. 부가적으로, 부가의 핵산 (예를 들어, 공여자)은 또한, 이들 벡터를 통하여 전달될 수 있다. 플라스미드 벡터, 레트로바이러스 벡터, 렌티바이러스 벡터, 아데노바이러스 벡터, 폭스바이러스 벡터; 헤르페스바이러스 벡터 및 아데노 관련 바이러스 벡터 등을 포함하지만, 이에 제한되지 않는 임의의 벡터 시스템을 사용할 수 있다. 또한, 미국 특허 번호 6,534,261; 6,607,882; 6,824,978; 6,933,113; 6,979,539; 7,013,219; 및 7,163,824 참조, 그 전문이 본원에 참조로 포함된다. 더욱이, 이들 벡터 중 어느 것이 하나 이상의 DNA-결합 단백질-코딩 서열 및/또는 부가의 핵산을 적절히 포함할 수 있다는 것이 명백할 것이다. 따라서, 본원에 기재된 바와 같은 하나 이상의 DNA-결합 단백질이 세포 내로 도입되고, 부가의 DNA가 적절히 도입되는 경우, 이들은 동일한 벡터 또는 상이한 벡터 상에서 수행될 수 있다. 다수의 벡터가 사용되는 경우, 각각의 벡터는 1개 또는 다수의 DNA-결합 단백질 및 원하는 만큼의 부가의 핵산을 코딩하는 서열을 포함할 수 있다.
- [0161] 통상적인 바이러스 및 비-바이러스 기반 유전자 전이 방법을 사용하여 세포 (예를 들어, 포유류 세포) 및 표적 조직에 조작된 DNA-결합 단백질을 코딩하는 핵산을 도입하과, 원하는 만큼의 부가의 뉴클레오티드 서열을 공동 도입할 수 있다. 이러한 방법을 또한 사용하여 핵산 (예를 들어, DNA-결합 단백질 및/또는 공여자를 코딩함)을 시험관 내에서 세포에 투여할 수 있다. 특정 실시양태에서, 핵산은 생체내 또는 생체외 유전자 요법 사용을 위하여 투여된다. 비-바이러스 벡터 전달 시스템은 DNA 플라스미드, 네이키드 핵산, 및 전달 비히클, 예컨대 리포솜, 지질 나노입자 또는 폴록사머와 복합체를 형성한 핵산을 포함한다. 바이러스 벡터 전달 시스템은 세포로의 전달 후에 에피솜성 또는 통합된 게놈을 갖는 DNA 및 RNA 바이러스를 포함한다. 유전자 요법 절차에 관한 고찰을 위해, 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., in *Current Topics in Microbiology and Immunology* Doerfler and Boehm (eds.) (1995); and Yu et al., *Gene Therapy* 1:13-26 (1994)]을 참조할 수 있다.
- [0162] 핵산의 비-바이러스 전달 방법은 전기천공, 리포펙션, 미세주사, 바이오리스팅, 비로솜, 리포솜, 지질 나노입자, 면역리포솜, 다가양이온 또는 지질:핵산 접합체, 네이키드 DNA, mRNA, 인공 비리온, 및 DNA의 작용제-증강된 흡수를 포함한다. 예를 들어, 소니트론(Sonitron) 2000 시스템 (Rich-Mar)을 사용하는 초음파천공이 또한, 핵산의 전달을 위해 사용될 수 있다. 바람직한 실시양태에서, 하나 이상의 핵산이 mRNA로서 전달된다. 또한 바람직한 것은 번역 효율 및/또는 mRNA 안정성을 증가시키기 위해 캡핑된 mRNA를 사용하는 것이다. 특히 바람직한 것은 ARCA (항-역전 캡 유사체) 캡 또는 그의 변이체이다. 미국 특허 번호 7,074,596 및 8,153,773 참조, 본원에 참조로 포함된다.
- [0163] 부가의 예시적인 핵산 전달 시스템은 아막사 바이오시스템즈(Amaxa Biosystems; 독일 쾰른), 맥스사이트, 인크.(Maxcyte, Inc.; 미국 매릴랜드주 록빌), BTX 분자 전달 시스템(BTX Molecular Delivery Systems;

미국 매사추세츠주 홀리스톤) 및 코페르니쿠스 테라퓨틱스 인크.(Copernicus Therapeutics Inc.)에 의해 제공된 것을 포함한다 (예를 들어, US6008336 참조). 리포펙션은, 예를 들어, US 5,049,386, US 4,946,787; 및 US 4,897,355에 기재되어 있고, 리포펙션 시약은 상업적으로 판매된다 (예를 들어, 트랜스펙탐(Transfectam)TM, 리포펙틴(Lipofectin)TM, 및 리포펙타민(Lipofectamine)TM RNAiMAX). 폴리뉴클레오티드의 효율적인 수용체-인식 리포펙션에 적합한 양이온성 및 중성 지질은 펠그너(Felgner), WO 91/17424, WO 91/16024의 것을 포함한다. 세포 (생체의 투여) 또는 표적 조직 (생체내 투여)에 전달될 수 있다.

[0164] 표적화된 리포솜, 예컨대 면역지질 복합체를 포함한 지질:핵산 복합체의 제조는 관련 기술분야의 통상의 기술자에게 널리 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 번호 4,186,183, 4,217,344, 4,235,871, 4,261,975, 4,485,054, 4,501,728, 4,774,085, 4,837,028, 및 4,946,787 참조).

[0165] 부가의 전달 방법은 EnGeneIC 전달 비히클 (EDV) 내로 전달될 핵산의 패키징을 이용하는 것을 포함한다. 이들 EDV는 항체의 하나의 아암이 표적 조직에 대한 특이성을 갖고 있고 다른 아암이 EDV에 대한 특이성을 갖고 있는 이중특이적 항체를 사용하여 표적 조직에 특이적으로 전달된다. 항체는 표적 세포 표면에 EDV를 가져온 다음 EDV를 세포내 이입에 의해 세포 내로 가져온다. 일단 세포 내에 있으면, 그 내용물이 방출된다 (문헌 [MacDiarmid et al. (2009) *Nature Biotechnology* 27(7) p. 643] 참조).

[0166] 조작된 DNA-결합 단백질을 코딩하는 핵산, 및/또는 원하는 만큼의 공여자 (예를 들어 CAR 또는 ACTR)의 전달을 위하여 RNA 또는 DNA 바이러스 기반 시스템을 사용하는 것은 바이러스가 체내의 특이적 세포를 표적으로 하고 바이러스 페이로드(payload)를 핵으로 트래핑하기 위한 고도로 진화된 프로세스를 활용한다. 바이러스 벡터는 환자에게 직접적으로 투여될 수 있거나 (생체내) 또는 시험관 내에서 세포를 처리하기 위해 사용될 수 있고, 이와 같이 변형된 세포가 환자에게 투여된다 (생체외). 핵산을 전달하기 위한 통상적인 바이러스 기반 시스템은 유전자 전이를 위한 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노 관련, 백시니아 및 단순 포진 바이러스 벡터를 포함하지만, 이에 제한되지 않는다. 숙주 게놈에서의 통합은 레트로바이러스, 렌티바이러스, 및 아데노 관련 바이러스 유전자 전이 방법을 이용하여 가능하며, 이로써 종종 삽입된 트랜스진의 장기간 발현이 초래된다. 부가적으로, 많은 상이한 세포 유형 및 표적 조직에서 높은 형질도입 효율이 관찰되었다.

[0167] 레트로바이러스의 지향성은 외래 외막 단백질을 혼입시켜, 표적 세포의 잠재적 표적 집단을 확장시킴으로써 변경될 수 있다. 렌티바이러스 벡터는 비-분할 세포를 형질도입 또는 감염시킬 수 있고 전형적으로 높은 바이러스 역가를 생성시킬 수 있는 레트로바이러스 벡터이다. 레트로바이러스 유전자 전이 시스템의 선택은 표적 조직에 좌우된다. 레트로바이러스 벡터는 6-10 kb 이하의 외래 서열에 대한 패키징 성능을 지닌 시스 작용성 장 말단 반복 서열로 구성된다. 최소한의 시스 작용성 LTR은 벡터의 복제 및 패키징에 충분하므로, 이어서 이를 사용하여 치료용 유전자를 표적 세포 내로 통합시켜 영구적인 트랜스진 발현을 제공한다. 광범위하게 사용된 레트로바이러스 벡터는 무린 백혈병 바이러스 (MuLV), 긴팔원숭이 유인원 백혈병 바이러스 (GaLV), 유인원의 면역결핍 바이러스 (SIV), 인간 면역결핍 바이러스 (HIV), 및 그의 조합에 근거한 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., *Viol.* 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).

[0168] 일시적인 발현이 바람직한 적용에서는, 아데노바이러스 기반 시스템을 사용할 수 있다. 아데노바이러스 기반 벡터는 많은 세포 유형에서 매우 높은 형질도입 효율을 나타낼 수 있고 세포 분열을 필요로 하지 않는다. 이러한 벡터를 이용하여, 높은 역가 및 높은 수준의 발현이 수득되었다. 이러한 벡터는 비교적 단순한 시스템에서 다량으로 생성될 수 있다. 아데노 관련 바이러스 ("AAV") 벡터는 또한, 예를 들어, 핵산 및 펩티드의 시험관내 생성에서, 및 생체내 및 생체외 유전자 요법 절차의 경우에, 세포를 표적 핵산으로 형질도입하기 위해 사용된다 (예를 들어, 문헌 [West et al., *Virology* 160:38-47 (1987)]; 미국 특허 번호 4,797,368; WO 93/24641; 문헌 [Kotin, *Human Gene Therapy* 5:793-801 (1994); Muzyczka, *J. Clin. Invest.* 94:1351 (1994)] 참조). 재조합 AAV 벡터의 구축이 수많은 공개에 기재되어 있으며, 이러한 공개는 미국 특허 번호 5,173,414; 문헌 [Tratschin et al., *Mol. Cell. Biol.* 5:3251-3260 (1985); Tratschin, et al., *Mol. Cell. Biol.* 4:2072-2081 (1984); Hermonat & Muzyczka, *PNAS USA* 81:6466-6470 (1984); and Samulski et al., *J. Virol.* 63:03822-3828 (1989)]을 포함한다.

- [0169] 형질도입 작용제를 생성시키기 위해 헬퍼 세포주 내로 삽입된 유전자에 의한 결합있는 벡터의 보안을 포함하는 접근 방식을 활용하는 적어도 6가지의 바이러스 벡터 접근 방식이 임상 시험에서 유전자 전이를 위해 현재 이용 가능하다.
- [0170] pLASN 및 MFG-S가 임상 시험에 사용되었던 레트로바이러스 벡터의 예이다 (Dunbar et al., *Blood* 85:3048-305 (1995); Kohn et al., *Nat. Med.* 1:1017-102 (1995); Malech et al., *PNAS USA* 94:22 12133-12138 (1997)). PA317/pLASN은 유전자 요법 시험에 사용된 첫 번째 치료용 벡터였다 (Blaese et al., *Science* 270:475-480 (1995)). 50% 이상의 형질도입 효율이 MFG-S 패키징된 벡터에 대하여 관찰되었다 (Ellem et al., *Immunol Immunother.* 44(1):10-20 (1997); Dranoff et al., *Hum. Gene Ther.* 1:111-2 (1997)).
- [0171] 재조합 아데노 관련 바이러스 벡터 (rAAV)는 결합있는 비병원성의 파르보바이러스 아데노 관련 유형 2 바이러스에 근거한 유망한 대체 유전자 전달 시스템이다. 모든 벡터는 트랜스진 발현 카세트를 플랭킹하는 AAV 145 bp 역위 말단 반복 서열만을 보유하고 있는 플라스미드로부터 유래된다. 형질도입된 세포의 계능 내로의 통합에 기인한 효율적인 유전자 전이 및 안정적인 트랜스진 전달이 상기 벡터 시스템에 대한 중요한 특징이다 (Wagner et al., *Lancet* 351:9117 1702-3 (1998), Kearns et al., *Gene Ther.* 9:748-55 (1996)). 다른 AAV 혈청형, 예를 들어 AAV1, AAV3, AAV4, AAV5, AAV6, AAV8, AAV8.2, AAV9 및 AAVrh10 및 가짜형 AAV, 예컨대 AAV2/8, AAV2/5 및 AAV2/6이 또한, 본 발명에 따라서 사용될 수 있다.
- [0172] 복제-결핍된 재조합 아데노바이러스 벡터 (Ad)는 높은 역가로 생성될 수 있고 수많은 상이한 세포 유형을 용이하게 감염시킬 수 있다. 대부분의 아데노바이러스 벡터는 트랜스진이 Ad E1a, E1b, 및/또는 E3 유전자를 대체하도록 조작되고; 연속해서 상기 복제 결합있는 벡터는 트랜스로 결실된 유전자 기능을 공급하는 인간 293 세포에서 번식된다. Ad 벡터는 여러 가지 유형의 조직을 생체내에서 형질도입시킬 수 있으며, 이는 비-분할, 분화된 세포, 예컨대 간, 신장 및 근육에서 발견된 것을 포함한다. 통상적인 Ad 벡터는 큰 운반 능력을 가지고 있다. 임상 시험에서 Ad 벡터를 사용하는 것의 한 예는 근육내 주사를 이용하여 항종양 면역을 위한 폴리뉴클레오티드 요법을 포함하였다 (Sterman et al., *Hum. Gene Ther.* 7:1083-9 (1998)). 임상 시험에서 유전자 전이를 위하여 아데노바이러스 벡터를 사용하는 것의 부가 예는 하기 문헌을 포함한다 [Rosenecker et al., *Infection* 24:1 5-10 (1996); Sterman et al., *Hum. Gene Ther.* 9:7 1083-1089 (1998); Welsh et al., *Hum. Gene Ther.* 2:205-18 (1995); Alvarez et al., *Hum. Gene Ther.* 5:597-613 (1997); Topf et al., *Gene Ther.* 5:507-513 (1998); Sterman et al., *Hum. Gene Ther.* 7:1083-1089 (1998)].
- [0173] 패키징 세포를 사용하여 숙주 세포를 감염시킬 수 있는 바이러스 입자를 형성시킨다. 이러한 세포는 아데노바이러스 및 AAV를 패키징하는 293 세포, 및 레트로바이러스를 패키징하는 ψ 2 세포 또는 PA317 세포를 포함한다. 유전자 요법에 사용된 바이러스 벡터는 통상적으로, 핵산 벡터를 바이러스 입자 내로 패키징하는 생산자 세포주에 의해 생성된다. 이러한 벡터는 전형적으로, 패키징 및 숙주 내로의 후속 통합 (적용가능한 경우)에 필요한 최소한의 바이러스 서열을 함유하고, 다른 바이러스 서열은 발현될 단백질을 코딩하는 발현 카세트에 의해 대체된다. 누락된 바이러스 기능은 패키징 세포주에 의해 트랜스로 공급된다. 예를 들어, 유전자 요법에 사용된 AAV 벡터는 전형적으로, 패키징 및 숙주 계능 내로의 통합에 필요한 AAV 계능으로부터의 역위 말단 반복 (ITR) 서열만을 보유한다. 바이러스 DNA는 다른 AAV 유전자, 즉 rep 및 cap를 코딩하는 헬퍼 플라스미드를 함유하지만, ITR 서열이 결여된 세포주에서 패키징된다. 이러한 세포주는 또한, 헬퍼로서 아데노바이러스로 감염된다. 헬퍼 바이러스는 AAV 벡터의 복제, 및 헬퍼 플라스미드로부터의 AAV 유전자의 발현을 증진시킨다. 헬퍼 플라스미드는 ITR 서열의 결여로 인해 상당한 양으로 패키징되지 않는다. 아데노바이러스로의 오염은, 예를 들어, 아데노바이러스가 AAV보다 더 감수성인 열 처리에 의해 감소될 수 있다. 또한, AAV는 바콜로바이러스 시스템을 사용하여 제조될 수 있다 (예를 들어, 미국 특허 6,723,551 및 7,271,002 참조).
- [0174] 293 또는 바콜로바이러스 시스템으로부터의 AAV 입자의 정제는 전형적으로, 상기 바이러스를 생산하는 세포를 성장시킨 다음, 세포 상등액으로부터 바이러스 입자를 수집하거나 또는 세포를 용해시키고 조약한 용해물로부터 바이러스를 수집하는 것을 포함한다. 그런 다음, AAV는 이온 교환 크로마토그래피 (예를 들어, 미국 특허 7,419,817 및 6,989,264 참조), 이온 교환 크로마토그래피 및 CsCl 밀도 원심분리 (예를 들어, PCT 공개 W02011094198A10), 면역친화 크로마토그래피 (예를 들어 W02016128408) 또는 AVB 세파로스 (예를 들어, GE 헬스케어 라이프 사이언스(GE Healthcare Life Sciences))를 이용하는 정제를 포함한, 관련 기술분야에 공지된 방법에 의해 정제된다.
- [0175] 많은 유전자 요법 적용에서는, 유전자 요법 벡터가 고도의 특이성으로 특정한 조직 유형에 전달되는 것이 바람직하다. 따라서, 바이러스 벡터는 바이러스의 외부 표면 상에 바이러스 코트 단백질을 수반한 융합 단백질로서

리간드를 발현시킴으로써 소정의 세포 유형에 대한 특이성을 갖도록 변형될 수 있다. 관심 세포 유형 상에 존재하는 것으로 공지된 수용체에 대한 친화성을 갖는 리간드가 선택된다. 예를 들어, 문헌 [Han et al., *Proc. Natl. Acad. Sci. USA* 92:9747-9751 (1995)]에는 gp70과 융합된 인간 헤레굴린을 발현하도록 물로니 뮌 백혈병 바이러스를 변형시킬 수 있고, 재조합 바이러스가 인간 표피 성장 인자 수용체를 발현하는 특성의 인간 유방암 세포를 감염시킨다고 보고되었다. 이러한 원리는 표적 세포가 수용체를 발현하고 바이러스가 세포 표면 수용체에 대한 리간드를 포함하는 융합 단백질을 발현하는 다른 바이러스-표적 세포 쌍으로 확장될 수 있다. 예를 들어, 섬유상 파지는 사실상 임의의 선택된 세포성 수용체에 대한 특이적 결합 친화성을 갖는 항체 단편 (예를 들어, Fab 또는 Fv)을 표시하도록 조작될 수 있다. 상기 설명이 주로 바이러스 벡터에 적용되긴 하지만, 동일한 원리가 비-바이러스 벡터에도 적용될 수 있다. 이러한 벡터는 특이적 표적 세포에 의한 흡수를 선호하는 특이적 흡수 서열을 함유하도록 조작될 수 있다.

[0176] 유전자 요법 벡터는, 전형적으로 아래에 기재되는 바와 같은, 전신 투여 (예를 들어, 정맥내, 복강내, 근육내, 피하 또는 두개내 주입) 또는 국소 적용에 의해 개별 환자에게 투여함으로써 생체내 전달될 수 있다. 대안적으로, 벡터는 생체의 세포, 예컨대 개별 환자로부터 체외 이식된 세포 (예를 들어, 림프구, 골수 흡인물, 조직 생검) 또는 만능 공여자 조혈 줄기 세포에 전달한 다음, 통상적으로 벡터가 혼입된 세포에 대한 선택 후에, 상기 세포를 환자 내로 다시 체내 이식할 수 있다.

[0177] 진단, 연구, 이식을 위한 생체의 세포 형질감염 또는 (예를 들어, 형질감염된 세포의 숙주 유기체 내로의 재-주입을 통하여) 유전자 요법하기 위한 생체의 세포 형질감염은 관련 기술분야의 통상의 기술자에게 널리 공지되어 있다. 바람직한 실시양태에서는, 세포를 대상 유기체로부터 분리하고, DNA-결합 단백질 핵산 (유전자 또는 cDNA)으로 형질감염시킨 다음, 이를 대상 유기체 (예를 들어, 환자)에게 다시 재주입한다. 생체의 형질감염에 적합한 다양한 세포 유형은 관련 기술분야의 통상의 기술자에게 널리 공지되어 있고 (예를 들어, 문헌 [Freshney et al., *Culture of Animal Cells, A Manual of Basic Technique* (3rd ed. 1994)] 참조), 환자로부터 세포를 분리하고 배양하는 방법에 관한 논의에 대해서는 상기 문헌에 인용된 참고문헌을 참조할 수 있다.

[0178] 한 실시양태에서, 세포 형질감염 및 유전자 요법을 위한 생체의 절차에 줄기 세포가 사용된다. 줄기 세포를 사용하는 경우의 이점은 이들이 시험관 내에서 다른 세포 유형으로 분화될 수 있거나, 또는 이들이 골수에 생착될 포유류 (예컨대, 세포의 공여자) 내로 도입될 수 있다는 것이다. 시토키인, 예컨대 GM-CSF, IFN- γ 및 TNF- α 를 사용하여 시험관 내에서 CD34+ 세포를 임상적으로 중요한 면역 세포 유형으로 분화시키는 방법이 공지되어 있다 (문헌 [Inaba et al., *J. Exp. Med.* 176:1693-1702 (1992)] 참조).

[0179] 줄기 세포는 공지된 방법을 사용하여 형질도입 및 분화시키기 위해 분리된다. 예를 들어, 줄기 세포는 원치 않는 세포, 예컨대 CD4+ 및 CD8+ (T 세포), CD45+ (panB 세포), GR-1 (과립구), 및 Iad (분화된 항원 제시 세포)에 결합하는 항체로 골수 세포를 패닝함으로써 골수 세포로부터 분리된다 (문헌 [Inaba et al., *J. Exp. Med.* 176:1693-1702 (1992)] 참조).

[0180] 일부 실시양태에서는 변형시켰던 줄기 세포를 사용할 수도 있다. 예를 들어, 아포토시스에 대해 저항성이 되도록 만든 뉴런의 줄기 세포가 치료 조성물로서 사용될 수 있으며, 여기서 상기 줄기 세포는 또한, 본 발명의 ZFP TF를 함유한다. 아포토시스에 대한 저항성은, 예를 들어, 줄기 세포 내에서 BAX- 또는 BAK-특이적 ZFN (미국 특허 출원 번호 12/456,043 참조)을 사용하여 BAX 및/또는 BAK를 녹아웃시킴으로써, 또는 예를 들어, 카스파제-6 특이적 ZFN을 또한 사용하여 카스파제에서 붕괴되는 것에 의해 발생될 수 있다. 이들 세포는 TCR을 조절하는 것으로 공지되어 있는 ZFP TF로 형질감염시킬 수 있다.

[0181] 치료용 DNA-결합 단백질 (또는 이들 단백질을 코딩하는 핵산)을 함유하는 벡터 (예를 들어, 레트로바이러스, 아데노바이러스, 리포솜 등)는 또한, 생체내에서 세포의 형질도입을 위하여 유기체에 직접적으로 투여될 수 있다. 대안적으로, 네이티브 DNA를 투여할 수 있다. 투여는 주사, 주입, 국소 적용 및 전기천공을 포함하지만, 이에 제한되지 않는, 분자를 혈액 또는 조직 세포와 궁극적으로 접촉시켜 이들 내로 도입하는데 정상적으로 사용되는 어떠한 경로에 의해서도 수행된다. 이러한 핵산의 적합한 투여 방법은 관련 기술분야의 통상의 기술자에게 이용가능하고 널리 공지되어 있으며, 2가지 이상의 경로를 사용하여 특정한 조성물을 투여할 수 있긴 하지만, 특정한 경로가 종종, 또 다른 경로보다 더 즉각적이고 더 유효한 반응을 제공할 수 있다.

[0182] DNA를 조혈 줄기 세포 내로 도입하는 방법이, 예를 들어, 미국 특허 번호 5,928,638에 개시된다. 트랜스젠을 조혈 줄기 세포, 예를 들어, CD34+ 세포 내로 도입하는데 유용한 벡터는 아데노바이러스 유형 35를 포함한다.

[0183] 트랜스젠을 면역 세포 (예를 들어, T-세포) 내로 도입하는데 적합한 벡터는 비-통합 렌티바이러스 벡터를 포함

한다. 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] 참조.

[0184] 제약상 허용되는 담체는 부분적으로, 투여되는 특정한 조성물에 의해서 결정될 뿐만 아니라 이러한 조성물을 투여하는데 사용되는 특정한 방법에 의해 결정된다. 따라서, 아래에 기재되는 바와 같이, 제약 조성물의 광범위한 적합한 제형이 이용가능하다 (예를 들어, 문헌 [Remington's Pharmaceutical Sciences, 17th ed., 1989] 참조).

[0185] 상기 기재된 바와 같이, 본원에 개시된 방법 및 조성물은 원핵 세포, 진균 세포, 고세균 세포, 식물 세포, 곤충 세포, 동물 세포, 척추 동물 세포, 포유류 세포 및 인간 세포를 포함하지만, 이에 제한되지 않는 임의의 유형의 세포 (임의의 유형의 줄기 세포 및 T-세포 포함)에 사용될 수 있다. 단백질 발현에 적합한 세포주는 관련 기술 분야의 통상의 기술자에게 공지되어 있고, 이는 COS, CHO (예를 들어, CHO-S, CHO-K1, CHO-DG44, CHO-DUXB11), VERO, MDCK, WI38, V79, B14AF28-G3, BHK, HaK, NSO, SP2/0-Ag14, HeLa, HEK293 (예를 들어, HEK293-F, HEK293-H, HEK293-T), perC6, 곤충 세포, 예컨대 스포도페라 푸기페르다 (Sf), 및 진균 세포, 예컨대 사카로미세스, 피키아 및 쉬조사카로미세스를 포함하지만, 이에 제한되지 않는다. 이들 세포주의 자손, 변이체 및 유도체가 사용될 수도 있다.

[0186] **적용**

[0187] 본원에 개시된 조성물 및 방법은 B2M 조정이 바람직한 치료 및 연구 적용 분야를 포함하지만 이에 제한되지 않는, B2M 발현 및/또는 기능성을 조정하는 것이 요망되는 임의의 적용 분야에 사용될 수 있다. 예를 들어, 본원에 개시된 조성물은 하나 이상의 외인성 CAR, 외인성 TCR, 외인성 ACTR 또는 다른 암-특이적 수용체 분자를 발현하는 입양 세포 요법을 위해 변형된 T 세포에서의 내인성 B2M의 발현을 붕괴시킴으로써, 암을 치료 및/또는 예방하기 위해 생체내 및/또는 생체외 (세포 요법)에서 사용될 수 있다. 또한, 이러한 환경에서는, 특정 세포 내에서의 B2M 발현을 조정하는 것이, 건강하고 비표적화된 조직과의 원치않는 교차 반응 (즉, 이식편 대 숙주 반응)의 위험을 없애거나 또는 실질적으로 감소시킬 수 있다.

[0188] 방법 및 조성물은 또한, 줄기 세포 내에서의 B2M 유전자를 조정시킨 (변형시킨) 줄기 세포 조성물을 포함하고, 상기 세포는 ACTR 및/또는 CAR 및/또는 단리된 또는 조작된 TCR을 추가로 포함한다. 예를 들어, B2M 녹아웃 또는 녹다운 조정된 동종이계 조혈 줄기 세포는 골수 제거 후에 HLA 일치하지 않는 환자 내로 도입될 수 있다. 이들 변형된 HSC는 환자의 재-집락 형성을 허용할 것이지만, 잠재적인 GvHD를 유발하지는 않을 것이다. 이와 같이 도입된 세포는 또한, 기저 질환을 치료하기 위한 후속 요법 (예를 들어, 화학요법 내성) 동안 도움을 주는 다른 변형을 가질 수 있다. HLA 널 세포는 또한, 외상 환자를 둔 응급실 상황에서 "쉽게 사용할 수 있는" 요법으로서 사용된다.

[0189] 본 발명의 방법 및 조성물은 또한, 관련 장애의 연구를 허용해주는 시험관내 및 생체내 모델, 예를 들어, B2M 및/또는 HLA 및 관련 장애의 동물 모델의 디자인 및 이행에 유용하다.

[0190] 본원에 언급된 모든 특허, 특허 출원 및 공개는 그 전문이 본원에 참조로 포함된다.

[0191] 비록 명확성 및 이해의 목적을 위한 예 및 한 예시로서 개시내용이 일부 상세히 제공되었지만, 본 개시내용의 요지 또는 범위를 벗어나지 않고서도 다양한 변화 및 변형이 실시될 수 있다는 것이 관련 기술분야의 통상의 기술자에게 명백할 것이다. 따라서, 전술된 개시내용 및 하기 실시예는 제한되는 것으로 간주되지 않아야 한다.

[0192] 실시예

[0193] **실시예 1: B2M-특이적 뉴클레아제의 디자인**

[0194] B2M 유전자에 이중 가닥 파단의 부위 특이적 도입을 가능하게 하는 B2M-특이적 ZFN을 구축하였다. ZFN은 문헌 [Urnov et al. (2005) *Nature* 435(7042):646-651, Lombardo et al. (2007) *Nat Biotechnol.* Nov;25(11):1298-306], 및 미국 특허 공개 2008-0131962, 2015-016495, 2014-0120622, 2014-0301990 및 미국 특허 8,956,828에 본질적으로 기재된 바와 같이 디자인하였다. ZFN 쌍은 B2M 유전자의 불변 영역 내의 상이한 부위를 표적화하였다 (도 1 참조). 예시적인 ZFN 쌍에 대한 인식 나선 뿐만 아니라 표적 서열이 아래 표 1에 제시된다. B2M 징크 핑거 디자인의 표적 부위가 첫 번째 칼럼에 제시된다. ZFP 인식 나선에 의해 표적화되는 표적 부위 내의 뉴클레오티드가 대문자로 표시되고; 비-표적화된 뉴클레오티드가 소문자로 표시된다. FokI 뉴클레아제 도메인과 ZFP DNA 결합 도메인을 연결하기 위해 사용된 링커가 또한 제시된다 (미국 특허 공개 20150132269 참조). 예를

들어, 도메인 링커 L0의 아미노산 서열은 DNA 결합 도메인-QLVKS-FokI 뉴클레아제 도메인 (서열식별번호: 3)이다. 유사하게, 도메인 링커 N7a에 대한 아미노산 서열은 FokI 뉴클레아제 도메인-SGTPHEVGVTIL-DNA 결합 도메인 (서열식별번호: 4)이고, N6a는 FokI 뉴클레아제 도메인-SGAQGSTLDF-DNA 결합 도메인 (서열식별번호: 5)이다.

표 1: B2M 징크 핑거 디자인

ZFN 명칭 표적 서열	F1	F2	F3	F4	F5	F6	도메인 링커
SBS57071 5' gcCACGGA gCGAGACA TCTCGgcc cgaa (SEQ ID NO:6)	RSDDLK (SEQ ID NO:49)	DSSARKK (SEQ ID NO:50)	DRSNLSR (SEQ ID NO:51)	QRTHLRD (SEQ ID NO:52)	QSGHLAR (SEQ ID NO:53)	DSSNREA (SEQ ID NO:54)	L0
SBS57531 5'gaGTAG CGcGAGCA CAGCtaag gccacg (SEQ ID NO:7)	AQCCLFH (SEQ ID NO:55)	DQSNLRA (SEQ ID NO:56)	RSANLTR (SEQ ID NO:57)	RSDDLTR (SEQ ID NO:58)	QSGSLTR (SEQ ID NO:59)	NA	N6a
SBS57362 5'tcCAGC AGAGAATG GAAAGTca aatctc (SEQ ID NO:8)	LNHHLQQ (SEQ ID NO:60)	QSGNLAR (SEQ ID NO:61)	RSDTLA (SEQ ID NO:62)	QNAHRKT (SEQ ID NO:63)	RSDNLSE (SEQ ID NO:64)	KPYNLRT (SEQ ID NO:65)	N6a
SBS57376 5'ttTCCT GAATTGCT ATGTGTct gggttt (SEQ ID NO:9)	TRDHLST (SEQ ID NO:66)	RSDARTN (SEQ ID NO:67)	QSSDLR (SEQ ID NO:68)	HRSSLKN (SEQ ID NO:69)	QSSHLTR (SEQ ID NO:70)	DSSDRKK (SEQ ID NO:71)	L0
SBS57017 5'tgTCGG ATgGATGA AACCCAGa cacata (SEQ ID NO:10)	RSDNLSE (SEQ ID NO:64)	ASKTRTN (SEQ ID NO:72)	QSGNLAR (SEQ ID NO:61)	TSGNLTR (SEQ ID NO:73)	TSGNLTR (SEQ ID NO:73)	RIQDLNK (SEQ ID NO:74)	N7a
SBS57327 5' taGCAATT CAGGAaT TTGACTtt ccat (SEQ ID NO:11)	DRSNLSR (SEQ ID NO:51)	ARWYLDK (SEQ ID NO:75)	QSGNLAR (SEQ ID NO:61)	AKWNLDA (SEQ ID NO:76)	QQHVLQN (SEQ ID NO:77)	QNATRTK (SEQ ID NO:78)	L0
SBS57328 5'taGCAA TTCAGGAA ATTtgact	TNQLHW (SEQ ID NO:79)	QSGNLAR (SEQ ID NO:61)	RSDNLRE (SEQ ID NO:80)	ASHVLNA (SEQ ID NO:81)	QNATRTK (SEQ ID NO:78)	NA	L0

ttccat (SEQ ID NO:11)							
SBS57332 5'tgTCGG ATgGATGA AACCCAGa cacata (SEQ ID NO:10)	RSDNLSE (SEQ ID NO:64)	ASKTRTN (SEQ ID NO:72)	QSGNLAR (SEQ ID NO:61)	TSANLSR (SEQ ID NO:82)	TSGNLTR (SEQ ID NO:73)	RTEDRLA (SEQ ID NO:83)	N6a
SBS57469 5'tgTCGG ATGGATGA aACCCAGa cacata (SEQ ID NO:10)	RSDNLSE (SEQ ID NO:64)	ASKTRTN (SEQ ID NO:72)	YTSSLCY (SEQ ID NO:84)	QSGHLSR (SEQ ID NO:85)	TSGNLTR (SEQ ID NO:73)	RIQDLNK (SEQ ID NO:74)	N7a
SBS57331 5'tgTCGG ATgGATGA AACCCAGa cacata (SEQ ID NO:10)	RSDNLSE (SEQ ID NO:64)	ASKTRKN (SEQ ID NO:86)	QSGNLAR (SEQ ID NO:61)	TSANLSR (SEQ ID NO:82)	TSGNLTR (SEQ ID NO:73)	RIQDLNK (SEQ ID NO:74)	N6a
SBS57326 5'taGCAA TTCAGGAA aTTTGACT ttccat (SEQ ID NO:11)	DRSNLSR (SEQ ID NO:51)	ARWYLDK (SEQ ID NO:75)	QSGNLAR (SEQ ID NO:61)	AKWNLDA (SEQ ID NO:76)	TTPVLVQ (SEQ ID NO:87)	QNATRTK (SEQ ID NO:78)	L0
SBS55822 5'caTCCG ACATTGAA GTTGACTt actgaa (SEQ ID NO:12)	DRSNLSR (SEQ ID NO:51)	FPGSRTR (SEQ ID NO:88)	QSGNLAR (SEQ ID NO:61)	WRISLAA (SEQ ID NO:89)	DRSNLSR (SEQ ID NO:51)	DSSDRKK (SEQ ID NO:71)	N7a
SBS57511 5'gaAGAA TGGAGAGA GAATTGaa aaagtg (SEQ ID NO:13)	DQSLLR (SEQ ID NO:90)	QSGNLAR (SEQ ID NO:61)	HLGLRD (SEQ ID NO:91)	RSANLTR (SEQ ID NO:57)	RSDVLST (SEQ ID NO:92)	QNAHRIK (SEQ ID NO:93)	L0
SBS57509 5'gaAGAA TGGAGAGA GAATTGaa aaagtg (SEQ ID NO:13)	DQSLLR (SEQ ID NO:90)	QSGNLAR (SEQ ID NO:61)	QSAHRKN (SEQ ID NO:94)	RSANLTR (SEQ ID NO:57)	RSDVLST (SEQ ID NO:92)	QNAHRIK (SEQ ID NO:93)	L0
SBS57482 5'caTCCG ACATTGAA GTTGACTt actgaa	DRSNLSR (SEQ ID NO:51)	FPGSRTR (SEQ ID NO:88)	QSGNLAR (SEQ ID NO:61)	HKLSLSI (SEQ ID NO:95)	DRSNLSR (SEQ ID NO:51)	DSSDRKK (SEQ ID NO:71)	N7a

[0197]

(SEQ ID NO:12)							
SBS57347 5'gaAGAA TGGAGAGA GAATTGaa aaagtg (SEQ ID NO:13)	DQSLIRT (SEQ ID NO:90)	QSGNLAR (SEQ ID NO:61)	QNAHRKT (SEQ ID NO:63)	RSANLTR (SEQ ID NO:57)	RSDVLST (SEQ ID NO:92)	QNAHRIK (SEQ ID NO:93)	L0
SBS57296 5'ctGAAG AATGGAGA GAGaattg aaaaag (SEQ ID NO:14)	RSANLTR (SEQ ID NO:57)	QSAHRKN (SEQ ID NO:94)	RHSHLTS (SEQ ID NO:96)	QSGNLAR (SEQ ID NO:61)	QSNQLAV (SEQ ID NO:97)	NA	N7a
SBS57322 5'aaAAG TGGAGCAT TCAGActt gtcttt (SEQ ID NO:15)	DRSNLSR (SEQ ID NO:51)	QSADRTK (SEQ ID NO:98)	TNQNRT (SEQ ID NO:99)	RSANLTR (SEQ ID NO:57)	RSDSLSV (SEQ ID NO:100)	QANRKT (SEQ ID NO:101)	L0
SBS57323 5' aaAAGTG GAGCATTC AGActtgt cttt (SEQ ID NO:15)	DRSNLSR (SEQ ID NO:51)	QSADRTK (SEQ ID NO:98)	LKQNLDA (SEQ ID NO:103)	RSANLTR (SEQ ID NO:57)	RSDSLSV (SEQ ID NO:100)	QANRKT (SEQ ID NO:101)	L0
SBS57447 5'ctGAAG AATGGAGA GAGaattg aaaaag (SEQ ID NO:14)	RSANLTR (SEQ ID NO:57)	QSAHRKN (SEQ ID NO:94)	RHSHLTS (SEQ ID NO:96)	QSGNLAR (SEQ ID NO:61)	QRGNLWT (SEQ ID NO:102)	NA	N7a

[0198]

[0199]

모든 ZFN을 시험하였으며, 이는 그의 표적 부위에 결합하는 것으로 밝혀졌으며, 뉴클레아제로서 활성인 것으로 밝혀졌다.

[0200]

에스. 피오게네스(*S. pyogenes*) CRISPR/Cas9 시스템에 대한 가이드 RNA가 또한, B2M 유전자를 표적화하도록 구축되었다. B2M 유전자 내의 표적 서열 뿐만 아니라 가이드 RNA 서열이 아래 표 2A에 표시된다. 모든 가이드 RNA를 CRISPR/Cas9 시스템에서 시험하며, 이는 활성인 것으로 밝혀졌다. 가이드 서열 중 일부의 5' 말단에서의 소문자 "g"는 PAM 서열에서 작용하는 부가된 G 뉴클레오티드를 표시한다.

[0201] 표 2A: 인간 B2M의 불변 영역에 대한 가이드 RNA

명칭	가닥	표적 (5' → 3')	gRNA (5' → 3')
B2M-Gr1073	f	GGCCGAGATGTCTCGCTCCGTGG (SEQ ID NO:16)	GGCCGAGATGTCTCGCTCCG (SEQ ID NO:104)
B2M-Gr1074	r	CGCGAGCACAGCTAAGGCCACGG (SEQ ID NO:17)	gCGCGAGCACAGCTAAGGCCA (SEQ ID NO:105)
B2M-Gr1080	r	GAGTAGCGCGAGCACAGCTAAGG (SEQ ID NO:18)	GAGTAGCGCGAGCACAGCTA (SEQ ID NO:106)
B2M-Gr1107	f	CTCGCGCTACTCTCTCTTTCTGG (SEQ ID NO:19)	gCTCGCGCTACTCTCTCTTTC (SEQ ID NO:107)
B2M-Gr1112	f	GCTACTCTCTCTTCTGGCCTGG (SEQ ID NO:20)	GCTACTCTCTCTTCTGGCC (SEQ ID NO:108)
B2M-Gr1115	f	ACTCTCTCTTTCTGGCCTGGAGG (SEQ ID NO:21)	gACTCTCTCTTCTGGCCTGG (SEQ ID NO:109)
B2M-Gr1114	r	ACTCACGCTGGATAGCCTCCAGG (SEQ ID NO:22)	gACTCACGCTGGATAGCCTCC (SEQ ID NO:110)
B2M-Gr1126	r	AGGGTAGGAGAGACTCACGCTGG (SEQ ID NO:23)	gAGGGTAGGAGAGACTCACGC (SEQ ID NO:111)
B2M-Gr4942	r	CGTGAGTAAACCTGAATCTTTGG (SEQ ID NO:24)	gCGTGAGTAAACCTGAATCTT (SEQ ID NO:112)
B2M-Gr4948	f	CTCAGGTACTCCAAAGATTCAGG (SEQ ID NO:25)	gCTCAGGTACTCCAAAGATTC (SEQ ID NO:113)
B2M-Gr4969	r	TTTGACTTTCCATTCTCTGCTGG (SEQ ID NO:26)	gTTTGACTTTCCATTCTCTGC (SEQ ID NO:114)
B2M-Gr4976	f	TCACGTCATCCAGCAGAGAATGG (SEQ ID NO:27)	gTCACGTCATCCAGCAGAGAA (SEQ ID NO:115)
B2M-Gr4995	r	ACCCAGACACATAGCAATTCAGG (SEQ ID NO:28)	gACCCAGACACATAGCAATTC (SEQ ID NO:116)
B2M-Gr5009	f	TTCCTGAATTGCTATGTGTCTGG (SEQ ID NO:29)	gTTCCTGAATTGCTATGTGTC (SEQ ID NO:117)
B2M-Gr5010	f	TCCTGAATTGCTATGTGTCTGGG (SEQ ID NO:30)	gTCCTGAATTGCTATGTGTCT (SEQ ID NO:118)
B2M-Gr5023	r	AAGTCAACTTCAATGTCGGATGG (SEQ ID NO:31)	gAAGTCAACTTCAATGTCGGA (SEQ ID NO:119)
B2M-Gr5027	r	CAGTAAGTCAACTTCAATGTCGG (SEQ ID NO:32)	gCAGTAAGTCAACTTCAATGT (SEQ ID NO:120)
B2M-Gr5051	f	GAAGTTGACTTACTGAAGAATGG (SEQ ID NO:33)	GAAGTTGACTTACTGAAGAA (SEQ ID NO:121)
B2M-Gr5071	f	TGGAGAGAGAATTGAAAAAGTGG (SEQ ID NO:34)	gTGGAGAGAGAATTGAAAAAG (SEQ ID NO:122)
B2M-Gr5098	f	TTCAGACTTGTCTTTCAGCAAGG (SEQ ID NO:35)	gTTCAGACTTGTCTTTCAGCA (SEQ ID NO:123)
B2M-Gr5103	f	ACTTGTCTTTCAGCAAGGACTGG (SEQ ID NO:36)	gACTTGTCTTTCAGCAAGGAC (SEQ ID NO:124)
B2M-Gr5141	r	ATACTCATCTTTTTCAGTGGGGG (SEQ ID NO:37)	gATACTCATCTTTTTCAGTGG (SEQ ID NO:125)
B2M-Gr5142	r	CATACTCATCTTTTTCAGTGGGG (SEQ ID NO:38)	gCATACTCATCTTTTTCAGTG (SEQ ID NO:126)
B2M-Gr5143	r	GCATACTCATCTTTTTCAGTGGG (SEQ ID NO:39)	GCATACTCATCTTTTTCAGT (SEQ ID NO:127)
B2M-Gr5144	r	GGCATACTCATCTTTTTCAGTGG (SEQ ID NO:40)	GGCATACTCATCTTTTTCAG (SEQ ID NO:128)

[0202]

B2M-Gr5165	r	AGTCACATGGTTCACACGGCAGG (SEQ ID NO:41)	gAGTCACATGGTTCACACGGC (SEQ ID NO:129)
B2M-Gr5169	r	ACAAAGTCACATGGTTCACACGG (SEQ ID NO:42)	gACAAAGTCACATGGTTCACA (SEQ ID NO:130)
B2M-Gr5178	r	TGGGCTGTGACAAAGTCACATGG (SEQ ID NO:43)	gTGGGCTGTGACAAAGTCACA (SEQ ID NO:131)
B2M-Gr5197	r	TTACCCCACTTAACTATCTTGGG (SEQ ID NO:44)	gTTACCCCACTTAACTATCTT (SEQ ID NO:132)
B2M-Gr5198	r	CTTACCCCACTTAACTATCTTGG (SEQ ID NO:45)	gCTTACCCCACTTAACTATCT (SEQ ID NO:133)
B2M-Gr5208	f	CACAGCCCAAGATAGTTAAGTGG (SEQ ID NO:46)	gCACAGCCCAAGATAGTTAAG (SEQ ID NO:134)
B2M-Gr5209	f	ACAGCCCAAGATAGTTAAGTGGG (SEQ ID NO:47)	gACAGCCCAAGATAGTTAAGT (SEQ ID NO:135)
B2M-Gr5210	f	CAGCCCAAGATAGTTAAGTGGGG (SEQ ID NO:48)	gCAGCCCAAGATAGTTAAGTG (SEQ ID NO:136)

[0203]

[0204] TALEN은 B2M 로커스를 표적화하도록 만들었고, 이는 아래 표 2B에 제시된다. 모든 TALEN을 K562 세포에서 시험하였으며, 이는 활성인 것으로 밝혀졌다(표 2C 및 도 2b 참조).

[0205] 표 2B: B2M에 대해 특이적인 TALEN

SBS#	표적 부위 5' -> 3'	RVD N -> C
103049	atTCGGGCCGAGATGTCTCgc (SEQ ID NO:137)	NG-HD-NN-NN-NN-HD-HD-NN-NI-NN-NI-NG-NN-NG-HD-NG-HD
103050	gtAGCGCGAGCACAGCTAAgg (SEQ ID NO:138)	NI-NN-HD-NN-HD-NN-NI-NN-HD-NI-HD-NI-NN-HD-NG-NI-NI
103051	ctCCGTGGCCTTAGCTGTGct (SEQ ID NO:139)	HD-HD-NN-NG-NN-NN-HD-HD-NG-NG-NI-NN-HD-NG-NN-NG-NK
103052	ctCCAGGCCAGAAAGAGAgag (SEQ ID NO:140)	HD-HD-NI-NN-NN-HD-HD-NI-NN-NI-NI-NI-NN-NI-NN-NI-NK
103053	gtGGCCTTAGCTGTGCTCGcg (SEQ ID NO:141)	NN-NN-HD-HD-NG-NG-NI-NN-HD-NG-NN-NG-NN-HD-NG-HD-NK
103054	atAGCCTCCAGGCCAGAAAgag (SEQ ID NO:142)	NI-NN-HD-HD-NG-HD-HD-NI-NN-NN-HD-HD-NI-NN-NI-NI-NI
103055	ctTAGCTGTGCTCGCGCTAct (SEQ ID NO:143)	NG-NI-NN-HD-NG-NN-NG-NN-HD-NG-HD-NN-HD-NN-HD-NG-NI
103056	ctGGATAGCCTCCAGGCCAgag (SEQ ID NO:144)	NN-NN-NI-NG-NI-NN-HD-HD-NG-HD-HD-NI-NN-NN-HD-HD-NI
103057	ctGTGCTCGCGCTACTCTct (SEQ ID NO:145)	NN-NG-NN-HD-NG-HD-NN-HD-NN-HD-NG-NI-HD-NG-HD-NG-HD
103058	ctCACGCTGGATAGCCTCCag (SEQ ID NO:146)	HD-NI-HD-NN-HD-NG-NN-NN-NI-NG-NI-NN-HD-HD-NG-HD-HD
103059	ctACTCTCTCTTTCTGGCctg (SEQ ID NO:147)	NI-HD-NG-HD-NG-HD-NG-HD-NG-NG-NG-HD-NG-NN-NN-HD-HD
103060	gtAGGAGAGACTCACGCTGga (SEQ ID NO:148)	NI-NN-NN-NI-NN-NI-NN-NI-HD-NG-HD-NI-HD-NN-HD-NG-NK
103061	gtGTCTTTTCCCGATATTCct (SEQ ID NO:149)	NN-NG-HD-NG-NG-NG-NG-HD-HD-HD-NN-NI-NG-NI-NG-NG-HD
103062	gtGAGTAAACCTGAATCTTtg (SEQ ID NO:150)	NN-NI-NN-NG-NI-NI-NI-HD-HD-NG-NN-NI-NI-NG-HD-NG-NG
103063	ttTTCCCAGATATTCCTCAGgt (SEQ ID NO:151)	NG-NG-HD-HD-HD-NN-NI-NG-NI-NG-NG-HD-HD-NG-HD-NI-NK
103064	atGACGTGAGTAAACCTGAat (SEQ ID NO:152)	NN-NI-HD-NN-NG-NN-NI-NN-NG-NI-NI-NI-HD-HD-NG-NN-NI
103065	ttCCTCAGGTACTCCAAAGat (SEQ ID NO:153)	HD-HD-NG-HD-NI-NN-NN-NG-NI-HD-NG-HD-HD-NI-NI-NI-NK

[0206]

103066	ctCTGCTGGATGACGTGAGta (SEQ ID NO:154)	HD-NG-NN-HD-NG-NN-NN-NI-NG-NN-NI-HD-NN-NG-NN-NI-NK
103067	ctCAGGTACTCCAAAGATTca (SEQ ID NO:155)	HD-NI-NN-NN-NG-NI-HD-NG-HD-HD-NI-NI-NI-NN-NI-NG-NG
103068	atTCTCTGCTGGATGACGTga (SEQ ID NO:156)	NG-HD-NG-HD-NG-NN-HD-NG-NN-NN-NI-NG-NN-NI-HD-NN-NG
103069	gtACTCCAAGATTTCAGGttt (SEQ ID NO:157)	NI-HD-NG-HD-HD-NI-NI-NI-NN-NI-NG-NG-HD-NI-NN-NN-NG
103070	ttTCCATTCTCTGCTGGATga (SEQ ID NO:158)	NG-HD-HD-NI-NG-NG-HD-NG-HD-NG-NN-HD-NG-NN-NN-NI-NG
103071	ctCCAAGATTTCAGGTTTact (SEQ ID NO:159)	HD-HD-NI-NI-NI-NN-NI-NG-NG-HD-NI-NN-NN-NG-NG-NG-NI
103072	ttGACTTTCCATTCTCTGctg (SEQ ID NO:160)	NN-NI-HD-NG-NG-NG-HD-HD-NI-NG-NG-HD-NG-HD-NG-NN-HD
103073	ctCACGTCATCCAGCAGAGaa (SEQ ID NO:161)	HD-NI-HD-NN-NG-HD-NI-NG-HD-HD-NI-NN-HD-NI-NN-NI-NK
103074	atAGCAATTCAGGAAATTTga (SEQ ID NO:162)	NI-NN-HD-NI-NI-NG-NG-HD-NI-NN-NN-NI-NI-NI-NG-NG-NG
103075	ttCCTGAATTGCTATGTGTct (SEQ ID NO:163)	HD-HD-NG-NN-NI-NI-NG-NG-NN-HD-NG-NI-NG-NN-NG-NN-NG
103076	gtCAACTTCAATGTCGGATgg (SEQ ID NO:164)	HD-NI-NI-HD-NG-NG-HD-NI-NI-NG-NN-NG-HD-NN-NN-NI-NG
103077	ctATGTGTCTGGGTTTCATcc (SEQ ID NO:165)	NI-NG-NN-NG-NN-NG-HD-NG-NN-NN-NN-NG-NG-NG-HD-NI-NG
103078	ttCTTCAGTAAGTCAACTTca (SEQ ID NO:166)	HD-NG-NG-HD-NI-NN-NG-NI-NI-NN-NG-HD-NI-NI-HD-NG-NG
103079	atGTGTCTGGGTTTCATCCat (SEQ ID NO:167)	NN-NG-NN-NG-HD-NG-NN-NN-NN-NG-NG-NG-HD-NI-NG-HD-HD
103080	atTCTTCAGTAAGTCAACTtc (SEQ ID NO:168)	NG-HD-NG-NG-HD-NI-NN-NG-NI-NI-NN-NG-HD-NI-NI-HD-NG
103081	gtCTGGGTTTCATCCATCCga (SEQ ID NO:169)	HD-NG-NN-NN-NN-NG-NG-NG-HD-NI-NG-HD-HD-NI-NG-HD-HD
103082	ctCCATTCTTCAGTAAGTCaa (SEQ ID NO:170)	HD-HD-NI-NG-NG-HD-NG-NG-HD-NI-NN-NG-NI-NI-NN-NG-HD
103083	ttTCATCCATCCGACATTGaa (SEQ ID NO:171)	NG-HD-NI-NG-HD-HD-NI-NG-HD-HD-NN-NI-HD-NI-NG-NG-NK
103084	ttCTCTCTCCATTCTTCAGta (SEQ ID NO:172)	HD-NG-HD-NG-HD-NG-HD-HD-NI-NG-NG-HD-NG-NG-HD-NI-

[0207]

	(SEQ ID NO:172)	NK
103085	atCCATCCGACATTGAAGTtg (SEQ ID NO:173)	HD-HD-NI-NG-HD-HD-NN-NI-HD-NI-NG-NG-NN-NI-NI-NN-NG
103086	ttCAATTCTCTCTCCATTctt (SEQ ID NO:174)	HD-NI-NI-NG-NG-HD-NG-HD-NG-HD-NG-HD-NI-NG-NG-HD
103087	atCCGACATTGAAGTtGACTt (SEQ ID NO:175)	HD-HD-NN-NI-HD-NI-NG-NG-NN-NI-NI-NN-NG-NG-NN-NI-HD
103088	ctTTTTCAATTCTCTCTCCat (SEQ ID NO:176)	NG-NG-NG-NG-HD-NI-NI-NG-NG-HD-NG-HD-NG-HD-NG-HD-HD
103089	ttGAAGTTGACTTACTGAga (SEQ ID NO:177)	NN-NI-NI-NN-NG-NG-NN-NI-HD-NG-NG-NI-HD-NG-NN-NI-NI
103090	atGCTCCACTTTTTCAATTct (SEQ ID NO:178)	NN-HD-NG-HD-HD-NI-HD-NG-NG-NG-NG-NG-NG-HD-NI-NI-NG-NG
103091	gtTGACTTACTGAAGAATGga (SEQ ID NO:179)	NG-NN-NI-HD-NG-NG-NI-HD-NG-NN-NI-NI-NN-NI-NI-NG-NK
103092	gtCTGAATGCTCCACTTTTct (SEQ ID NO:180)	HD-NG-NN-NI-NI-NG-NN-HD-NG-HD-HD-NI-HD-NG-NG-NG-NG
103093	atGGAGAGAGAATTGAAAag (SEQ ID NO:181)	NN-NN-NI-NN-NI-NN-NI-NN-NI-NI-NG-NG-NN-NI-NI-NI-NI
103094	ctTGCTGAAAGACAAGTCTga (SEQ ID NO:182)	NG-NN-HD-NG-NN-NI-NI-NI-NN-NI-HD-NI-NI-NN-NG-HD-NG
103095	ttCAGACTTGCTTTTCAGCaa (SEQ ID NO:183)	HD-NI-NN-NI-HD-NG-NG-NN-NG-HD-NG-NG-NG-HD-NI-NN-HD
103096	gtGTAGTACAAGAGATAGaa (SEQ ID NO:184)	NN-NG-NI-NN-NG-NI-HD-NI-NI-NN-NI-NN-NI-NG-NI-NN-NI
103097	ctTGCTTTTCAGCAAGGACTg (SEQ ID NO:185)	NG-NN-NG-HD-NG-NG-NG-HD-NI-NN-HD-NI-NI-NN-NN-NI-HD
103098	atTCAGTGTAGTACAAGAGat (SEQ ID NO:186)	NG-HD-NI-NN-NG-NN-NG-NI-NN-NG-NI-HD-NI-NI-NN-NI-NK
103099	ctTTCAGCAAGGACTGGTctt (SEQ ID NO:187)	NG-NG-HD-NI-NN-HD-NI-NI-NN-NN-NI-HD-NG-NN-NN-NG-HD
103100	gtGAATTTCAGTGTAGTACAag (SEQ ID NO:188)	NN-NI-NI-NG-NG-HD-NI-NN-NG-NN-NG-NI-NN-NG-NI-HD-NI
103101	ctGGTCTTTCTATCTCTGta (SEQ ID NO:189)	NN-NN-NG-HD-NG-NG-NG-HD-NG-NI-NG-HD-NG-HD-NG-NG-NK
103102	ttTTTCAGTGGGGTGAATtc (SEQ ID NO:190)	NG-NG-NG-HD-NI-NN-NG-NN-NN-NN-NN-NG-NN-NI-NI-NG

[0208]

103103	ctATCTCTTGTACTACACTga (SEQ ID NO:191)	NI-NG-HD-NG-HD-NG-NG-NN-NG-NI-HD-NG-NI-HD-NI-HD-NG
103104	atACTCATCTTTTTCAGTGgg (SEQ ID NO:192)	NI-HD-NG-HD-NI-NG-HD-NG-NG-NG-NG-NG-HD-NI-NN-NG-NK
103105	ctACACTGAATTCACCCCCac (SEQ ID NO:193)	NI-HD-NI-HD-NG-NN-NI-NI-NG-NG-HD-NI-HD-HD-HD-HD
103106	ttCACACGGCAGGCATACTca (SEQ ID NO:194)	HD-NI-HD-NI-HD-NN-NN-HD-NI-NN-NN-HD-NI-NG-NI-HD-NG
103107	atTCACCCCCACTGAAAAga (SEQ ID NO:195)	NG-HD-NI-HD-HD-HD-HD-HD-NI-HD-NG-NN-NI-NI-NI-NI-NI
103108	gtCACATGGTTCACACGGCag (SEQ ID NO:196)	HD-NI-HD-NI-NG-NN-NN-NG-NG-HD-NI-HD-NI-HD-NN-NN-HD
103109	atTCACCCCCACTGAAAAga (SEQ ID NO:195)	NG-HD-NI-HD-HD-HD-HD-HD-NI-HD-NG-NN-NI-NI-NI-NI-NI
103110	gtCACATGGTTCACACGGCag (SEQ ID NO:186)	HD-NI-HD-NI-NG-NN-NN-NG-NG-HD-NI-HD-NI-HD-NN-NN-HD
103111	ctGAAAAAGATGAGTATGCct (SEQ ID NO:197)	NN-NI-NI-NI-NI-NI-NN-NI-NG-NN-NI-NN-NG-NI-NG-NN-HD
103112	ctGTGACAAAGTCACATGGtt (SEQ ID NO:198)	NN-NG-NN-NI-HD-NI-NI-NI-NN-NG-HD-NI-HD-NI-NG-NN-NK
103113	atGAGTATGCCTGCCGTGTga (SEQ ID NO:199)	NN-NI-NN-NG-NI-NG-NN-HD-HD-NG-NN-HD-HD-NN-NG-NN-NG
103114	ctATCTTGGGCTGTGACAAag (SEQ ID NO:200)	NI-NG-HD-NG-NG-NN-NN-NN-HD-NG-NN-NG-NN-NI-HD-NI-NI
103115	gtATGCCCTGCCGTGTGAACca (SEQ ID NO:201)	NI-NG-NN-HD-HD-NG-NN-HD-HD-NN-NG-NN-NG-NN-NI-NI-HD
103116	ttAACTATCTTGGGCTGTGac (SEQ ID NO:202)	NI-NI-HD-NG-NI-NG-HD-NG-NG-NN-NN-NN-HD-NG-NN-NG-NK
103117	gtGTGAACCATGTGACTTTgt (SEQ ID NO:203)	NN-NG-NN-NI-NI-HD-HD-NI-NG-NN-NG-NN-NI-HD-NG-NG-NG
103118	ttACCCCACTTAACATATCTtg (SEQ ID NO:204)	NI-HD-HD-HD-HD-NI-HD-NG-NG-NI-NI-HD-NG-NI-NG-HD-NG
103119	atGTGACTTTGTACAGCCca (SEQ ID NO:205)	NN-NG-NN-NI-HD-NG-NG-NG-NN-NG-HD-NI-HD-NI-NN-HD-HD

[0209]

[0210]

표 2B로부터의 TALEN을 반응당 각각의 TALEN 25, 100 또는 400 ng의 3가지 상이한 농도 하에 시험하였다. 시험된 모든 TALEN이 그의 표적 부위에 결합하는 것으로 밝혀졌고, 뉴클레아제로서 활성인 것으로 밝혀졌으며; 예시적인 데이터가 표 2C 및 도 2b에 제시된다.

[0211] 표 2C: K562 세포에서 TALEN 쌍의 활성

SBS#	%Indel-25ng	%Indel-100ng	%Indel-400ng
103049:103050	2.0	7.3	19.5
103051:103052	18.3	38.7	63.8
103053:103054	12.0	17.4	32.4
103055:103056	6.3	12.7	25.1
103057:103058	15.7	24.5	46.2
103061:103062	3.1	5.7	22.9
103063:103064	1.9	4.0	14.8
103065:103066	7.8	13.7	41.2
103067:103068	14.1	25.6	49.3
103069:103070	2.4	5.0	27.9
103071:103072	1.5	3.6	13.3
103073:103074	0.1	0.5	3.0
103075:103076	0.3	0.5	1.5
103077:103078	2.0	5.8	17.1
103079:103080	15.3	30.1	42.3
103081:103082	5.2	16.2	27.5
103083:103084	7.3	12.2	32.2
103085:103086	0.3	1.3	3.9
103087:103088	0.7	4.4	10.5
103089:103090	1.3	8.7	16.1
103091:103092	14.3	33.5	48.5
103093:103094	2.4	7.6	20.2
103095:103096	12.0	23.5	42.0
103097:103098	10.0	28.3	52.0
103099:103100	1.9	7.5	15.3
103101:103102	3.3	7.0	15.5
103103:103104	15.8	29.3	44.9
103105:103106	2.2	5.9	14.2
103107:103108	1.0	2.1	4.2
103109:103110	1.0	3.8	8.0
103111:103112	11.3	30.2	26.5
103113:103114	13.5	22.2	26.6
103115:103116	29.8	41.0	66.1
103117:103118	5.8	20.6	45.7
103119:103120	14.5	40.9	57.7

[0212]

[0213] 따라서, 본원에 기재된 뉴클레아제 (예를 들어, ZFP, TALE 또는 sgRNA DNA-결합 도메인을 포함하는 뉴클레아제)는 그의 표적 부위에 결합하고 B2M 유전자를 절단함으로써, 서열식별번호: 6-48 또는 137-205 중 어느 것을 포함하는 B2M 유전자 내에서의 유전적 변형을 만들며, 이는 이들 서열 (서열식별번호: 6-48 또는 137-205) 중 어느 것 이내에서의 변형 (삽입 및/또는 결실); 이들 유전자 서열의 1-50개 (예를 들어, 1 내지 10개) 염기 쌍 이내에서의 변형; (이량체의 경우) 쌍을 이룬 표적 부위의 표적 부위 간의 변형; 및/또는 하기 서열: GGCCTTA, TCAAATT, TCAAAT, TTAAGA 및/또는 AATTGAA 중 하나 이상 이내에서의 변형을 포함한다 (도 1 참조).

[0214] 더욱이, DNA-결합 도메인 (ZFP, TALE 및 sgRNA) 모두는 그의 표적 부위에 결합하였고, 이는 또한, 하나 이상의 전사 조절성 도메인과 연합될 때 활성의 조작된 전사 인자로 공식화된다.

[0215] 실시예 2: T 세포에서의 B2M-특이적 ZFN 활성

[0216] B2M-특이적 ZFN 쌍을 대상으로 하여 뉴클레아제 활성을 알아보기 위하여 인간 T 세포에서 시험하였다. 이러한 ZFN을 코딩하는 mRNA를 정제된 T 세포 내로 형질감염시켰다. 간략하게 설명하면, T 세포를 백혈구성분 채집술 생성물로부터 획득하였고, 밀테니이 클리니MACS(Miltenyi CliniMACS) 시스템 (CD4 및 CD8 이중 선택)을 사용하여 정제하였다. 이어서, 이들 세포를 제조업자의 프로토콜에 따라서 다이나비드(Dynabeads) (썬코퍼서

(ThermoFisher))를 사용하여 활성화시켰다. 활성화시킨지 3일 후, 상기 세포를 표준 프로토콜에 따라서 BTX 96 웰 전기천공기 (BTX)를 사용하여 2회 용량의 mRNA (총 2개 ZFN 중의 2 또는 6 μ g)로 형질감염시켰다. 이어서, 활성화 후 총 10일 동안 부가의 7일 동안 세포를 확장시켰다. 세포를 제7일에 꺼내고, 심층 서열 분석 (Miseq; 일루미나(Illumina))을 사용하여 표적에 적중한 B2M 변형에 관하여 분석하였으며, 제10일에는 HLA-A, -B 및 -C 염색을 이용하여 FACS 분석하였다.

[0217] B2M-특이적 ZFN 쌍은 모두 T 세포에서 활성이었고, 6 μ g 및 2 μ g mRNA 용량 각각에 대하여 평균 89% 및 83%를 유발시켰다 (도 2 참조). 사용된 쌍 및 위치 (도 1에 도시됨)가 아래 표 3에 열거된다.

[0218] 표 3: B2M 특이적 ZFN 쌍 및 표적 부위

SBS 쌍	부위, 쌍
57071 및 57531	A1
57362 및 57376	C1
57017 및 57327	D1
57017 및 57328	D2
57332 및 57327	D3
57469 및 57327	D4
57469 및 57328	D5
57331 및 57326	D6
55822 및 57511	E1
55822 및 57509	E2
57482 및 57511	E3
55822 및 57347	E4
57296 및 57322	G1
57296 및 57323	G2
57447 및 57322	G3

[0219]

[0220] 유사하게, ZFN으로 처리된 T 세포는 HLA A, B 및 C의 발현을 상실하였으며, 여기서 FACS 분석은 6 μ g 및 2 μ g mRNA 용량 각각에서 평균 81% 및 67% HLA 음성 T 세포를 나타내었다 (도 3 참조).

[0221] 실시예 3: 시험관 내에서 B2M에 대항한 가이드 RNA의 활성

[0222] 본 실험에서는, Cas9를 pVAX 플라스미드 상에 공급하였고, sgRNA를 U6 프로모터의 제어하에 특정 플라스미드 상에 공급하였다. 이들 플라스미드를 각각 100 ng 또는 각각 400 ng 하에 혼합하였고, 실험당 2e5개 세포와 혼합하였다. 상기 세포를 아막사(Amaxa) 시스템을 사용하여 형질감염시켰다. 간략하게는, 아막사 형질감염 키트를 사용하였고, 핵산을 표준 아막사 서틀 프로토콜을 이용하여 형질감염시켰다. 형질감염 후, 상기 세포를 실온 하에 10분 동안 정치시켜 둔 다음, 미리 가온된 RPMI에 재현탁시켰다. 그런 다음, 세포를 37°C 하의 표준 조건에서 성장시켰다. 형질감염 후 7일에 게놈 DNA를 단리시키고, 이를 대상으로 MiSeq 분석을 수행하였다.

[0223] 아래에 제시된 데이터 (표 4)는 2개 용량의 가이드 RNA에서 검출된 indel (삽입 및 결실)의 퍼센트를 표시하고, 상이한 가이드 RNA가 표적화된 부위에서 절단을 유도시켰다는 것을 표시하였다. 숫자는 두 실험의 평균을 나타낸다. 모든 가이드는 활성이었다.

[0224]

표 4: B2M에 대한 CRISPR/Cas 시스템의 활성

가이드	평균 100 ng (% indel)	평균 400 ng (% indel)	GFP (% indel)
B2M-Gf1073	31.12	55.85	0.20
B2M-Gr1074	35.21	55.86	0.24
B2M-Gr1080	24.83	62.22	0.17
B2M-Gf1107	4.32	52.68	0.19
B2M-Gf1112	9.99	22.09	0.11
B2M-Gf1115	26.52	28.97	0.16
B2M-Gr1114	16.93	53.88	0.12
B2M-Gr1126	15.25	55.92	0.04
B2M-Gr4942	3.05	48.28	0.15
B2M-Gf4948	1.62	12.18	0.11
B2M-Gr4969	1.68	10.11	0.16
B2M-Gf4976	7.47	12.65	0.14
B2M-Gr4995	1.14	31.89	0.16
B2M-Gf5009	0.93	8.87	0.10
B2M-Gf5010	0.30	5.84	0.16
B2M-Gr5023	2.66	2.41	0.22
B2M-Gr5027	14.90	13.13	0.19
B2M-Gf5051	12.98	42.24	0.25
B2M-Gf5071	5.16	47.59	0.44
B2M-Gf5098	4.17	25.60	0.35
B2M-Gf5103	4.00	24.85	0.45
B2M-Gr5141	3.53	18.55	0.32
B2M-Gr5142	3.26	18.22	0.35
B2M-Gr5143	1.79	20.51	0.35
B2M-Gr5144	7.04	6.99	0.34
B2M-Gr5165	6.48	26.44	0.56
B2M-Gr5169	4.62	34.59	0.44
B2M-Gr5178	3.62	30.98	0.61
B2M-Gr5197	2.32	23.26	0.52
B2M-Gr5198	1.87	20.26	0.53
B2M-Gf5208	6.20	17.64	0.55
B2M-Gf5209	10.09	35.92	0.17
B2M-Gf5210	8.47	21.72	0.21

[0225]

[0226]

실시예 4: 1차 T 세포에서 B2M 및 TCR의 이중 녹아웃

[0227]

본원에 기재된 B2M 쌍을 또한, TCRA에 대해 특이적인 ZFN과 조합하여 시험하였다 (아래 표 5 및 미국 가출원 62/269,365 및 62/306,500 참조). 세포를 실시예 2에 기재된 바와 같이 수득하고 처리하였다. ZFN 쌍 (B2M의 경우 SBS#57017/SBS#57327 및 TCRA의 경우 SBS#55254/SBS#55248)을 코딩하는 mRNA를 제조업자의 지시에 따라서 맥스사이트 기기를 사용하여 세포 내로 전기천공하였다. 간략하게는, T 세포를 제0일에 활성화시키고, 제3일에 ZFN-코딩 mRNA로 처리하였으며, 여기서 세포 밀도는 3×10^7 개 세포/mL였다. 전기천공 후 밤새 30°C 저온 충격이 뒤따랐다. 제4일에는, 세포를 계수하고, 생존 능력에 관하여 검정하였으며, 0.5×10^6 개 세포/mL가 되도록 희석시키고, 이를 37°C 하의 배양물로 옮겼다. 제7일에는, 세포를 계수하고 다시 검정하며, 0.5×10^6 개 세포/mL가 되도록 재희석시켰다. 제10일 및 제14일에는, FACS 및 MiSeq 심층 서열 분석을 위하여 세포의 일정 부분을 수거하였다.

[0228] 표 5: TCRA ZFN

ZFN 명칭 표적 서열	F1	F2	F3	F4	F5	F6	도메인 링커
SBS55266 5'tcAAGCTGG TCGAGaAAAGC Tttgaaac (SEQ ID NO:206)	QSSDLRS (SEQ ID NO:68)	QSGNRTT (SEQ ID NO:208)	RSANLAR (SEQ ID NO:209)	DRSALAR (SEQ ID NO:210)	RSDVLSE (SEQ ID NO:211)	KHSTRRV (SEQ ID NO:212)	N7c
SBS53853 5'aaCAGGTAA GACAGGGGTCT Agcctggg (SEQ ID NO:207)	TMHQRVE (SEQ ID NO:213)	TSGHLSR (SEQ ID NO:214)	RSDHLTQ (SEQ ID NO:215)	DSANLSR (SEQ ID NO:216)	QSGSLTR (SEQ ID NO:59)	AKWNLDA (SEQ ID NO:76)	L0
SBS55254 5'ctCCTGAAA GTGGCCGGgtt taatctgc (SEQ ID NO:217)	RSDHLST (SEQ ID NO:219)	DRSHLAR (SEQ ID NO:220)	LKQHLNE (SEQ ID NO:221)	QSGNLAR (SEQ ID NO:61)	HNSSLKD (SEQ ID NO:222)	N/A	L0
SBS55248 5'agGATTCGG AACCAATCAC tgacaggt (SEQ ID NO:218)	DQSNLRA (SEQ ID NO:56)	TSSNRKT (SEQ ID NO:223)	LQOTLAD (SEQ ID NO:224)	QSGNLAR (SEQ ID NO:61)	RREDLIT (SEQ ID NO:225)	TSSNLSR (SEQ ID NO:226)	L0

[0229]

[0230]

세포를 4개 군으로 분리시켰다: ZFN 없는 대조군, TCRA ZFN 단독, B2M ZFN 단독, 및 TCRA ZFN + B2M ZFN. FACS 분석에 의하면, TCRA 단독에 대항한 ZFN (180 μ g/ μ L ZFN mRNA) 또는 B2M 단독에 대항한 ZFN (180 μ g/ μ L의 ZFN mRNA) 세트는 높은 절단 속도를 제공한다 (TCRA ZFN의 경우 96% CD3 마킹 및 B2M ZFN의 경우 HLA 마킹의 92% 녹아웃 (도 4)). 세포를 양쪽 유형의 ZFN 쌍 (둘 다 180 μ g/ μ L)으로 처리한 경우, 세포의 82%가 CD3과 HLA 마킹 둘 다를 상실하였다.

[0231]

유사한 군의 세포를 또한, 제시된 바와 같은 가변적 양의 TCRA-특이적 ZFN (60-250 μ g/ μ L) 플러스 B2M ZFN (60 μ g/ μ L)으로 처리하였고, 제10일 및 제14에는 MiSeq 심층 서열 분석 (일루미나) 및 FACS 분석을 수행하였으며, 그 결과가 아래 표 6에 제시된다. 그 결과는 NHEJ-매개된 삽입 및 결실에 의해 검출된 바와 같은 높은 비율의 이중 녹아웃이 이들 ZFN의 경우에 관찰되었다는 것을 표시한다.

[0232]

표 6: TCRA/B2M 이중 녹아웃에 대한 FACS 및 miSeq 분석

	FACS						MiSeq			
	D10			D14			D10		D14	
	TCRA-	B2M-	이중-	TCRA-	B2M-	이중-	TCRA-	B2M-	TCRA	B2M-
모의	2.3	2.9	2.1	3.0	3.8	2.1	0.6	1.8	0.2	1.4
TRAC 180	67.4	1.8	1.5	40.1	1.0	0.5	66.0	0.8	52.4	2.9
B2M 180	3.4	90.5	3.1	1.5	89.5	1.4	0.9	64.6	0.2	74.7
T180 : B180	60.4	84.3	57.3	43.9	78.8	41.8	63.2	73.0	51.1	63.4
T240 : B120	73.3	85.3	68.3	56.9	80.2	53.2	73.0	79.5	66.0	66.7
T120 : B240	79.4	80.5	71.2	70.6	73.9	63.2	79.0	72.1	69.2	61.3

[0233]

[0234]

따라서, 그 데이터는 B2M 및 TCRA의 이중 녹아웃이 표적 또는 그 근처에서 및/또는 본원에 기재된 뉴클레아제의 표적 서열 및/또는 절단 부위 (B2M 서열 TCAAT (도 1에서의 부위 D) 및 SBS#55254/SBS#55248 TCRA-특이적 쌍에 대한 두 표적 서열 사이의 TCRA 서열 CCTTC 포함)의 1-50개 (예를 들어, 1 내지 10개) 염기 쌍 (쌍을 이룬

표적 부위 사이 포함) 이내에서 양쪽 유전자를 불활성화시켰다는 것을 입증해 준다.

[0235] **실시예 5: 표적화된 통합을 이용한 B2M 및 TCRA의 이중 녹아웃**

[0236] 상기 기재된 바와 같은 뉴클레아제 (실시예 4 참조)를 사용하여 B2M 및 TCRA를 불활성화시켰고 (실시예 5 참조), 표적화된 통합을 통하여 공여자 (트랜스젠)를 TCRA 또는 B2M 로커스 내로 도입하였다. 본 실험에서는, TCRA-특이적 ZFN 쌍이 TCRA-특이적 ZFN 표적 부위 사이의 서열 TTGAAA를 포함하는 SBS#55266/SBS#53853이었고 (표 5), B2M 쌍은 B2M-특이적 ZFN 표적 부위 사이의 서열 TCAAAT를 포함하는 SBS#57332/SBS#57327이었다 (표 1).

[0237] 간략하게는, T-세포 (AC-TC-006)를 해동시키고, X-vivo15 T-세포 배양 배지에서 CD3/28 다이나비드 (1:3 세포:비드 비)로 활성화시켰다 (제0일). 배양한지 2일 후 (제2일), AAV 공여자 (TCRA 또는 B2M 유전자에 대한 상동성 아암 및 GFP 트랜스젠을 포함함)를 세포 배양물에 부가하였으며, 단 공여자를 수반하지 않은 대조군을 또한 유지시켰다. 그 다음 날 (제3일), mRNA 전달을 통하여 TCRA 및 B2M ZFN을 하기 5개 군에 부가하였다:

[0238] (a) 군 1 (TCRA 및 B2M ZFN 단독, 공여자 없음): TCRA 120 ug/mL; B2M 단독 60 ug/mL;

[0239] (b) 군 2 (TCRA 및 B2M ZFN 및 TCRA 상동성 아암을 수반한 공여자): TCRA 120 ug/mL; B2M 60 ug/mL 및 AAV (TCRA-hPGK-eGFP-클론 E2) 1E5vg/세포;

[0240] (c) 군 3 (TCRA 및 B2M ZFN 및 TCRA 상동성 아암을 수반한 공여자): TCRA 120 ug/mL; B2M 60 ug/mL; 및 AAV (TCRA-hPGK-eGFP-클론 E2) 3E4vg/세포;

[0241] (d) 군 4 (TCRA 및 B2M ZFN 및 B2M 상동성 아암을 수반한 공여자): TCRA 120 ug/mL; B2M 60 ug/mL 및 AAV (pAAV B2M 부위 D hPGK GFP) 1E5vg/세포;

[0242] (e) 군 5 (TCRA 및 B2M ZFN 및 B2M 상동성 아암을 수반한 공여자): TCRA 120 ug/mL; B2M 60 ug/mL 및 AAV (pAAV B2M 부위 D hPGK GFP) 3E4vg/세포.

[0243] 모든 실험을 미국 출원 번호 15/347,182에 기재된 바와 같은 프로토콜 (극도의 저온 충격)을 이용하여 3e7개 세포/ml 세포 밀도로 시행하였고, 전기천공 후 30°C에서 밤새 저온 충격으로 배양하였다.

[0244] 그 다음 날 (제4일), 세포를 0.5e6개 세포/ml가 되도록 희석시키고, 이를 37°C 하의 배양물로 옮겼다. 3일 후 (제7일), 세포를 0.5e6개 세포/ml가 되도록 다시 희석시켰다. 3일 및 7일 더 배양한 후 (각각 제10일 및 제14일), FACS 및 MiSeq 분석을 위하여 세포를 수거하였다 (0.5e6개 세포/ml가 되도록 희석됨).

[0245] 도 5에 도시된 바와 같이, GFP 발현은 표적 통합이 성공적이었다는 것과, 뉴클레아제 부위 이내 (또는 TTGAAA 및 TCAAAT 이내 및/또는 쌍을 이룬 표적 부위 사이를 포함한, 뉴클레아제 표적 부위의 1 내지 50개 염기 쌍 이내)에 B2M 및 TCRA 변형 (삽입 및/또는 결실)을 포함하는 유전적으로 변형된 세포가 수득되었다는 것을 표시하였다.

[0246] CAR 트랜스젠을 B2M 및/또는 TCRA 로커스 내로 통합시켜, CAR을 발현하는 이중 B2M/TCRA 녹아웃을 창출시키는 실험을 또한 수행하였다.

[0247] 본원에 언급된 모든 특허, 특허 출원 및 공개는 그 전문이 본원에 참조로 포함된다.

[0248] 비록 명확성 및 이해의 목적을 위한 예 및 한 예시로서 개시내용이 일부 상세히 제공되었지만, 본 개시내용의 요지 또는 범위를 벗어나지 않고서도 다양한 변화 및 변형이 실시될 수 있다는 것이 관련 기술분야의 통상의 기술자에게 명백할 것이다. 따라서, 전술된 설명과 실시예는 제한되는 것으로 간주되지 않아야 한다.

예손 1

5' cgagatgctcgcctccgt**A**gcttgctcgcctactctcttcttgccctgaggctatccagc

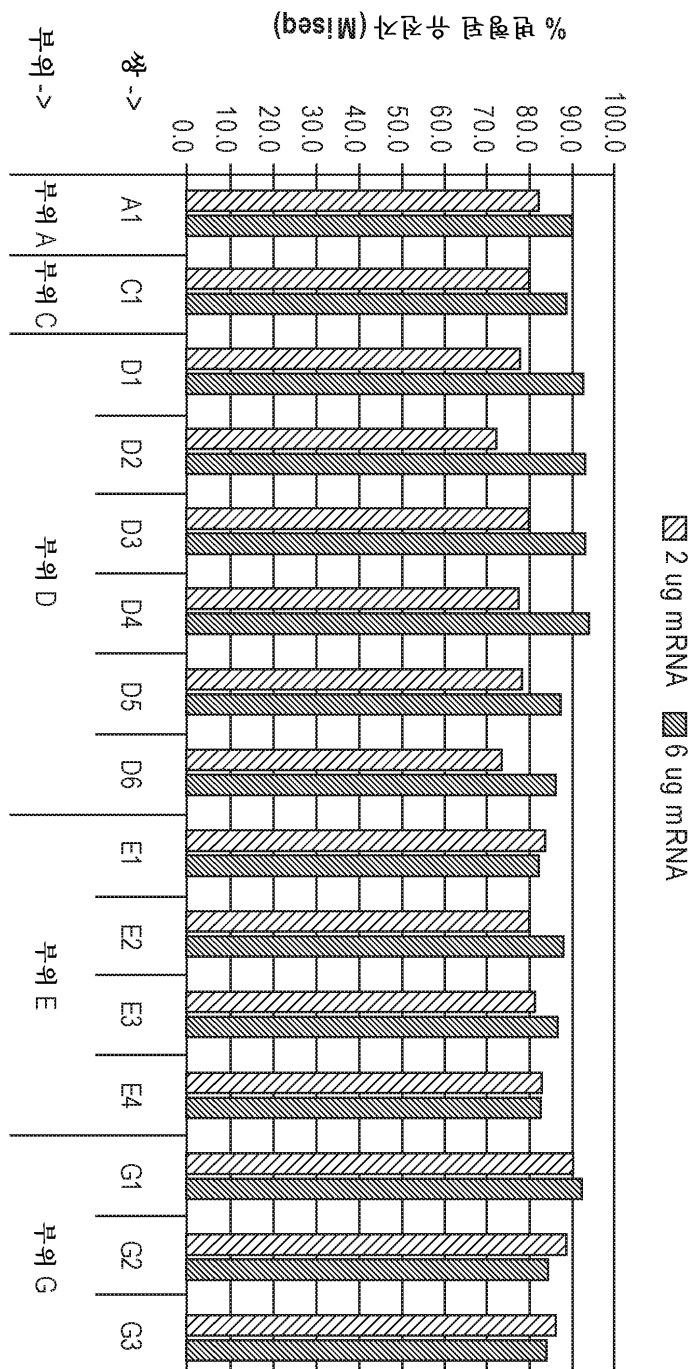
예손 2

5' gta^Cctccaagattcaggttactcacg^Dtcattccagcagagaatggaag^Ctcaatttcctgaattgc^Dtcaatttctgg
gtttcattccatccgacattgaagtga^Ecttactga^Gaagaatggagagagaattgaa^Gaaagtggagcattcagacttgct

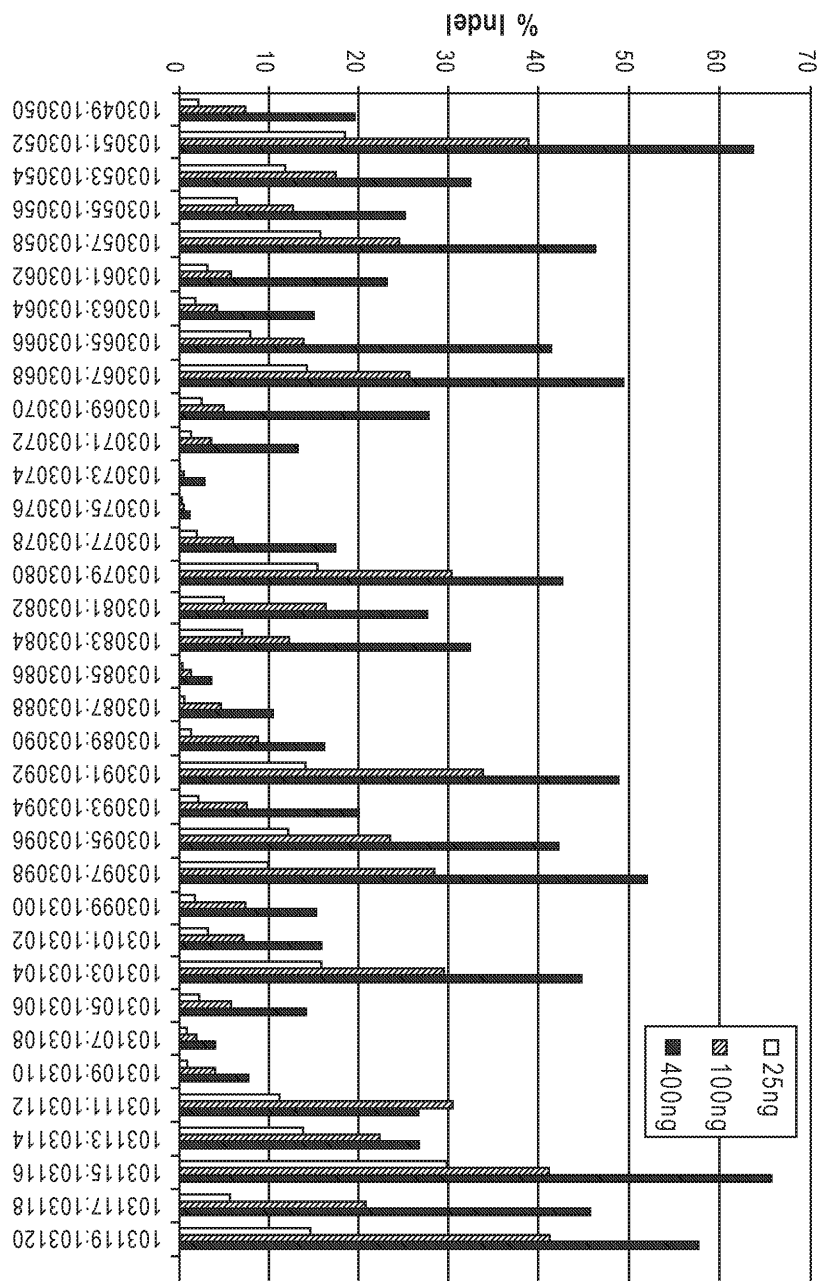
ttcagcagagactggtcttctatctcttgta^Gctacactgaattcacc^Cccactgaaaaagatgagtatgccctgccgtg
tgaaccattgta^Gcttgtcacagcccaagatagttaa^Ggtgg

도면
도면1

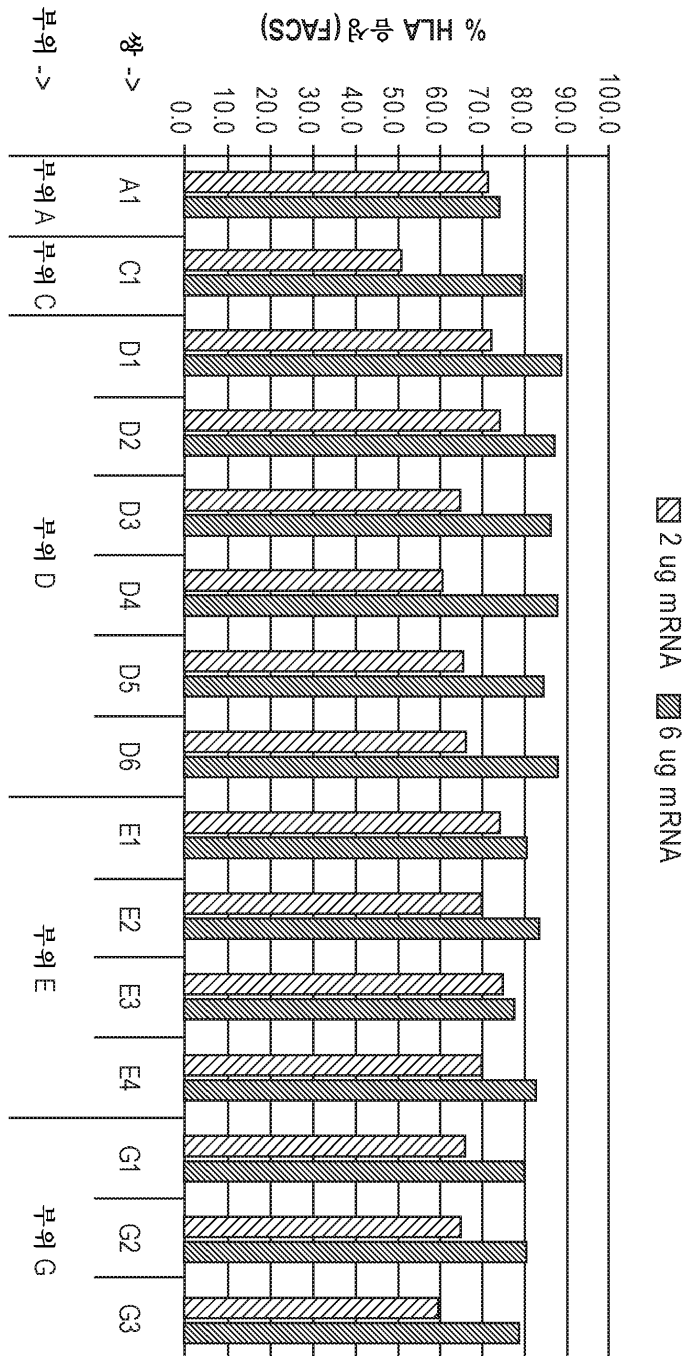
도면2a



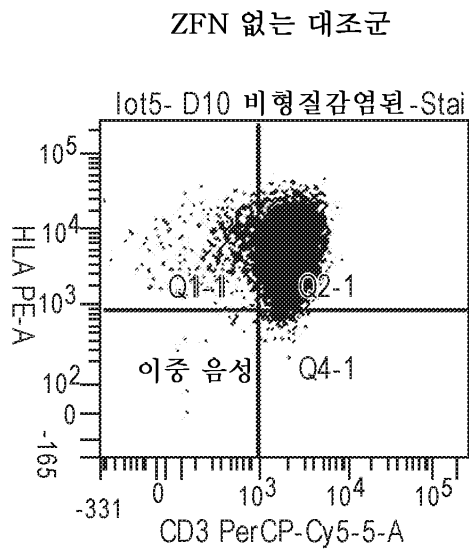
도면2b



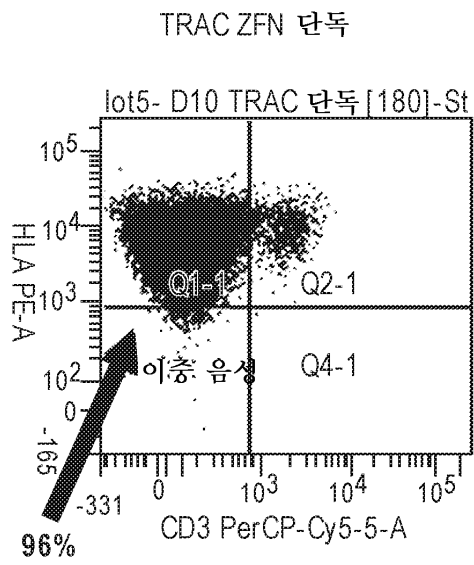
도면3



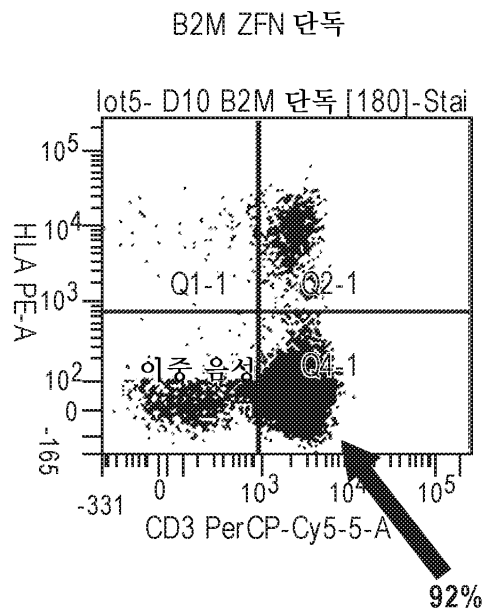
도면4a



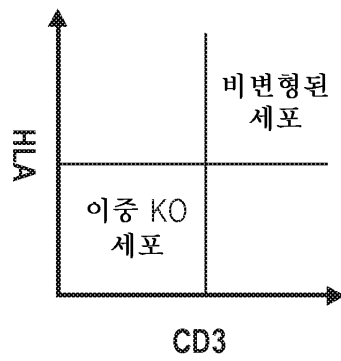
도면4b



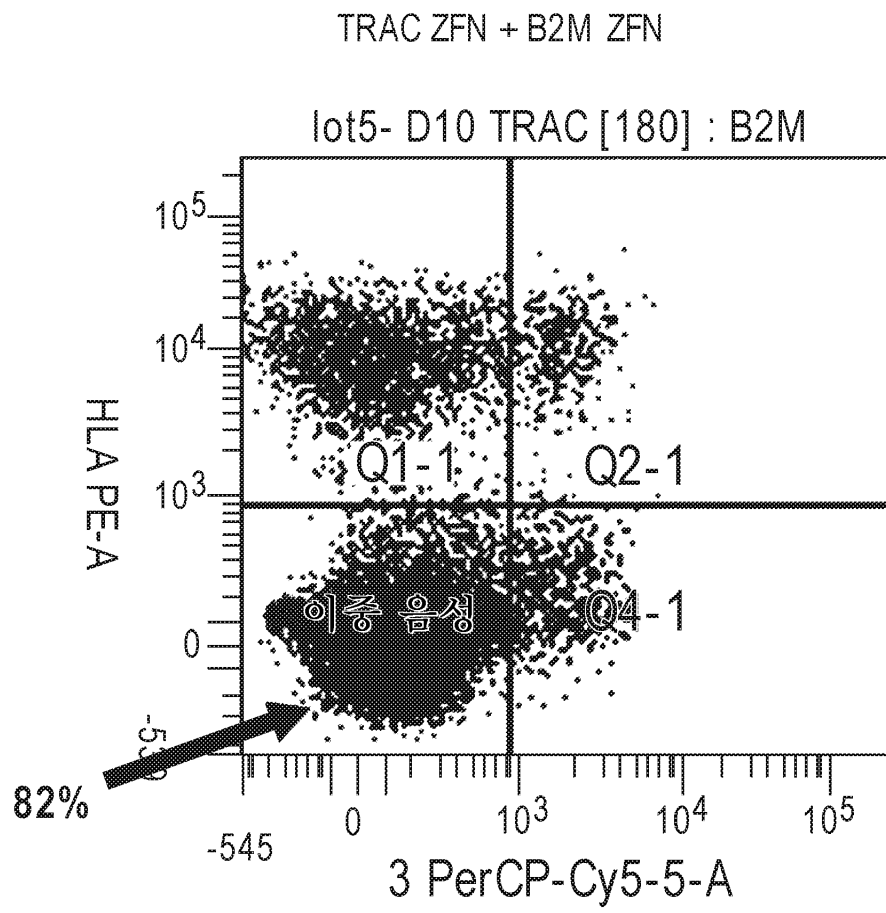
도면4c



도면4d



도면4e



도면5

FACS					
D10					
ug/mL	TRAC-	B2M-	이중-	총 GFP+	이중- GFP+
모의	0.5	0.1	0.0	0.0	0.0
TRAC/B2M KO 단독	85.0	83.6	80.0	0.0	0.0
TRAC/B2M KO + 1E5vg/세포 TRAC 로커스 AAV 공여자	91.9	92.7	89.3	80.8	83.0
TRAC/B2M KO + 3E4vg/세포 TRAC 로커스 AAV 공여자	91.2	93.4	89.1	71.9	74.3
TRAC/B2M KO + 1E5vg/세포 B2M 로커스 AAV 공여자	88.2	90.5	86.4	54.9	59.6
TRAC/B2M KO + 3E4vg/세포 B2M 로커스 AAV 공여자	89.8	92.2	87.9	43.2	46.7

서열목록

SEQUENCE LISTING

<110> SANGAMO BIOSCIENCES, INC.

<120> TARGETED DISRUPTION OF THE MHC CELL RECEPTOR

<130> 8325-0147.40

<140> PCT/US2016/066988

<141> 2016-12-15

<150> 62/329,439

<151> 2016-04-29

<150> 62/305,097

<151> 2016-03-08

<150> 62/269,410

<151> 2015-12-18

<160> 227

<170> PatentIn version 3.5

<210> 1

<211> 71

<212> DNA

<213> Homo sapiens

<400> 1

cgagatgtct cgctccgtgg ccttagctgt gctcgcgcta ctctctcttt ctggcctgga	60
ggctatccag c	71

<210> 2

<211> 280

<212> DNA

<213> Homo sapiens

<400> 2

gtactccaaa gattcagggt tactcacgtc atccagcaga gaatggaaag tcaaatttcc	60
tgaattgtct aaatttctgg gtttcatcca tccgacattg aagttgactt actgaagaat	120
ggagagagaa ttgaaaaagt ggagcattca gacttgtctt tcagcaagga ctggtctttc	180
tatctcttgt actacactga attcaccccc actgaaaaag atgagtatgc ctgccgtgtg	240
aacctgtga ctttgtcaca gcccaagata gttaagtggg	280

<210> 3

<211> 5

<212> PRT

<213> Artificial Sequence

<220

><223> Description of Artificial Sequence: Synthetic

peptide

<400> 3

Gln Leu Val Lys Ser

1 5

<210> 4

<211> 12

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 4

Ser Gly Thr Pro His Glu Val Gly Val Tyr Thr Leu

1 5 10

<210> 5

<211> 10

<212> PRT

<213> Artificial Sequence

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

peptide

<400>

> 5

Ser Gly Ala Gln Gly Ser Thr Leu Asp Phe

1 5 10

<210> 6

<211> 28

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 6

gccacggagc gagacatctc ggcccgaa 28

<210> 7

<211> 28

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 7

gagtagcgcg agcacagcta aggccacg 28

<210> 8

<211> 28

<212> DNA

<213> Artificial Sequence

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

<400> 8

tccagcagag aatggaaagt caaatttc 28

<210> 9

<211> 28

<212> DNA

<213> Artificial Sequence

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

<400> 9

tttctgaat tgctatgtgt ctgggttt 28

<210> 10

<211> 28

<212> DNA

<213> Artificial Sequence

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

<400> 10

tgtcgatgg atgaaaccca gacacata 28

<210> 11

<211> 28

<212> DNA

<213> Artificial Sequence

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

<400> 11

tagcaattca ggaaatttga ctttccat 28

<210> 12

<211> 28

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 12

catccgacat tgaagttgac ttactgaa 28

<210> 13

<211> 28

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 13

gaagaatgga gagagaattg aaaaagtg 28

<210> 14

<211> 28

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 14

ctgaagaatg gagagagaat tgaaaaag 28

<210> 15

<211> 28

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 15

aaaaagtgga gcattcagac ttgtcttt 28

<210> 16

<211> 23

<212> DNA

<213> Artificial Sequence

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

<400> 16

ggccgagatg tctcgctccg tgg 23

<210> 17

<211> 23

<212> DNA

<213> Artificial Sequence

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

<400> 17

cgcgagcaca gctaaggcca cgg 23

<210> 18

<211> 23

<212> DNA

<213> Artificial Sequence

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

<400> 18

gagtagcgcg agcacagcta agg 23

<210> 19

<211> 23

<212> DNA

<213> Artificial Sequence

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

<400> 19

ctcgcgctac tctctctttc tgg 23

<210> 20

<211> 23

<212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 20
 gctactctct ctttctggcc tgg 23
 <210> 21
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 21
 actctctctt tctggcctgg agg 23
 <210> 22
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 22
 actcacgtg gatagcctcc agg 23
 <210> 23
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 23
 agggtaggag agactcacgc tgg 23
 <210> 24
 <211> 23
 <212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 24

cgtgagtaaa cctgaatctt tgg 23

<210> 25

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 25

ctcaggtact ccaaagattc agg 23

<210> 26

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 26

tttgactttc cattctctgc tgg 23

<210> 27

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 27

tcacgtcatc cagcagagaa tgg 23

<210> 28

<211> 23

<212> DNA

<213> Artificial Sequence

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

<400> 28

accagacac atagcaattc agg 23

<210> 29

<211> 23

<212> DNA

<213> Artificial Sequence

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

<400> 29

ttcctgaatt gctatgtgtc tgg 23

<210> 30

<211> 23

<212> DNA

<213> Artificial Sequence

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

<400> 30

tcctgaattg ctatgtgtct ggg 23

<210> 31

<211> 23

<212> DNA

<213> Artificial Sequence

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

<400> 31

aagtcaactt caatgtcgga tgg 23

<210> 32

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 32

cagtaagtca acttcaatgt cgg

23

<210> 33

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 33

gaagttgact tactgaagaa tgg

23

<210> 34

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 34

tggagagaga attgaaaaag tgg

23

<210> 35

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 35

ttcagacttg tctttcagca agg

23

<210> 36

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 36
acttgctctt cagcaaggac tgg 23
<210> 37
<211> 23
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 37
atactcatct ttttcagtgg ggg 23
<210> 38
<211> 23
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 38
catactcatc tttttcagtg ggg 23
<210> 39
<211> 23
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 39
gcatactcat ctttttcagt ggg 23
<210> 40
<211> 23
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 40

ggcataactca tctttttcag tgg 23

<210> 41

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 41

agtcacatgg ttcacacggc agg 23

<210> 42

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 42

acaaagtcac atggttcaca cgg 23

<210> 43

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 43

tgggctgtga caaagtcaca tgg 23

<210> 44

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 44

ttacccact taactatctt ggg 23

<210> 45

<211> 23

<212> DNA

<213> Artificial Sequence

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

<400> 45

cttaccacac ttaactatct tgg

23

<210> 46

<211> 23

<212> DNA

<213> Artificial Sequence

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

<400> 46

cacagcccaa gatagttaag tgg

23

<210> 47

<211> 23

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 47

acagcccaag atagttaagt ggg

23

<210> 48

<211> 23

<212> DNA

<213> Artificial Sequence

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

<400> 48

cagcccaaga tagttaagt ggg

23

<210> 49

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 49

Arg Ser Asp Asp Leu Ser Lys

1 5

<210> 50

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 50

Asp Ser Ser Ala Arg Lys Lys

1 5

<210> 51

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 51

Asp Arg Ser Asn Leu Ser Arg

1 5

<210> 52

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

peptide

<400> 52

Gln Arg Thr His Leu Arg Asp

1 5
 <210> 53
 <211> 7
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 peptide

<400> 53
 Gln Ser Gly His Leu Ala Arg

1 5
 <210> 54
 <211> 7
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 peptide

<400> 54
 Asp Ser Ser Asn Arg Glu Ala

1 5
 <210> 55
 <211> 7
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 peptide

<400> 55
 Ala Gln Cys Cys Leu Phe His

1 5
 <210> 56
 <211> 7
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 peptide

<400> 56

Asp Gln Ser Asn Leu Arg Ala

1 5

<210> 57

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

peptide

<400> 57

Arg Ser Ala Asn Leu Thr Arg

1 5

<210> 58

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 58

Arg Ser Asp Asp Leu Thr Arg

1 5

<210> 59

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 59

Gln Ser Gly Ser Leu Thr Arg

1 5

<210> 60

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 60

Leu Asn His His Leu Gln Gln

1 5

<210> 61

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 61

Gln Ser Gly Asn Leu Ala Arg

1 5

<210> 62

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
peptide

<400> 62

Arg Ser Asp Thr Leu Ser Ala

1 5

<210> 63

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 63

Gln Asn Ala His Arg Lys Thr

1 5

<210> 64

<211> 7
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 peptide
 <400> 64
 Arg Ser Asp Asn Leu Ser Glu

1 5
 <210> 65
 <211> 7
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 peptide
 <400> 65
 Lys Pro Tyr Asn Leu Arg Thr

1 5
 <210> 66
 <211> 7
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 peptide
 <400> 66
 Thr Arg Asp His Leu Ser Thr

1 5
 <210> 67
 <211> 7
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence: Synthetic
 peptide
 <400> 67

Arg Ser Asp Ala Arg Thr Asn

1 5

<210> 68

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 68

Gln Ser Ser Asp Leu Ser Arg

1 5

<210> 69

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 69

His Arg Ser Ser Leu Lys Asn

1 5

<210> 70

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 70

Gln Ser Ser His Leu Thr Arg

1 5

<210> 71

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 71

Asp Ser Ser Asp Arg Lys Lys

1 5

<210> 72

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

peptide

<400> 72

Ala Ser Lys Thr Arg Thr Asn

1 5

<210> 73

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 73

Thr Ser Gly Asn Leu Thr Arg

1 5

<210> 74

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 74

Arg Ile Gln Asp Leu Asn Lys

1 5

<210> 75

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 75

Ala Arg Trp Tyr Leu Asp Lys

1 5

<210> 76

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 76

Ala Lys Trp Asn Leu Asp Ala

1 5

<210> 77

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

peptide

<400> 77

Gln Gln His Val Leu Gln Asn

1 5

<210> 78

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 78

Gln Asn Ala Thr Arg Thr Lys

1 5

<210> 79

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 79

Thr Asn Gln Ser Leu His Trp

1 5

<210> 80

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 80

Arg Ser Asp Asn Leu Arg Glu

1 5

<210> 81

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 81

Ala Ser His Val Leu Asn Ala

1 5

<210> 82

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
peptide

<400> 82

Thr Ser Ala Asn Leu Ser Arg

1 5

<210> 83

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 83

Arg Thr Glu Asp Arg Leu Ala

1 5

<210> 84

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 84

Tyr Thr Ser Ser Leu Cys Tyr

1 5

<210> 85

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 85

Gln Ser Gly His Leu Ser Arg

1 5

<210> 86

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 86

Ala Ser Lys Thr Arg Lys Asn

1 5

<210> 87

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic
peptide

<400> 87

Thr Thr Pro Val Leu Val Gln

1 5

<210> 88

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 88

Phe Pro Gly Ser Arg Thr Arg

1 5

<210> 89

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 89

Trp Arg Ile Ser Leu Ala Ala

1 5

<210> 90

<211> 7
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 peptide
 <400> 90
 Asp Gln Ser Leu Leu Arg Thr
 1 5
 <210> 91
 <211> 7
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 peptide
 <400> 91
 His Arg Leu Gly Leu Arg Asp
 1 5
 <210> 92
 <211> 7
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Description of Artificial Sequence: Synthetic
 peptide
 <400> 92
 Arg Ser Asp Val Leu Ser Thr
 1 5
 <210> 93
 <211> 7
 <212> PRT
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 peptide
 <400> 93
 Gln Asn Ala His Arg Ile Lys

1 5

<210> 94

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 94

Gln Ser Ala His Arg Lys Asn

1 5

<210> 95

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 95

His Lys Leu Ser Leu Ser Ile

1 5

<210> 96

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 96

Arg His Ser His Leu Thr Ser

1 5

<210> 97

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Synthetic

peptide

<400> 97

Gln Ser Asn Gln Leu Ala Val

1 5

<210> 98

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 98

Gln Ser Ala Asp Arg Thr Lys

1 5

<210> 99

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 99

Thr Asn Gln Asn Arg Ile Thr

1 5

<210> 100

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 100

Arg Ser Asp Ser Leu Ser Val

1 5

<210> 101

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 101

Gln Asn Ala Asn Arg Lys Thr

1 5

<210> 102

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 102

Gln Arg Gly Asn Leu Trp Thr

1 5

<210> 103

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 103

Leu Lys Gln Asn Leu Asp Ala

1 5

<210> 104

<211> 20

<212> DNA

<213> Artificial Sequence

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

<400> 104

ggccgagatg tctcgctccg

<210> 105

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 105

gcgcgagcac agctaaggcc a 21

<210> 106

<211> 20

<212> DNA

<213> Artificial Sequence

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

<400> 106

gagtagcgcg agcacagcta 20

<210> 107

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 107

gctcgcgcta ctctctcttt c 21

<210> 108

<211> 20

<212> DNA

<213> Artificial Sequence

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

<400> 108

gtactctct ctttctggcc 20

<210> 109

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 109

gactctctct ttctggcctg g 21

<210> 110

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 110

gactcacgct ggatagcctc c 21

<210> 111

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 111

gagggtagga gagactcacg c 21

<210> 112

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 112

gcgtgagtaa acctgaatct t 21

<210> 113

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 113

gctcaggtac tccaaagatt c

21

<210> 114

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 114

gtttgacttt ccattctctg c

21

<210> 115

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 115

gtcacgtcat ccagcagaga a

21

<210> 116

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 116

gaccagaca catagcaatt c

21

<210> 117

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 117

gttcctgaat tgctatgtgt c 21

<210> 118

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 118

gtcctgaatt gctatgtgtc t 21

<210> 119

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 119

gaagtcaact tcaatgtcgg a 21

<210> 120

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 120

gcagtaagtc aacttcaatg t 21

<210> 121

<211> 20

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 121
gaagttgact tactgaagaa 20

<210> 122
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 122
gtggagagag aattgaaaaa g 21

<210> 123
<211> 21
<212> DNA
<213> Artificial Sequence

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

<400> 123
gttcagactt gtctttcagc a 21

<210> 124
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 124
gacttgtctt tcagcaagga c 21

<210> 125
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 125

gatactcatc tttttcagt g 21

<210> 126

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 126

gcatactcat ctttttcagt g 21

<210> 127

<211> 20

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 127

gcatactcat ctttttcagt 20

<210> 128

<211> 20

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 128

ggcataactca tctttttcag 20

<210> 129

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 129

gagtcacatg gttcacacgg c 21

<210> 130

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 130

gacaaagtca catggttcac a

21

<210> 131

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 131

gtgggctgtg acaaagtcac a

21

<210> 132

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 132

gttaccacac ttaactatct t

21

<210> 133

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 133

gcttaccaca cttactatc t

21

<210> 134

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 134

gcacagccca agatagttaa g

21

<210> 135

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 135

gacagcccaa gatagttaag t

21

<210> 136

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 136

gcagcccaag atagttaagt g

21

<210> 137

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 137

attcgggccg agatgtctcg c

21

<210> 138

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 138

gtagcgcgag cacagctaag g 21

<210> 139

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 139

ctccgtggcc ttagctgtgc t 21

<210> 140

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 140

ctccaggcca gaaagagaga g 21

<210> 141

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 141

gtggccttag ctgtgctcgc g 21

<210> 142

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 142
atagcctcca ggccagaaag a 21
<210> 143
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 143
cttagctgtg ctgcgctac t 21
<210> 144
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 144
ctggatagcc tccaggccag a 21
<210> 145
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 145
ctgtgctcgc gctactctct c 21
<210> 146
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 146

ctcacgctgg atagcctcca g 21

<210> 147

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 147

ctactctctc tttctggcct g 21

<210> 148

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 148

gtaggagaga ctcacgctgg a 21

<210> 149

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 149

gtgtcttttc ccgatatcc t 21

<210> 150

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 150

gtgagtaaac ctgaatcttt g	21
<210> 151	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 151	
ttttcccgat attcctcagg t	21
<210> 152	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 152	
atgacgtgag taaacctgaa t	21
<210> 153	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 153	
ttcctcaggt actccaaaga t	21
<210> 154	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> Description of Artificial Sequence: Synthetic oligonucleotide	
<400> 154	
ctctgctgga tgacgtgagt a	21

<210> 155
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 155
 ctcaggtact ccaaagattc a 21

<210> 156
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 156
 attctctgct ggatgacgtg a 21

<210> 157
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 157
 gtactccaaa gattcaggtt t 21

<210> 158
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 158
 tttccattct ctgctggatg a 21

<210> 159

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 159

ctccaaagat tcaggtttac t 21

<210> 160

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 160

ttgactttcc attctctgct g 21

<210> 161

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 161

ctcacgtcat ccagcagaga a 21

<210> 162

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 162

atagcaattc aggaaatttg a 21

<210> 163

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 163

ttcctgaatt gctatgtgtc t 21

<210> 164

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 164

gtcaacttca atgtcggatg g 21

<210> 165

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 165

ctatgtgtct gggtttcatc c 21

<210> 166

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 166

ttcttcagta agtcaacttc a 21

<210> 167

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 167

atgtgtcttg gtttcatcca t

21

<210> 168

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 168

attcttcagt aagtcaactt c

21

<210> 169

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 169

gtctgggttt catccatccg a

21

<210> 170

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 170

ctccattctt cagtaagtca a

21

<210> 171

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide
 <400> 171
 ttctcatccat ccgacattga a 21
 <210> 172
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 172
 ttctctctcc attcttcagt a 21
 <210> 173
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 173
 atccatccga cattgaagtt g 21
 <210> 174
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 174
 ttcaattctc tctccattct t 21
 <210> 175
 <211> 21
 <212> DNA
 <213> Artificial Sequence

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

<400> 175
atccgacatt gaagttgact t 21
<210> 176
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 176
ctttttcaat tctctctcca t 21
<210> 177
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 177
ttgaagttga cttactgaag a 21
<210> 178
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 178
atgctccact ttttcaattc t 21
<210> 179
<211> 21
<212> DNA
<213> Artificial Sequence
<220><223> Description of Artificial Sequence: Synthetic
oligonucleotide
<400> 179

gttgacttac tgaagaatgg a 21

<210> 180

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 180

gtctgaatgc tccacttttt c 21

<210> 181

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 181

atggagagag aattgaaaaa g 21

<210> 182

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 182

cttgctgaaa gacaagtctg a 21

<210> 183

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 183

ttcagacttg tctttcagca a 21

<210> 184
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 184
 gtgtagtaca agagatagaa a 21
 <210> 185
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 185
 cttgtctttc agcaaggact g 21
 <210> 186
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 186
 attcagtgtgta gtacaagaga t 21
 <210> 187
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> Description of Artificial Sequence: Synthetic
 oligonucleotide
 <400> 187
 ctttcagcaa ggactggtct t 21
 <210> 188
 <211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 188

gtgaattcag ttagtagaa g

21

<210> 189

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 189

ctggctcttc tatctcttg a

21

<210> 190

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 190

tttttcagtg ggggtgaatt c

21

<210> 191

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 191

ctatctcttg tactacactg a

21

<210> 192

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 192

atactcatct ttttcagtgg g 21

<210> 193

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 193

ctacactgaa ttcaccccca c 21

<210> 194

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 194

ttcacacggc aggcatactc a 21

<210> 195

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 195

attcaccccc actgaaaaag a 21

<210> 196

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 196

gtcacatggt tcacacggca g

21

<210> 197

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 197

ctgaaaaaga tgagtatgcc t

21

<210> 198

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 198

ctgtgacaaa gtcacatggt t

21

<210> 199

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 199

atgagtatgc ctgccgtgtg a

21

<210> 200

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 200

ctatcttggg ctgtgacaaa g 21

<210> 201

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 201

gtatgcctgc cgtgtgaacc a 21

<210> 202

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 202

ttaactatct tgggctgtga c 21

<210> 203

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 203

gtgtgaacca tgtgactttg t 21

<210> 204

<211> 21

<212> DNA

<213> Artificial Sequence

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

oligonucleotide

<400> 204

ttacccact taactatctt g 21

<210> 205

<211> 21

<212> DNA

<213> Artificial Sequence

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

<400> 205

atgtgacttt gtcacagccc a 21

<210> 206

<211> 28

<212> DNA

<213> Artificial Sequence

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

<400> 206

tcaagctggt cgagaaaagc tttgaaac 28

<210> 207

<211> 28

<212> DNA

<213> Artificial Sequence

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

<400> 207

aacagtaag acaggggtct agcctggg 28

<210> 208

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 208

Gln Ser Gly Asn Arg Thr Thr

1 5

<210> 209

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 209

Arg Ser Ala Asn Leu Ala Arg

1 5

<210> 210

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 210

Asp Arg Ser Ala Leu Ala Arg

1 5

<210> 211

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 211

Arg Ser Asp Val Leu Ser Glu

1 5

<210> 212

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 212

Lys His Ser Thr Arg Arg Val

1 5

<210> 213

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 213

Thr Met His Gln Arg Val Glu

1 5

<210> 214

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 214

Thr Ser Gly His Leu Ser Arg

1 5

<210> 215

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 215

Arg Ser Asp His Leu Thr Gln

1 5

<210> 216

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 216

Asp Ser Ala Asn Leu Ser Arg

1 5

<210

> 217

<211> 28

<212> DNA

<213> Artificial Sequence

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

<400> 217

ctcctgaaag tggccgggtt taatctgc 28

<210> 218

<211> 28

<212> DNA

<213> Artificial Sequence

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

<400> 218

aggattcgga acccaatcac tgacaggt 28

<210> 219

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 219

Arg Ser Asp His Leu Ser Thr

1 5

<210> 220

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 220

Asp Arg Ser His Leu Ala Arg

1 5

<210> 221

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 221

Leu Lys Gln His Leu Asn Glu

1 5

<210> 222

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 222

His Asn Ser Ser Leu Lys Asp

1 5

<210> 223

<211> 7

<212> PRT

<213> Artificial Sequence

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

peptide

<400> 223

Thr Ser Ser Asn Arg Lys Thr

1 5

<210> 224

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 224

Leu Gln Gln Thr Leu Ala Asp

1 5

<210> 225

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 225

Arg Arg Glu Asp Leu Ile Thr

1 5

<210> 226

<211> 7

<212> PRT

<213> Artificial Sequence

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

<400> 226

Thr Ser Ser Asn Leu Ser Arg

1 5

<210> 227

<211> 9

<212> PRT

<213> Unknown

<220><223> Description of Unknown: "LAGLIDADG" family motif
peptide

<400> 227

Leu Ala Gly Leu Ile Asp Ala Asp Gly

1

5