(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0105947
(43) 공개일자 2021년08월27일

(51) 국제특허분류(Int. Cl.)
*C12N 15/85* (2006.01) *A61K 48/00* (2006.01)
*A61P 35/00* (2006.01) *C07K 14/705* (2006.01)
*C07K 14/725* (2006.01) *C07K 16/28* (2006.01)
*C07K 16/30* (2006.01)
(52) CPC특허분류
*C12N 15/85* (2013.01)
*A61K 48/00* (2013.01)
(21) 출원번호 10-2021-7022767
(22) 출원일자(국제) 2019년12월20일
심사청구일자 없음
(85) 번역문제출일자 2021년07월19일
(86) 국제출원번호 PCT/US2019/067758
(87) 국제공개번호 WO 2020/132396
국제공개일자 2020년06월25일
(30) 우선권주장
62/783,133 2018년12월20일 미국(US)
*(뒷면에 계속)*

(71) 출원인
**포세이다 테라퓨틱스, 인크.**
미국 92121 캘리포니아주 샌디에고 타운 센터 드라이브 9390 스위트 200
(72) 발명자
**오스터텍 에릭 엠.**
미국 92121 캘리포니아주 샌디에고 스윗 200 타운 센터 드라이브 9390
**셰드록 드본**
미국 92121 캘리포니아주 샌디에고 스윗 200 타운 센터 드라이브 9390
(74) 대리인
**김진희, 김태홍**

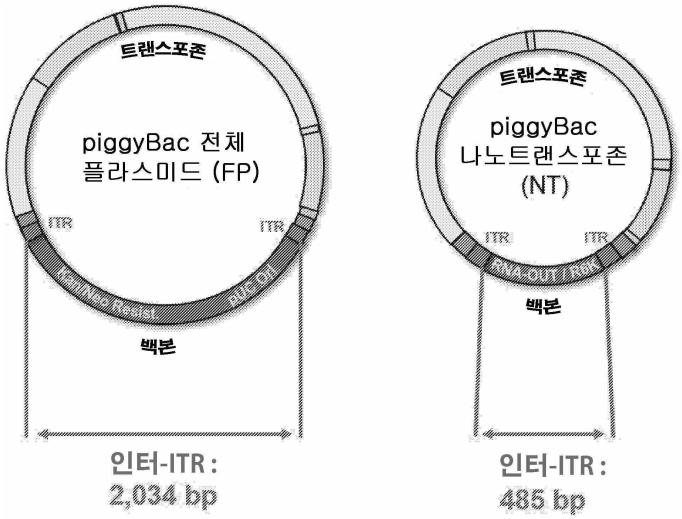
전체 청구항 수 : 총 54 항

(54) 발명의 명칭 **나노트랜스포존 조성물 및 사용 방법**

## (57) 요 약

(a) 제1 역위 말단 반복(ITR), (b) 제2 ITR 및 (c) 인트라-ITR 서열을 포함하는 제1 핵산 서열로서, 인트라-ITR 서열은 트랜스포존 서열을 포함하는 것인 제1 핵산 서열; 및 인터-ITR 서열을 포함하는 제2 핵산 서열로서, 인터-ITR 서열의 길이는 종점을 포함하여 1 내지 600개의 뉴클레오타이드인 것인 제2 핵산 서열을 포함하는 조성물이 개시된다. 바람직하게는, 조성물은 나노트랜스포존이다.

*대 표 도* - 도1



(52) CPC특허분류
***A61P 35/00*** (2018.01)
***C07K 14/7051*** (2013.01)
***C07K 14/70517*** (2013.01)
***C07K 14/70578*** (2013.01)
***C07K 16/2878*** (2013.01)
***C07K 16/3069*** (2013.01)
*C12N 2800/90* (2013.01)

(30) 우선권주장
62/815,335 2019년03월07일 미국(US)
62/815,845 2019년03월08일 미국(US)

# 명 세 서

## 청구범위

### 청구항 1

(a) 제1 역위 말단 반복(inverted terminal repeat, ITR), (b) 제2 ITR 및 (c) 인트라-ITR(intra-ITR) 서열을 포함하는 제1 핵산 서열로서, 인트라-ITR 서열은 트랜스포존(transposon) 서열을 포함하는 것인 제1 핵산 서열; 및

인터-ITR(inter-ITR) 서열을 포함하는 제2 핵산 서열로서, 인터-ITR 서열의 길이는 종점을 포함하여 1 내지 600개의 뉴클레오타이드인 것인 제2 핵산 서열

을 포함하는 조성물.

### 청구항 2

제1항에 있어서, 인터-ITR 서열의 길이는 종점을 포함하여 1 내지 100개의 뉴클레오타이드인 것인 조성물.

### 청구항 3

제1항에 있어서, 제1 핵산 서열은 복제 원점 서열을 추가로 포함하는 것인 조성물.

### 청구항 4

제1항에 있어서, 제2 핵산 서열은 복제 원점 서열을 추가로 포함하는 것인 조성물.

### 청구항 5

제3항 또는 제4항에 있어서, 복제 원점 서열의 길이는 1 내지 450개의 뉴클레오타이드인 것인 조성물.

### 청구항 6

제5항에 있어서, 복제 원점 서열은 R6K 복제 원점을 포함하는 것인 조성물.

### 청구항 7

제1항에 있어서, 제1 핵산은 제1 선택 마커를 코딩하는 서열을 추가로 포함하는 것인 조성물.

### 청구항 8

제1항에 있어서, 제2 핵산 서열은 제1 선택 마커를 코딩하는 서열을 추가로 포함하는 것인 조성물.

### 청구항 9

제7항 또는 제8항에 있어서, 제1 선택 마커의 길이는 1 내지 200개의 뉴클레오타이드인 것인 조성물.

### 청구항 10

제7항 또는 제8항에 있어서, 제1 선택 마커는 수크로스 선택 마커인 것인 조성물.

### 청구항 11

제7항 또는 제8항에 있어서, 수크로스 선택 마커는 RNA-OUT 선택 마커인 것인 조성물.

### 청구항 12

제1항에 있어서, 제1 핵산 서열은 재조합 부위, 절제 부위, 결찰 부위, 또는 이의 조합을 포함하지 않는 것인 조성물.

### 청구항 13

제1항에 있어서, 제2 핵산 서열은 재조합 부위, 절제 부위, 결찰 부위, 또는 이의 조합을 포함하지 않는 것인 조성물.

**청구항 14**

제1항에 있어서, 제1 핵산 서열은 외래 DNA를 코딩하는 서열을 포함하지 않는 것인 조성물.

**청구항 15**

제1항에 있어서, 제2 핵산 서열은 외래 DNA를 코딩하는 서열을 포함하지 않는 것인 조성물.

**청구항 16**

제1항에 있어서, 제1 핵산 서열은 적어도 하나의 외인성 서열 및 포유동물 세포에서 외인성 서열을 발현할 수 있는 프로모터를 코딩하는 서열을 추가로 포함하는 것인 조성물.

**청구항 17**

제16항에 있어서, 제1 핵산 서열은 인슐레이터(insulator)를 코딩하는 적어도 하나의 서열을 추가로 포함하는 것인 조성물.

**청구항 18**

제16항에 있어서, 제1 핵산 서열은 폴리아데노신(polyA) 서열을 추가로 포함하는 것인 조성물.

**청구항 19**

제16항에 있어서, 포유동물 세포에서 외인성 서열을 발현할 수 있는 프로모터를 코딩하는 서열은 인간 세포에서 외인성 서열을 발현할 수 있는 것인 조성물.

**청구항 20**

제19항에 있어서, 프로모터는 항시성 프로모터인 것인 조성물.

**청구항 21**

제19항에 있어서, 프로모터는 유도성 프로모터인 것인 조성물.

**청구항 22**

제16항에 있어서, 적어도 하나의 외인성 서열은 비자연발생 항원 수용체를 코딩하는 서열, 치료적 폴리펩타이드를 코딩하는 서열, 또는 이의 조합을 포함하는 것인 조성물.

**청구항 23**

제22항에 있어서, 비자연발생 항원 수용체는 키메라 항원 수용체(CAR)를 포함하는 것인 조성물.

**청구항 24**

제23항에 있어서, CAR은

(a) 항원 인식 영역을 포함하는 엑토도메인(ectodomain),

(b) 막관통 도메인, 및

(c) 적어도 하나의 공자극 도메인을 포함하는 엔도도메인(endodomain)

을 포함하는 것인 조성물.

**청구항 25**

제24항에 있어서, 항원 인식 영역은 적어도 하나의 단일 사슬 가변 단편(scFv), 단일 도메인 항체, 센티린(Centyrin), 또는 이의 조합을 포함하는 것인 조성물.

**청구항 26**

제25항에 있어서, 단일 도메인 항체는 VHH 또는 VH인 것인 조성물.

**청구항 27**

제24항에 있어서, 항원 인식 영역은 적어도 하나의 항-BCMA 센티린을 포함하는 것인 조성물.

**청구항 28**

제27항에 있어서, 항-BCMA 센티린은 서열번호: 29의 아미노산 서열을 포함하는 것인 조성물.

**청구항 29**

제24항에 있어서, 항원 인식 영역은 적어도 하나의 항-BCMA VH를 포함하는 것인 조성물.

**청구항 30**

제29항에 있어서, 항-BCMA VH는 서열번호: 97의 아미노산 서열을 포함하는 것인 조성물.

**청구항 31**

제24항에 있어서, 항원 인식 영역은 적어도 하나의 항-PSMA 센티린을 포함하는 것인 조성물.

**청구항 32**

제31항에 있어서, 항-PSMA 센티린은 서열번호: 94의 아미노산 서열을 포함하는 것인 조성물.

**청구항 33**

제24항에 있어서, 엑토도메인은 신호 펩타이드를 추가로 포함하는 것인 조성물.

**청구항 34**

제24항에 있어서, CAR은 항원 인식 영역과 막관통 도메인 사이에 힌지 영역을 추가로 포함하는 것인 조성물.

**청구항 35**

제24항에 있어서, 막관통 도메인은 CD8 막관통 도메인을 코딩하는 서열을 포함하는 것인 조성물.

**청구항 36**

제24항에 있어서, 적어도 하나의 공자극 도메인은 CD3ζ 공자극 도메인, 4-1BB 공자극 도메인, 또는 이의 조합을 포함하는 것인 조성물.

**청구항 37**

제24항에 있어서, 적어도 하나의 공자극 도메인은 CD3ζ 공자극 도메인 및 4-1BB 공자극 도메인을 포함하고, 4-1BB 공자극 도메인은 막관통 도메인과 CD3ζ 공자극 도메인 사이에 위치하는 것인 조성물.

**청구항 38**

제XX항에 있어서, 적어도 하나의 외인성 서열은 유도성 아폽토시스 촉진 폴리펩타이드를 코딩하는 서열, 제2 선택 마커를 코딩하는 서열, 키메라 자극 수용체(CSR)를 코딩하는 서열, 트랜스포사제 효소를 코딩하는 서열, 자가 절단 펩타이드를 코딩하는 서열, 또는 이의 조합을 포함하는 것인 조성물.

**청구항 39**

제38항에 있어서, 제2 선택 마커는 디하이드로폴레이트 환원효소(DHFR) 뮤테인 효소를 코딩하는 서열을 포함하는 것인 조성물.

**청구항 40**

제1항에 있어서, 조성물은 트랜스포존인 것인 조성물.

**청구항 41**

제40항에 있어서, 트랜스포존은 piggyBac 트랜스포존인 것인 조성물.

**청구항 42**

제1항의 조성물을 코딩하는 핵산 서열을 포함하는 폴리뉴클레오타이드.

**청구항 43**

제1항의 조성물을 포함하는 세포.

**청구항 44**

복수의 세포 집단이 제24항의 CAR을 발현하도록 변형된 것인 세포 집단.

**청구항 45**

제44항에 있어서, 복수의 변형된 세포는 복수의 변형된 면역 세포인 것인 세포 집단.

**청구항 46**

제44항에 있어서, 복수의 변형된 세포는 복수의 변형된 T-세포인 것인 세포 집단.

**청구항 47**

제44항에 있어서, 복수의 변형된 T-세포의 적어도 50%는 CD45RA 및 CD62L을 포함하는 하나 이상의 세포 표면 마커(들)를 발현하고 CD45RO를 포함하는 하나 이상의 세포 표면 마커(들)를 발현하지 않는 것인 세포 집단.

**청구항 48**

제43항의 세포를 포함하는 조성물.

**청구항 49**

제44항의 세포 집단을 포함하는 조성물.

**청구항 50**

제1항, 제48항 또는 제49항 중 어느 한 항의 조성물 및 약학적으로 허용가능한 담체를 포함하는 약학 조성물.

**청구항 51**

암 치료를 필요로 하는 대상체에서 암 치료에 사용하기 위한, 제1항, 제48항 또는 제49항 중 어느 한 항의 조성물 또는 제50항의 약학 조성물.

**청구항 52**

제51항에 있어서, 암은 BCMA 양성 암 또는 PSMA 양성 암인 것인 용도.

**청구항 53**

제51항에 있어서, 암은 원발성 종양, 전이성 암, 다중 내성 암, 진행성 종양 또는 재발성 암인 것인 용도.

**청구항 54**

제51항에 있어서, 암은 폐암, 뇌암, 두경부암, 유방암, 피부암, 간암, 췌장암, 위암, 결장암, 직장암, 자궁암, 자궁경부암, 난소암, 전립선암, 고환암, 피부암, 식도암, 림프종, 백혈병, 급성 백혈병, 급성 림프구성 백혈병(acute lymphoblastic leukemia, ALL), 급성 림프성 백혈병(acute lymphocytic leukemia), 급성 골수성 백혈병(acute myeloid leukemia, AML), 급성 골수성 백혈병(acute myelogenous leukemia), 만성 골수성 백혈병(chronic myelocytic leukemia, CML), 만성 림프성 백혈병(chronic lymphocytic leukemia, CLL), 털세포 백혈

병(hairy cell leukemia), 골수이형성 증후군(myelodyplastic syndrome, MDS), 호지킨병, 비호지킨 림프종, 또는 다발성 골수종인 것인 용도.

## 발명의 설명

### 기 술 분 야

[0001] 관련 출원에 대한 상호참조

[0002] 본 출원은 2018년 12월 20일에 출원된 U.S.S.N. 제62/783,133호; 2019년 3월 7일에 출원된 U.S.S.N. 제62/815,335호 및 2019년 3월 8일에 출원된 U.S.S.N. 제62/815,845호의 우선권 및 이익을 주장한다. 이들 각 출원의 내용은 그 전체가 참조로 본원에 포함되어 있다.

[0003] 서열 목록의 참조에 의한 포함

[0004] 2019년 12월 19일에 생성되고 295 MB 크기인 "POTH-047_001WO Seq Listing_ST25.txt"라는 명칭의 파일의 내용은 그 전체가 참조로 본원에 포함되어 있다.

[0005] 본 개시내용의 분야

[0006] 본 개시내용은 분자 생물학, 및 보다 구체적으로, 나노트랜스포존(nanotransposon), 나노트랜스포존을 포함하는 세포 조성물, 이의 제조 방법 및 사용 방법에 관한 것이다.

### 배 경 기 술

[0007] 유전자 요법에서 사용하기 위한 개선된 전위(transposition)의 조성물 및 방법에 대한 당업계의 오랜 충족되지 않은 요구가 있었다. 본 개시내용은 특히 유전자 요법을 위해 세포를 변형시키는 방법으로서 인간 세포에서 사용하기 위해 전위의 효능을 개선하기 위한 트랜스포존(transposon) 서열을 갖는 벡터에 대한 비자연발생 구조 개선을 포함하는 나노트랜스포존 조성물, 이들 조성물의 제조 방법 및 사용 방법을 제공한다.

### 발명의 내용

[0008] 본 개시내용은 (a) 제1 역위 말단 반복(inverted terminal repeat, ITR), (b) 제2 ITR 및 (c) 인트라-ITR(intra-ITR) 서열을 포함하는 제1 핵산 서열로서, 인트라-ITR 서열은 트랜스포존(transposon) 서열을 포함하는 것인 제1 핵산 서열; 및 인터-ITR(inter-ITR) 서열을 포함하는 제2 핵산 서열로서, 인터-ITR 서열의 길이는 종점을 포함하여 1 내지 600개의 뉴클레오타이드인 것인 제2 핵산 서열을 포함하는 조성물을 제공한다. 바람직한 양태에서, 인터-ITR 서열의 길이는 종점을 포함하여 1 내지 100개의 뉴클레오타이드이다. 조성물은 트랜스포존일 수 있거나 나노트랜스포존일 수 있다. 바람직한 양태에서, 트랜스포존은 piggyBac 트랜스포존이다.

[0009] 제1 핵산 서열 및/또는 제2 핵산 서열은 복제 원점 서열을 추가로 포함할 수 있다. 복제 원점 서열의 길이는 1 내지 450개의 뉴클레오타이드일 수 있다. 복제 원점 서열은 R6K 복제 원점을 포함할 수 있다.

[0010] 제1 핵산 서열 및/또는 제2 핵산 서열은 제1 선택 마커를 코딩하는 서열을 추가로 포함할 수 있다. 제1 선택 마커의 길이는 1 내지 200개의 뉴클레오타이드일 수 있다. 제1 선택 마커는 수크로스 선택 마커일 수 있다. 바람직한 양태에서, 수크로스 선택 마커는 RNA-OUT 선택 마커이다

[0011] 제1 핵산 서열 및/또는 제2 핵산 서열은 재조합 부위, 절제 부위, 결찰 부위, 또는 이의 조합을 포함하지 않을 수 있다. 제1 핵산 서열 및/또는 제2 핵산 서열은 외래 DNA를 코딩하는 서열을 포함하지 않을 수 있다.

[0012] 제1 핵산 서열은 적어도 하나의 외인성 서열 및 포유동물 세포에서 외인성 서열을 발현할 수 있는 프로모터를 코딩하는 서열을 추가로 포함할 수 있다. 제1 핵산 서열은 인슐레이터(insulator)를 코딩하는 적어도 하나의 서열을 추가로 포함할 수 있다. 제1 핵산 서열은 폴리아데노신(polyA) 서열을 추가로 포함할 수 있다. 포유동물 세포에서 외인성 서열을 발현할 수 있는 프로모터를 코딩하는 서열은 인간 세포에서 외인성 서열을 발현할 수 있다. 프로모터는 항시성 프로모터 또는 유도성 프로모터일 수 있다.

[0013] 적어도 하나의 외인성 서열은 비자연발생 항원 수용체를 코딩하는 서열, 치료적 폴리펩타이드를 코딩하는 서열, 또는 이의 조합을 포함할 수 있다. 바람직한 양태에서, 비자연발생 항원 수용체는 키메라 항원 수용체(CAR)를 포함한다. CAR은 (a) 항원 인식 영역을 포함하는 엑토도메인(ectodomain), (b) 막관통 도메인, 및 (c) 적어도 하나의 공자극 도메인을 포함하는 엔도도메인(endodomain)을 포함할 수 있다. 항원 인식 영역은 적어도 하나의

단일 사슬 가변 단편(scFv), 단일 도메인 항체, 센티린(Centyrin), 또는 이의 조합을 포함할 수 있다. 단일 도메인 항체는 VHH 또는 VH일 수 있다.

[0014] 항원 인식 영역은 적어도 하나의 항-BCMA 센티린을 포함할 수 있다. 바람직하게는, 항-BCMA 센티린은 서열번호: 29의 아미노산 서열을 포함한다. 항원 인식 영역은 적어도 하나의 항-BCMA VH를 포함할 수 있다. 바람직하게는, 항-BCMA VH는 서열번호: 97의 아미노산 서열을 포함한다. 항원 인식 영역은 적어도 하나의 항-PSMA 센티린을 포함할 수 있다. 바람직하게는, 항-PSMA 센티린은 서열번호: 94의 아미노산 서열을 포함한다.

[0015] 엑토도메인은 신호 펩타이드를 추가로 포함할 수 있다. CAR은 항원 인식 영역 및 막관통 도메인 사이에 힌지 영역을 추가로 포함한다. 막관통 도메인은 CD8 막관통 도메인을 코딩하는 서열을 포함할 수 있다. 적어도 하나의 공자극 도메인은 CD3ζ 공자극 도메인, 4-1BB 공자극 도메인, 또는 이의 조합을 포함할 수 있다. 적어도 하나의 공자극 도메인은 CD3ζ 공자극 도메인 및 4-1BB 공자극 도메인을 포함할 수 있고, 4-1BB 공자극 도메인은 막관통 도메인 및 CD3ζ 공자극 도메인 사이에 위치한다. 적어도 하나의 외인성 서열은 유도성 아폽토시스 촉진(proapoptotic) 폴리펩타이드를 코딩하는 서열, 제2 선택 마커를 코딩하는 서열, 키메라 자극 수용체(CSR)를 코딩하는 서열, 트랜스포사제(transposase) 효소를 코딩하는 서열, 자가 절단 펩타이드를 코딩하는 서열, 또는 이의 조합을 포함할 수 있다. 제2 선택 마커는 디하이드로폴레이트 환원효소(DHFR) 뮤테인 효소를 코딩하는 서열을 포함할 수 있다.

[0016] 본 개시내용은 또한 본원에 개시된 바와 같은 조성물(예컨대, 트랜스포존 또는 나노트랜스포존)을 코딩하는 핵산 서열을 포함하는 폴리뉴클레오타이드 및/또는 본원에 개시된 바와 같은 CAR을 코딩하는 핵산 서열을 포함하는 폴리뉴클레오타이드를 제공한다.

[0017] 본 개시내용은 또한 본원에 개시된 바와 같은 조성물(예컨대, 트랜스포존 또는 나노트랜스포존)을 포함하는 세포를 제공한다. 본 개시내용은 또한 복수의 집단이 본원에 개시된 바와 같은 CAR 또는 조성물(예컨대, 트랜스포존 또는 나노트랜스포존)을 발현하도록 변형된 복수의 집단을 제공한다. 일 양태에서, 복수의 변형된 세포는 복수의 변형된 면역 세포이다. 일 양태에서, 복수의 변형된 세포는 복수의 변형된 T-세포이다. 일 양태에서, 복수의 변형된 T-세포의 적어도 50%는 CD45RA 및 CD62L을 포함하는 하나 이상의 세포 표면 마커(들)를 발현하지만 CD45RO를 포함하는 하나 이상의 세포 표면 마커(들)를 발현하지 않는다.

[0018] 본 개시내용은 또한 본원에 개시된 바와 같은 CAR 또는 조성물(예컨대, 트랜스포존 또는 나노트랜스포존)을 포함하고 약학적으로 허용가능한 담체를 추가로 포함하는 약학 조성물을 제공한다.

[0019] 본 개시내용은 또한 본원에 개시된 바와 같은 조성물(예컨대, 트랜스포존 또는 나노트랜스포존), CAR, 세포, 세포 집단 또는 약학 조성물 중 어느 것의 치료적 유효량을 투여함으로써 이를 필요로 하는 대상체에서 증식 장애를 치료하는 방법을 제공한다. 일 양태에서, 증식 장애는 암이다. 암은 BCMA 양성 암 또는 PSMA 양성 암일 수 있다. 암은 원발성 종양, 전이성 암, 다중 내성 암(multiply resistant cancer), 진행성 종양 또는 재발성 암일 수 있다. 암은 고형 종양 또는 혈액암일 수 있다. 암은 폐암, 뇌암, 두경부암, 유방암, 피부암, 간암, 췌장암, 위암, 결장암, 직장암, 자궁암, 자궁경부암, 난소암, 전립선암, 고환암, 피부암, 식도암, 림프종, 백혈병, 급성 백혈병, 급성 림프구성 백혈병(acute lymphoblastic leukemia, ALL), 급성 림프성 백혈병(acute lymphocytic leukemia), 급성 골수성 백혈병(acute myeloid leukemia, AML), 급성 골수성 백혈병(acute myelogenous leukemia), 만성 골수성 백혈병(chronic myelocytic leukemia, CML), 만성 림프성 백혈병(chronic lymphocytic leukemia, CLL), 털세포 백혈병(hairy cell leukemia), 골수이형성 증후군(myelodyplastic syndrome, MDS), 호지킨병, 비호지킨 림프종, 또는 다발성 골수종일 수 있다.

## *도면의 간단한 설명*

[0020] 특허 또는 출원 파일은 컬러로 제작된 적어도 하나의 도면을 포함한다. 컬러 도면(들)이 있는 본 특허 또는 특허 출원의 사본은 요청하고 필요한 비용을 지불하면 특허청에 의해 제공될 것이다.

도 1은 piggyBac 전체 플라스미드 및 piggyBac 나노트랜스포존(NT)의 맵을 비교하는 한 쌍의 개략도이다.

도 2는 인간 범(pan) T 세포에서 piggyBac NT를 이용한 개선된 전위를 도시하는 그래프이다.

도 3은 piggyBac NT 및 piggyBac 짧은 NT의 맵을 비교하는 한 쌍의 개략도이다.

도 4는 인간 범 T 세포에서 piggyBac 전위가 인터-ITR 서열을 감소시킴으로써(예컨대, ITR에 측접하는 거리를 감소시킴으로써) 향상된다는 것을 보여주는 그래프이다.

도 5는 인간 범 T 세포에서 항-BCMA 키메라 항원 수용체(CAR) NT 및 항-PSMA CAR NT를 이용하여 증가된 전위를 보여주는 한 쌍의 그래프이다.

도 6은 항-BCMA CAR NT 또는 항-PSMA CAR NT를 사용하여 생산된 인간 CAR-T 세포가 표적 종양 세포를 사멸시킬 수 있다는 것을 보여주는 일련의 그래프이다.

도 7은 항-BCMA CAR NT 또는 항-PSMA CAR NT를 사용하여 생산된 인간 CAR-T 세포가 표현형 조성에서 대등하다는 것을 보여주는 일련의 그래프이다.

도 8은 항-BCMA CAR NT 또는 항-PSMA CAR NT를 사용하여 생산된 인간 CAR-T 세포가 유사한 통합된 카피 수를 갖는다는 것을 보여주는 일련의 그래프이다.

도 9는 단량체성 NT 순도가 인간 범 T 세포에서 전위 효율과 관련된다는 것을 입증하는 겔 전기영동 분석의 사진이다.

도 10은 단량체성 NT 순도가 인간 범 T 세포에서 전위 효율과 관련된다는 것을 보여주는 한 쌍의 그래프이다.

도 11은 쥣과 이종이식편 모델을 사용하여 스트레스 용량에서 전장 플라스미드(FLP) 대 NT에 의해 전달될 때 P-PSMA-101 트랜스포존의 임상전 평가를 보여주는 개략도이다.

도 12는 FLP 대 NT에 의해 전달될 때 P-PSMA-101 트랜스포존이 처리된 마우스의 종양 부피 평가를 보여주는 일련의 그래프이다.

도 13은 BCMA 센티린 CAR(CARTyrin)을 코딩하는 P-BCMA-101 piggyBac NT를 도시하는 개략도이다. 나노트랜스포존은 ITR #1, 인슐레이터 #1, EF1알파 프로모터, BMCA CARTyrin, SV40 PA, 인슐레이터 #2, 및 ITR #2를 코딩한다. 서열은 또한 나노트랜스포존 요소 RNA-OUT 및 R6K 원점을 코딩한다.

도 14는 PSMA CARTyrin을 코딩하는 P-PSMA-101 piggyBac NT를 도시하는 개략도이다. 나노트랜스포존은 ITR #1, 인슐레이터 #1, EF1알파 프로모터, PSMA CARTyrin, SV40 PA, 인슐레이터 #2, 및 ITR #2를 코딩한다. 서열은 또한 나노트랜스포존 요소 RNA-OUT 및 R6K 원점을 코딩한다.

도 15는 BCMA VH CAR(VCAR)을 코딩하는 P-BCMA-ALLO1 piggyBac 나노트랜스포존을 도시하는 개략도이다. 나노트랜스포존은 ITR #1, 인슐레이터 #1, EF1알파 프로모터, BMCA VCAR, SV40 PA, 인슐레이터 #2, 및 ITR #2를 코딩한다. 서열은 또한 나노트랜스포존 요소 RNA-OUT 및 R6K 원점을 코딩한다.

임의의 상호참조되거나 관련된 특허 또는 출원을 포함하는 본원에 인용된 모든 문서는 명시적으로 배제되거나 제한되지 않는 한 모든 목적을 위해 그 전체가 본원에 참조로 포함된다. 임의의 문서의 인용은 그것이 본원에 개시되거나 청구된 임의의 발명에 대한 선행 기술이라는 것을 인정하거나, 단독으로 또는 임의의 다른 참고문헌 또는 참고문헌들과 조합하여 이러한 임의의 발명을 교시, 제시 또는 개시한다는 것을 인정하는 것이 아니다. 또한, 이 문서에서의 용어의 임의의 의미 또는 정의가 참조로 포함된 문서에서의 동일한 용어의 임의의 의미 또는 정의와 충돌하는 경우, 이 문서에서 상기 용어에 할당된 의미 또는 정의가 우선할 것이다.

## *발명을 실시하기 위한 구체적인 내용*

[0021] 본 개시내용은 나노트랜스포존(nanotransposon), 나노트랜스포존을 포함하는 조성물 및 세포, 나노트랜스포존을 제조하는 방법 및 본원에 기재된 나노트랜스포존, 조성물 및 세포를 사용하는 방법을 제공한다.

[0022] 본 개시내용의 나노트랜스포존은 나노트랜스포존의 인터-ITR 서열을 최소화하여 제1 및 제2 ITR 서열을 가능한 가깝게 함으로써 전위 효능 및 효율을 증가시키도록 설계된다. 본 개시내용의 나노트랜스포존 및 나노트랜스포존을 포함하는 조성물은 모든 세포 유형에서 효과적이지만, 이들은 인간 세포에서 사용하기에 특히 효과적이다. 본원에 기재된 바와 같이, 본 개시내용의 나노트랜스포존은 전위를 증가시키고, 결과적으로 인간 세포로의 유전자 전달을 복수의 세포에서 충분히 높은 백분율의 세포로 증가시키는 데 사용될 수 있다.

[0023] 이론에 구속되기를 바라지 않지만, 인터-ITR 서열 또는 거리를 최소화함으로써, 상응하는 트랜스포사제(transposase)가 두 ITR 서열을 더 합칠 수 있어 나노트랜스포존로부터 인트라-ITR 서열의 절제를 증가시키고/또는 인트라-ITR 서열을 표적 부위로 통합시키는 것을 증가시킬 수 있다. 또한, 본 개시내용의 바람직한 양태에서, 나노트랜스포존, 이의 백본 및/또는 인터-ITR 서열은 외래 DNA 서열을 포함하지 않는다. 외래 DNA의 부족은 특히 비-나노트랜스포존과 비교할 때 전위 효능 및 효율을 추가로 개선한다.

[0024] **본 개시내용의 조성물**

[0025] 본 개시내용은 (a) 제1 역위 말단 반복(ITR) 또는 제1 ITR을 코딩하는 서열, (b) 제2 ITR 또는 제2 ITR을 코딩하는 서열, 및 (c) 인트라-ITR 서열 또는 인트라-ITR을 코딩하는 서열을 포함하는 제1 핵산 서열로서, 인트라-ITR 서열은 트랜스포존 서열 또는 트랜스포존을 코딩하는 서열을 포함하는 것인 제1 핵산 서열; 및 인터-ITR 서열 또는 인터-ITR을 코딩하는 서열을 포함하는 제2 핵산 서열로서, 인터-ITR 서열의 길이는 700개 이하의 뉴클레오타이드인 것인 제2 핵산 서열을 포함하는 조성물을 제공한다. 제2 핵산 서열은 또한 본원에서 백본 영역 또는 비통합 영역으로 지칭된다. 일 양태에서, 조성물은 원형 DNA 또는 선형 DNA이다. 일 양태에서, 조성물은 플라스미드 또는 벡터이다. 일 양태에서, 조성물은 트랜스포존이다. 바람직한 양태에서, 조성물은 나노트랜스포존이다.

[0026] 일부 양태에서, 인터-ITR 서열의 길이는 650개 이하의 뉴클레오타이드, 600개 이하의 뉴클레오타이드, 550개 이하의 뉴클레오타이드, 500개 이하의 뉴클레오타이드, 450개 이하의 뉴클레오타이드, 400개 이하의 뉴클레오타이드, 350개 이하의 뉴클레오타이드, 300개 이하의 뉴클레오타이드, 250개 이하의 뉴클레오타이드, 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 100개의 이하의 뉴클레오타이드, 50개 이하의 뉴클레오타이드, 25개 이하의 뉴클레오타이드, 또는 10개 이하의 뉴클레오타이드이다. 일부 양태에서, 제2 핵산 서열의 길이는 700개 이하의 뉴클레오타이드, 650개 이하의 뉴클레오타이드, 600개 이하의 뉴클레오타이드, 550개 이하의 뉴클레오타이드, 500개 이하의 뉴클레오타이드, 450개 이하의 뉴클레오타이드, 400개 이하의 뉴클레오타이드, 350개 이하의 뉴클레오타이드, 300개 이하의 뉴클레오타이드, 250개 이하의 뉴클레오타이드, 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 100개 이하의 뉴클레오타이드, 50개 이하의 뉴클레오타이드, 25개 이하의 뉴클레오타이드, 또는 10개 이하의 뉴클레오타이드이다.

[0027] 본 개시내용은 (a) 제1 역위 말단 반복(ITR) 또는 제1 ITR을 코딩하는 서열, (b) 제2 ITR 또는 제2 ITR을 코딩하는 서열, 및 (c) 인트라-ITR 서열 또는 인트라-ITR을 코딩하는 서열을 포함하는 제1 핵산 서열로서, 인트라-ITR 서열은 트랜스포존 서열 또는 트랜스포존을 코딩하는 서열을 포함하는 것인 제1 핵산 서열; 및 인터-ITR 서열 또는 인터-ITR을 코딩하는 서열을 포함하는 제2 핵산 서열로서, 인터-ITR 서열의 길이는 종점을 포함하여 1 내지 700개의 뉴클레오타이드인 것인 제2 핵산 서열을 포함하는 조성물을 제공한다. 제2 핵산 서열은 또한 본원에서 백본 영역 또는 비통합 영역으로 지칭된다. 일 양태에서, 조성물은 원형 DNA 또는 선형 DNA이다. 일 양태에서, 조성물은 플라스미드 또는 벡터이다. 일 양태에서, 조성물은 트랜스포존이다. 바람직한 양태에서, 조성물은 나노트랜스포존이다.

[0028] 일부 양태에서, 인터-ITR 서열의 길이는 1 내지 650개의 뉴클레오타이드, 1 내지 600개의 뉴클레오타이드, 1 내지 550개의 뉴클레오타이드, 1 내지 500개의 뉴클레오타이드, 1 내지 450개의 뉴클레오타이드, 1 내지 400개의 뉴클레오타이드, 1 내지 350개의 뉴클레오타이드, 1 내지 300개의 뉴클레오타이드, 1 내지 250개의 뉴클레오타이드, 1 내지 200개의 뉴클레오타이드, 1 내지 150개의 뉴클레오타이드, 1 내지 100개의 뉴클레오타이드, 1 내지 50개의 뉴클레오타이드, 1 내지 25개의 뉴클레오타이드 또는 1 내지 10개의 뉴클레오타이드이고, 각 범위는 종점을 포함한다. 일부 양태에서, 제2 핵산 서열의 길이는 1 내지 650개의 뉴클레오타이드, 1 내지 600개의 뉴클레오타이드, 1 내지 550개의 뉴클레오타이드, 1 내지 500개의 뉴클레오타이드, 1 내지 450개의 뉴클레오타이드, 1 내지 400개의 뉴클레오타이드, 1 내지 350개의 뉴클레오타이드, 1 내지 300개의 뉴클레오타이드, 1 내지 250개의 뉴클레오타이드, 1 내지 200개의 뉴클레오타이드, 1 내지 150개의 뉴클레오타이드, 1 내지 100개의 뉴클레오타이드, 1 내지 50개의 뉴클레오타이드, 1 내지 25개의 뉴클레오타이드 또는 1 내지 10개의 뉴클레오타이드이고, 각 범위는 종점을 포함한다.

[0029] 일부 양태에서, 인터-ITR 서열의 길이는 1 내지 25개의 뉴클레오타이드, 1 내지 50개의 뉴클레오타이드, 25 내지 50개의 뉴클레오타이드, 50 내지 100개의 뉴클레오타이드, 100 내지 150개의 뉴클레오타이드, 150 내지 200개의 뉴클레오타이드, 200 내지 250개의 뉴클레오타이드, 250 내지 300개의 뉴클레오타이드, 300 내지 350개의 뉴클레오타이드, 350 내지 400개의 뉴클레오타이드, 400 내지 450개의 뉴클레오타이드, 450 내지 500개의 뉴클레오타이드, 500 내지 550개의 뉴클레오타이드, 550 내지 600개의 뉴클레오타이드, 600 내지 650개의 뉴클레오타이드, 650 내지 700개의 뉴클레오타이드이고, 각 범위는 종점을 포함한다. 일부 양태에서, 제2 핵산 서열의 길이는 1 내지 25개의 뉴클레오타이드, 1 내지 50개의 뉴클레오타이드, 25 내지 50개의 뉴클레오타이드, 50 내지 100개의 뉴클레오타이드, 100 내지 150개의 뉴클레오타이드, 150 내지 200개의 뉴클레오타이드, 200 내지 250개의 뉴클레오타이드, 250 내지 300개의 뉴클레오타이드, 300 내지 350개의 뉴클레오타이드, 350 내지 400개의 뉴클레오타이드, 400 내지 450개의 뉴클레오타이드, 450 내지 500개의 뉴클레오타이드, 500 내지 550개의 뉴클레오타이드, 550 내지 600개의 뉴클레오타이드, 600 내지 650개의 뉴클레오타이드, 650 내지 700개의 뉴클레

오타이드이고, 각 범위는 종점을 포함한다.

[0030] 본 개시내용의 짧은 나노트랜스포존(NTS)을 포함하는 일부 양태에서, 인터-ITR 서열의 길이는 1 내지 10개의 뉴클레오타이드, 10 내지 20개의 뉴클레오타이드, 20 내지 30개의 뉴클레오타이드, 30 내지 40개의 뉴클레오타이드, 40 내지 50개의 뉴클레오타이드, 50 내지 60개의 뉴클레오타이드, 60 내지 70개의 뉴클레오타이드, 70 내지 80개의 뉴클레오타이드, 80 내지 90개의 뉴클레오타이드, 또는 90 내지 100개의 뉴클레오타이드이고, 각 범위는 종점을 포함한다. 본 개시내용의 짧은 나노트랜스포존(NTS)을 포함하는 일부 양태에서, 핵산 서열의 길이는 1 내지 10개의 뉴클레오타이드, 10 내지 20개의 뉴클레오타이드, 20 내지 30개의 뉴클레오타이드, 30 내지 40개의 뉴클레오타이드, 40 내지 50개의 뉴클레오타이드, 50 내지 60개의 뉴클레오타이드, 60 내지 70개의 뉴클레오타이드, 70 내지 80개의 뉴클레오타이드, 80 내지 90개의 뉴클레오타이드, 또는 90 내지 100개의 뉴클레오타이드이고, 각 범위는 종점을 포함한다.

[0031] 일부 양태에서, 인트라-ITR 서열의 길이는 100개 이상의 뉴클레오타이드, 500개 이상의 뉴클레오타이드, 1000개 이상의 뉴클레오타이드, 1500개 이상의 뉴클레오타이드, 2000개 이상의 뉴클레오타이드, 2500개 이상의 뉴클레오타이드, 3000개 이상의 뉴클레오타이드, 3500개 이상의 뉴클레오타이드, 4000개 이상의 뉴클레오타이드, 4500개 이상의 뉴클레오타이드, 5000개 이상의 뉴클레오타이드, 5500개 이상의 뉴클레오타이드, 6000개 이상의 뉴클레오타이드, 6500개 이상의 뉴클레오타이드, 7000개 이상의 뉴클레오타이드, 7500개 이상의 뉴클레오타이드, 8000개 이상의 뉴클레오타이드, 8500개 이상의 뉴클레오타이드, 9000개 이상의 뉴클레오타이드, 9500개 이상의 뉴클레오타이드, 10000개 이상의 뉴클레오타이드(10 킬로베이스(kb)), 50000개 이상의 뉴클레오타이드(50 kb), 100000개 이상의 뉴클레오타이드(100 kb), 150000개 이상의 뉴클레오타이드(150 kb), 200000개 이상의 뉴클레오타이드(200 kb), 250000개 이상의 뉴클레오타이드(250 kb), 300000개 이상의 뉴클레오타이드(300 kb), 350000개 이상의 뉴클레오타이드(350 kb), 400000개 이상의 뉴클레오타이드(400 kb), 450000개 이상의 뉴클레오타이드(450 kb), 500000개 이상의 뉴클레오타이드(50 kb), 또는 이들 사이의 임의의 수의 뉴클레오타이드이다. 일부 양태에서, 제2 핵산 서열의 길이는 100개 이상의 뉴클레오타이드, 500개 이상의 뉴클레오타이드, 1000개 이상의 뉴클레오타이드, 1500개 이상의 뉴클레오타이드, 2000개 이상의 뉴클레오타이드, 2500개 이상의 뉴클레오타이드, 3000개 이상의 뉴클레오타이드, 3500개 이상의 뉴클레오타이드, 4000개 이상의 뉴클레오타이드, 4500개 이상의 뉴클레오타이드, 5000개 이상의 뉴클레오타이드, 5500개 이상의 뉴클레오타이드, 6000개 이상의 뉴클레오타이드, 6500개 이상의 뉴클레오타이드, 7000개 이상의 뉴클레오타이드, 7500개 이상의 뉴클레오타이드, 8000개 이상의 뉴클레오타이드, 8500개 이상의 뉴클레오타이드, 9000개 이상의 뉴클레오타이드, 9500개 이상의 뉴클레오타이드, 10000개 이상의 뉴클레오타이드(10 킬로베이스(kb)), 50000개 이상의 뉴클레오타이드(50 kb), 100000개 이상의 뉴클레오타이드(100 kb), 150000개 이상의 뉴클레오타이드(150 kb), 200000개 이상의 뉴클레오타이드(200 kb), 250000개 이상의 뉴클레오타이드(250 kb), 300000개 이상의 뉴클레오타이드(300 kb), 350000개 이상의 뉴클레오타이드(350 kb), 400000개 이상의 뉴클레오타이드(400 kb), 450000개 이상의 뉴클레오타이드(450 kb), 500000개 이상의 뉴클레오타이드(50 kb), 또는 이들 사이의 임의의 수의 뉴클레오타이드이다.

[0032] 조성물은 복제 원점 서열 또는 복제 서열을 코딩하는 서열을 추가로 포함할 수 있다. 제1 핵산 서열 또는 제2 핵산 서열은 복제 원점 서열 또는 복제 서열을 코딩하는 서열을 추가로 포함할 수 있다. 바람직하게는, 제1 핵산 서열은 복제 원점 서열 또는 복제 서열을 코딩하는 서열을 포함한다.

[0033] 일부 양태에서, 복제 원점 서열의 길이는 450개 이하의 뉴클레오타이드, 400개 이하의 뉴클레오타이드, 350개 이하의 뉴클레오타이드, 300개 이하의 뉴클레오타이드, 250개 이하의 뉴클레오타이드, 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 100개 이하의 뉴클레오타이드, 50개 이하의 뉴클레오타이드, 25개 이하의 뉴클레오타이드, 또는 10개 이하의 뉴클레오타이드이다. 일부 양태에서, 복제 원점 서열의 길이는 1 내지 450개의 뉴클레오타이드, 1 내지 400개의 뉴클레오타이드, 1 내지 350개의 뉴클레오타이드, 1 내지 300개의 뉴클레오타이드, 1 내지 250개의 뉴클레오타이드, 1 내지 200개의 뉴클레오타이드, 1 내지 150개의 뉴클레오타이드, 1 내지 100개의 뉴클레오타이드, 1 내지 50개의 뉴클레오타이드, 1 내지 25개의 뉴클레오타이드, 또는 1 내지 10개의 뉴클레오타이드이고, 각 범위는 종점을 포함한다.

[0034] 복제 원점 서열은 R6K 복제 원점을 포함할 수 있다. R6K 복제 원점은 R6K 감마 복제 원점을 포함할 수 있다. 복제 원점 서열은 미니 복제 원점을 포함할 수 있다. 미니 복제 원점은 R6K 미니 복제 원점을 포함할 수 있다. R6K 미니 복제 원점은 R6K 감마 미니 복제 원점을 포함할 수 있다. R6K 감마 미니 복제 원점의 길이는 281개의 뉴클레오타이드(281개의 염기쌍)이고, 서열번호: 15의 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는

구성된다.

[0035] 조성물은 제1 선택 마커 또는 제1 선택 마커를 코딩하는 서열을 추가로 포함할 수 있다. 제1 핵산 서열 또는 제2 핵산 서열은 제1 선택 마커 또는 제1 선택 마커를 코딩하는 서열을 추가로 포함할 수 있다. 바람직하게는, 제1 핵산 서열은 제1 선택 마커 또는 제1 선택 마커를 코딩하는 서열을 포함할 수 있다.

[0036] 일부 양태에서, 제1 선택 마커의 길이는 450개 이하의 뉴클레오타이드, 200개 이하의 뉴클레오타이드, 150개 이하의 뉴클레오타이드, 100개 이하의 뉴클레오타이드, 50개 이하의 뉴클레오타이드, 25개 이하의 뉴클레오타이드, 또는 10개 이하의 뉴클레오타이드이다. 일부 양태에서, 제1 선택 마커의 길이는 1 내지 200개의 뉴클레오타이드, 1 내지 150개의 뉴클레오타이드, 1 내지 100개의 뉴클레오타이드, 1 내지 50개의 뉴클레오타이드, 1 내지 25개의 뉴클레오타이드, 또는 1 내지 10개의 뉴클레오타이드이고, 각 범위는 종점을 포함한다.

[0037] 제1 선택 마커는 수크로스 선택 마커, 형광 마커, 세포 표면 마커, 또는 이의 조합을 포함할 수 있다. 바람직한 양태에서, 제1 선택 마커는 수크로스 선택 마커를 포함하거나, 본질적으로 구성되거나, 또는 구성된다. 바람직한 양태에서, 수크로스 선택 마커는 RNA-OUT 선택 마커를 포함한다. RNA-OUT 선택 마커의 길이는 139개의 뉴클레오타이드(139개의 염기쌍)이고 서열번호: 16의 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다.

[0038] 제1 ITR을 코딩하는 서열 또는 제2 ITR을 코딩하는 서열은 TTAA, TTAT, 또는 TTAX 인식 서열을 포함할 수 있다. 제1 ITR을 코딩하는 서열 또는 제2 ITR을 코딩하는 서열은 적어도 2개, 적어도 3개, 적어도 4개, 적어도 5개, 적어도 6개, 적어도 7개, 적어도 8개, 적어도 9개, 적어도 10개, 적어도 11개, 적어도 12개, 적어도 13개, 적어도 14개, 적어도 15개, 적어도 16개, 적어도 17개, 적어도 18개, 적어도 19개, 또는 적어도 20개의 뉴클레오타이드를 포함할 수 있다.

[0039] 제1 ITR을 코딩하는 서열 또는 제2 ITR을 코딩하는 서열은 서열번호: 24와 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함할 수 있거나, 본질적으로 구성될 수 있거나, 또는 구성될 수 있다. 제1 ITR을 코딩하는 서열 또는 제2 ITR을 코딩하는 서열은 서열번호: 25와 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함할 수 있거나, 본질적으로 구성될 수 있거나, 또는 구성될 수 있다. 제1 ITR을 코딩하는 서열 또는 제2 ITR을 코딩하는 서열은 서열번호: 26과 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함할 수 있거나, 본질적으로 구성될 수 있거나, 또는 구성될 수 있다. 제1 ITR을 코딩하는 서열 또는 제2 ITR을 코딩하는 서열은 서열번호: 27과 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함할 수 있거나, 본질적으로 구성될 수 있거나, 또는 구성될 수 있다. 제1 ITR을 코딩하는 서열 또는 제2 ITR을 코딩하는 서열은 서열번호: 2와 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함할 수 있거나, 본질적으로 구성될 수 있거나, 또는 구성될 수 있다. 제1 ITR을 코딩하는 서열 또는 제2 ITR을 코딩하는 서열은 서열번호: 14와 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함할 수 있거나, 본질적으로 구성될 수 있거나, 또는 구성될 수 있다.

[0040] 일 양태에서, 제1 ITR을 코딩하는 서열은 서열번호: 24의 핵산 서열을 포함하고, 제2 ITR을 코딩하는 제2 서열은 서열번호: 25의 핵산 서열을 포함한다. 일 양태에서, 제1 ITR을 코딩하는 서열은 서열번호: 24의 핵산 서열을 포함하고, 제2 ITR을 코딩하는 제2 서열은 서열번호: 26의 핵산 서열을 포함한다. 일 양태에서, 제1 ITR을 코딩하는 서열은 서열번호: 24의 핵산 서열을 포함하고, 제2 ITR을 코딩하는 제2 서열은 서열번호: 27의 핵산 서열의 핵산 서열을 포함한다.

[0041] 제1 핵산 서열은 적어도 하나의 외인성 서열 및 포유동물 세포에서 외인성 서열을 발현할 수 있는 적어도 하나의 프로모터를 추가로 포함할 수 있다. 바람직한 양태에서, 프로모터는 인간 세포에서 외인성 서열을 발현할 수 있다. 바람직한 양태에서, 조성물의 트랜스포존 서열은 적어도 하나의 외인성 서열 및 포유동물 세포에서 외인성 서열을 발현할 수 있는 적어도 하나의 프로모터를 포함한다.

[0042] 프로모터는 항시성 프로모터일 수 있다. 프로모터는 유도성 프로모터일 수 있다. 프로모터는 세포-유형 또는 조직-유형 특이적 프로모터일 수 있다. 프로모터는 EF1a 프로모터(서열번호: 4), CMV 프로모터, MND 프로모터, SV40 프로모터, PGK1 프로모터, Ubc 프로모터, CAG 프로모터, H1 프로모터, 또는 U6 프로모터일 수 있다. 바람직한 양태에서, 프로모터는 EF1a 프로모터이다. 일 양태에서, 제1 핵산 서열은 포유동물 세포에서 제1 외인성

서열을 발현할 수 있는 제1 프로모터를 코딩하는 제1 서열, 및 포유동물 세포에서 제2 외인성 서열을 발현할 수 있는 제2 프로모터를 코딩하는 제2 서열을 포함하고, 제1 프로모터는 항시성 프로모터이고, 제2 프로모터는 유도성 프로모터이다. 일 양태에서, 제1 프로모터를 코딩하는 제1 서열 및 제2 프로모터를 코딩하는 제2 서열은 반대 방향으로 배향된다.

[0043] 적어도 하나의 외인성 서열은 비자연발생 항원 수용체를 코딩하는 서열, 치료적 폴리펩타이드를 코딩하는 서열, 또는 이의 조합을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. 비자연발생 항원 수용체는 키메라 항원 수용체(CAR), T 세포 수용체(TCR), 키메라 자극 수용체(CSR), HLA 부류 I 조직적합성 항원, 알파 사슬 E 재조합 폴리펩타이드(HLA-E), 베타-2-마이크로글로불린(B2M) 재조합 폴리펩타이드, 또는 이의 조합을 포함할 수 있다. TCR, CSR, HLA-E 및 B2M은 본원에 상세히 기재되어 있다. 바람직한 양태에서, 비자연발생 항원 수용체는 CAR을 포함한다.

[0044] 적어도 하나의 외인성 서열은 유도성 아폽토시스 촉진(proapoptotic) 폴리펩타이드를 코딩하는 서열을 추가로 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있다. 유도성 아폽토시스 촉진 폴리펩타이드는 본원에 상세히 기재되어 있다.

[0045] 적어도 하나의 외인성 서열은 제2 선택 마커를 코딩하는 서열을 추가로 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있다. 제2 선택 마커는 세포 생존력 및 생존에 필수적인 유전자 생성물을 코딩할 수 있다. 제2 선택 마커는 선택적인 세포 배양 조건에 의해 도전될 때 세포 생존력 및 생존에 필수적인 유전자 생성물을 코딩할 수 있다. 선택적인 세포 배양 조건은 세포 생존력 또는 생존에 유해한 화합물을 포함할 수 있고, 유전자 생성물은 화합물에 대한 내성을 부여한다. 선택 유전자의 비제한적인 예는 *neo*(네오마이신에 대한 내성을 부여함), DHFR(디하이드로폴레이트 환원효소를 코딩하고 메토트렉세이트에 대한 내성을 부여함), TYMS(티미딜레이트 합성효소를 코딩함), MGMT(O(6)-메틸구아닌-DNA 메틸트랜스퍼라아제를 코딩함), 다중약물 내성 유전자(MDR1), ALDH1(알데하이드 탈수소효소 1 패밀리, 구성원 A1을 코딩함), FRANCF, RAD51C(RAD51 Paralog C를 코딩함), GCS(글루코실세라미드 신타아제를 코딩함), NKX2.2(NK2 Homeobox 2를 코딩함), 또는 이의 임의의 조합을 포함한다.

[0046] 제2 선택 마커는 검출가능한 마커일 수 있다. 검출가능한 마커는 형광 마커, 세포 표면 마커 또는 대사 마커일 수 있다. 바람직한 양태에서, 제2 선택 마커는 디하이드로폴레이트 환원효소(DHFR) 뮤테인 효소를 코딩하는 서열을 포함한다. DHFR 뮤테인 효소는 서열번호: 52의 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. DHFR 뮤테인 효소는 서열번호: 53 또는 서열번호: 11의 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다. DHFR 뮤테인 효소의 아미노산 서열은 위치 80, 113, 또는 153 중 하나 이상에서 돌연변이를 추가로 포함할 수 있다. DHFR 뮤테인 효소의 아미노산 서열은 위치 80에서 페닐알라닌(F) 또는 류신(L)의 치환, 위치 113에서 류신(L) 또는 발린(V)의 치환, 및 위치 153에서 발린(V) 또는 아스파르트산(D)의 치환 중 하나 이상을 포함할 수 있다.

[0047] 적어도 하나의 외인성 서열은 적어도 하나의 자가 절단 펩타이드를 코딩하는 서열을 추가로 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있다. 예를 들어, 자가 절단 펩타이드는 CAR 및 유도성 아폽토시스 촉진 폴리펩타이드 사이에 위치할 수 있거나, 자가 절단 펩타이드는 CAR 및 제2 선택 마커 사이에 위치할 수 있다.

[0048] 적어도 하나의 외인성 서열은 적어도 2개의 자가 절단 펩타이드를 코딩하는 서열을 추가로 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있다. 예를 들어, 제1 자가 절단 펩타이드는 CAR의 상류 또는 바로 상류에 위치하고, 제2 자가 절단 펩타이드는 CAR의 하류 또는 바로 하류에 위치에 위치하거나; 또는 제1 자가 절단 펩타이드 및 제2 자가 절단 펩타이드는 CAR 측면에 위치한다. 예를 들어, 제1 자가 절단 펩타이드는 유도성 아폽토시스 촉진 폴리펩타이드의 상류 또는 바로 상류에 위치하고, 제2 자가 절단 펩타이드는 유도성 아폽토시스 촉진 폴리펩타이드의 하류 또는 바로 하류에 위치하거나; 또는 제1 자가 절단 펩타이드 및 제2 자가 절단 펩타이드는 유도성 아폽토시스 촉진 폴리펩타이드 측면에 위치한다. 예를 들어, 제1 자가 절단 펩타이드는 제2 선택 마커의 상류 또는 바로 상류에 위치하고, 제2 자가 절단 펩타이드는 제2 선택 마커의 하류 또는 바로 하류에 위치하거나; 또는 제1 자가 절단 펩타이드 및 제2 자가 절단 펩타이드는 제2 선택 마커 측면에 위치한다.

[0049] 자가 절단 펩타이드의 비제한적인 예는 T2A 펩타이드, GSG-T2A 펩타이드, E2A 펩타이드, GSG-E2A 펩타이드, F2A 펩타이드, GSG-F2A 펩타이드, P2A 펩타이드, 또는 GSG-P2A 펩타이드를 포함한다. T2A 펩타이드는 서열번호: 54와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. GSG-T2A 펩타이드는 서열번호: 55와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노

산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. GSG-T2A 폴리펩타이드는 서열번호: 7, 서열번호: 8, 서열번호: 10, 서열번호: 56과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나 구성되는 폴리뉴클레오타이드에 의해 코딩된다. E2A 펩타이드는 서열번호: 57과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. GSG-E2A 펩타이드는 서열번호: 58과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. F2A 펩타이드는 서열번호: 59와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. GSG-F2A 펩타이드는 서열번호: 60과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. P2A 펩타이드는 서열번호: 61과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. GSG-P2A 펩타이드는 서열번호: 62와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다.

[0050] 적어도 하나의 외인성 서열 및 포유동물 세포에서 외인성 서열을 발현할 수 있는 적어도 하나의 프로모터를 포함하는 제1 핵산 서열은 인슐레이터(insulator)를 코딩하는 적어도 하나의 서열을 추가로 포함할 수 있다. 일 양태에서, 제1 핵산 서열은 제1 인슐레이터를 코딩하는 제1 서열 및 제2 인슐레이터를 코딩하는 제2 서열을 포함할 수 있다. 일부 구현예에서, 제1 또는 제2 인슐레이터를 코딩하는 서열은 서열번호: 3 또는 서열번호: 13과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다.

[0051] 적어도 하나의 외인성 서열 및 포유동물 세포에서 외인성 서열을 발현할 수 있는 적어도 하나의 프로모터를 포함하는 제1 핵산 서열은 폴리아데노신(polyA) 서열을 추가로 포함할 수 있다. 적어도 하나의 외인성 서열, 포유동물 세포에서 외인성 서열을 발현할 수 있는 적어도 하나의 프로모터 및 인슐레이터를 코딩하는 적어도 하나의 서열을 포함하는 제1 핵산 서열은 폴리아데노신(polyA) 서열을 추가로 포함할 수 있다. polyA 서열은 바이러스 polyA 서열로부터 단리되거나 유래될 수 있다. polyA 서열은 (SV40) polyA 서열로부터 단리되거나 유래될 수 있다. 일부 구현예에서, 제1 또는 제2 인슐레이터를 코딩하는 서열은 서열번호: 12와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다.

[0052] 일 양태에서, 조성물은 외래 DNA를 코딩하는 서열을 포함하지 않는다. 일 양태에서, 제1 핵산 서열은 외래 DNA를 코딩하는 서열을 포함하지 않는다. 일 양태에서, 제2 핵산 서열은 외래 DNA를 코딩하는 서열을 포함하지 않는다. 일 양태에서, 조성물은 외래 DNA를 코딩하는 서열을 포함한다. 일 양태에서, 제1 핵산 서열은 외래 DNA를 코딩하는 서열을 포함한다. 일 양태에서, 제2 핵산 서열은 외래 DNA를 코딩하는 서열을 포함한다. 외래 DNA는 외인성 서열이 발현될 포유동물 세포와 동일한 유기체로부터 유래되거나 수득되지 않은 DNA 서열이다. 예를 들어, 외래 DNA는 포유동물이 아닌 바이러스로부터의 DNA일 수 있거나; 또는 외래 DNA는 포유동물이 아닌 파충류로부터의 DNA일 수 있다. 또 다른 양태에서, 외래 DNA는 하나의 포유동물로부터 유래될 수 있지만, 상기 포유동물은 외인성 서열이 발현될 포유동물과 상이하다. 예를 들어, 외래 DNA는 인간이 아닌 랫트로부터 유래된다.

[0053] 일 양태에서, 조성물은 재조합 부위, 절제 부위, 결찰 부위, 또는 이의 조합을 포함하지 않는다. 일 양태에서, 조성물은 재조합 사건, 절제 사건, 결찰 사건, 또는 이의 조합의 생성물을 포함하지 않는다. 일 양태에서, 조성물은 재조합 사건, 절제 사건, 결찰 사건, 또는 이의 조합으로부터 유래되지 않는다.

[0054] 일 양태에서, 제1 핵산 서열은 재조합 부위, 절제 부위, 결찰 부위, 또는 이의 조합을 포함하지 않는다. 일 양태에서, 제1 핵산 서열은 재조합 사건, 절제 사건, 결찰 사건, 또는 이의 조합의 생성물을 포함하지 않는다. 일 양태에서, 제1 핵산 서열은 재조합 사건, 절제 사건, 결찰 사건, 또는 이의 조합으로부터 유래되지 않는다.

[0055] 일 양태에서, 제2 핵산 서열은 재조합 부위, 절제 부위, 결찰 부위, 또는 이의 조합을 포함하지 않는다. 일 양태에서, 제2 핵산 서열은 재조합 사건, 절제 사건, 결찰 사건, 또는 이의 조합의 생성물을 포함하지 않는다. 일 양태에서, 제2 핵산 서열은 재조합 사건, 절제 사건, 결찰 사건, 또는 이의 조합으로부터 유래되지 않는다.

[0056] 재조합 부위는 재조합 사건으로부터 비롯되는 서열을 포함할 수 있거나, 재조합 사건의 생성물인 서열을 포함할 수 있거나, 또는 재조합효소의 활성(예컨대, 재조합효소 부위)을 포함할 수 있다.

[0057] **키메라 항원 수용체(CAR)**

[0058] 본 개시내용은 또한 CAR을 포함하는 조성물(예컨대, 나노트랜스포존)을 제공하며, CAR은 항원 인식 영역을 포함하는 엑토도메인; 막관통 도메인, 및 적어도 하나의 공자극 도메인을 포함하는 엔도도메인을 포함한다. CAR은 항원 인식 도메인 및 막관통 도메인 사이에 힌지 영역을 추가로 포함할 수 있다.

[0059] 항원 인식 영역은 적어도 하나의 단일 사슬 가변 단편(scFv), 센티린, 단일 도메인 항체, 또는 이의 조합을 포함할 수 있다. 일 양태에서, 적어도 하나의 단일 도메인 항체는 VHH이다. 일 양태에서, 적어도 하나의 단일 도메인 항체는 VH이다.

[0060] scFv

[0061] 본 개시내용의 조성물(예컨대, 트랜스포존 또는 나노트랜스포존)은 CAR을 포함할 수 있고; 일부 양태에서, CAR의 항원 인식 영역은 특정 표적 단백질/항원을 인식하고 이에 결합하기 위한 하나 이상의 scFv 조성물을 포함할 수 있다. 항원 인식 영역은 적어도 2개의 scFv를 포함할 수 있다. 항원 인식 영역은 적어도 3개의 scFv를 포함할 수 있다. 일 양태에서, 본 개시내용의 CAR은 2개의 구별되는 항원에 특이적으로 결합하는 적어도 2개의 scFv를 포함하는 이중특이적 CAR이다.

[0062] scFv 조성물은 항체의 중쇄 가변 영역 및 경쇄 가변 영역을 포함한다. scFv는 면역글로불린의 중쇄(VH) 및 경쇄(VL)의 가변 영역의 융합 단백질이며, VH 및 VL 도메인은 짧은 펩타이드 링커를 이용하여 연결된다. scFv는 불변 영역의 제거 및 링커의 도입에도 불구하고 원래의 면역글로불린의 특이성을 유지할 수 있다. 일부 양태에서, 링커 폴리펩타이드는 서열번호: 33의 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. 링커 폴리펩타이드는 서열번호: 34의 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩될 수 있다.

[0063] 센티린

[0064] 본 개시내용의 조성물(예컨대, 트랜스포존 또는 나노트랜스포존)은 CAR을 포함할 수 있고; 일부 양태에서, CAR의 항원 인식 영역은 특정 표적 단백질/항원을 인식하고 이에 결합하기 위한 하나 이상의 센티린 조성물을 포함할 수 있다. 항원에 특이적으로 결합하는 센티린은 세포(예컨대, 세포독성 면역 세포)의 특이성을 특정 항원으로 향하게 하는 데 사용될 수 있다. 센티린을 포함하는 CAR은 본원에서 CARTyrin으로 지칭된다.

[0065] 본 개시내용의 센티린은 단백질 스캐폴드를 포함할 수 있고, 스캐폴드는 항원에 특이적으로 결합할 수 있다. 본 개시내용의 센티린은 적어도 하나의 피브로넥틴 유형 III(FN3) 도메인의 공통 서열을 포함하는 단백질 스캐폴드를 포함할 수 있고, 스캐폴드는 항원에 특이적으로 결합할 수 있다. 적어도 하나의 피브로넥틴 유형 III(FN3) 도메인은 인간 단백질로부터 유래될 수 있다. 인간 단백질은 테나신-C(Tenascin-C)일 수 있다. 공통 서열은 서열번호: 84와 적어도 74%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되거나, 또는 공통 서열은 서열번호: 85와 적어도 74%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. 공통 서열은 서열번호: 86과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0066] 공통 서열은 (a) 공통 서열의 위치 13-16에 아미노산 잔기 TEDS(서열번호: 87)를 포함하거나 구성되는 A-B 루프; (b) 공통 서열의 위치 22-28에 아미노산 잔기 TAPDAAF(서열번호: 88)를 포함하거나 구성되는 B-C 루프; (c) 공통 서열의 위치 38-43에 아미노산 잔기 SEKVGE(서열번호: 89)를 포함하거나 구성되는 C-D 루프; (d) 공통 서열의 위치 51-54에 아미노산 잔기 GSER(서열번호: 90)를 포함하거나 구성되는 D-E 루프; (e) 공통 서열의 위치 60-64에 아미노산 잔기 GLKPG(서열번호: 91)를 포함하거나 구성되는 E-F 루프; (f) 공통 서열의 위치 75-81에 아미노산 잔기 KGGHRSN(서열번호: 92)을 포함하거나 구성되는 F-G 루프; 또는 (g) (a)-(f)의 임의의 조합 내의 하나 이상의 위치에서 변형될 수 있다. 본 개시내용의 센티린은 적어도 5개의 피브로넥틴 유형 III(FN3) 도메인, 적어도 10개의 피브로넥틴 유형 III(FN3) 도메인 또는 적어도 15개의 피브로넥틴 유형 III(FN3) 도메인의 공통 서열을 포함할 수 있다.

[0067] 용어 "항체 모방체"는 표적 서열에 특이적으로 결합하고 자연발생 항체와 구별되는 구조를 갖는 유기 화합물을 기술하는 것으로 의도된다. 항체 모방체는 단백질, 핵산, 또는 소분자를 포함할 수 있다. 본 개시내용의 항체 모방체가 특이적으로 결합하는 표적 서열은 항원일 수 있다. 항체 모방체는, 비제한적으로, 우수한 용해도, 조

직 침투, 열 및 효소에 대한 안정성(예컨대, 효소 분해에 대한 내성), 및 더 낮은 생산 비용을 포함하여 항체보다 우수한 특성을 제공할 수 있다. 예시적인 항체 모방체는, 비제한적으로, 아피바디(affibody), 아필린(afflilin), 아피머(affimer), 아피틴(affitin), 알파바디(alphabody), 안티칼린(anticalin), 및 아비머(avimer)(결합력 다량체(avidity multimer)로도 알려짐), DARPin(Designed Ankyrin Repeat Protein), 피노머(Fynomer), 쿠니츠(Kunitz) 도메인 펩타이드, 및 모노바디를 포함한다.

[0068] 본 개시내용의 아피바디 분자는 임의의 디설파이드 다리가 없는 하나 이상의 알파 나선을 포함하거나 구성되는 단백질 스캐폴드를 포함한다. 바람직하게는, 본 개시내용의 아피바디 분자는 3개의 알파 나선을 포함하거나 구성된다. 예를 들어, 본 개시내용의 아피바디 분자는 면역글로불린 결합 도메인을 포함할 수 있다. 본 개시내용의 아피바디 분자는 단백질 A의 Z 도메인을 포함할 수 있다.

[0069] 본 개시내용의 아필린 분자는, 예를 들어, 감마-B 크리스탈린 또는 유비퀴틴의 노출된 아미노산의 변형에 의해 생산된 단백질 스캐폴드를 포함한다. 아필린 분자는 항원에 대한 항체의 친화성을 기능적으로 모방하지만, 항체를 구조적으로 모방하지 않는다. 아필린을 제조하는 데 사용된 임의의 단백질 스캐폴드에서, 용매 또는 적절하게 접힌 단백질 분자 내의 가능한 결합 파트너에 접근가능한 아미노산은 노출된 아미노산으로 간주된다. 이들 노출된 아미노산 중 어느 하나 이상은 표적 서열 또는 항원에 특이적으로 결합하도록 변형될 수 있다.

[0070] 본 개시내용의 아피머 분자는 특이적 표적 서열에 대해 높은 친화성 결합 부위를 제공하는 펩타이드 루프를 나타내도록 조작된 매우 안정한 단백질을 포함하는 단백질 스캐폴드를 포함한다. 본 개시내용의 예시적인 아피머 분자는 시스타틴 단백질 또는 이의 3차 구조에 기초한 단백질 스캐폴드를 포함한다. 본 개시내용의 예시적인 아피머 분자는 역평행 베타-시트 위에 놓인 알파-나선을 포함하는 공통 3차 구조를 공유할 수 있다.

[0071] 본 개시내용의 아피틴 분자는 인공 단백질 스캐폴드를 포함하고, 이의 구조는, 예를 들어, DNA 결합 단백질(예컨대, DNA 결합 단백질 Sac7d)로부터 유래될 수 있다. 본 개시내용의 아피틴은 항원의 전체 또는 일부일 수 있는 표적 서열에 선택적으로 결합한다. 본 개시내용의 예시적인 아피틴은 DNA 결합 단백질의 결합 표면 상에 하나 이상의 아미노산 서열을 무작위화하고 생성된 단백질을 리보솜 디스플레이 및 선택에 적용함으로써 제조된다. 본 개시내용의 아피틴의 표적 서열은, 예를 들어, 게놈에서 또는 펩타이드, 단백질, 바이러스, 또는 박테리아의 표면 상에서 발견될 수 있다. 일부 양태에서, 아피틴 분자는 효소의 특정 억제제로서 사용될 수 있다. 본 개시내용의 아피틴 분자는 내열성 단백질 또는 이의 유도체를 포함할 수 있다.

[0072] 본 개시내용의 알파바디 분자는 또한 세포-침투 알파바디(Cell-Penetrating Alphabody, CPAB)로 지칭될 수 있다. 본 개시내용의 알파바디 분자는 다양한 표적 서열(항원 포함)에 결합하는 작은 단백질(전형적으로 10 kDa 미만)을 포함한다. 알파바디 분자는 세포내 표적 서열에 도달하고 결합할 수 있다. 구조적으로, 본 개시내용의 알파바디 분자는 단일 사슬 알파 나선(자연발생 코일형 코일 구조와 유사함)을 형성하는 인공 서열을 포함한다. 본 개시내용의 알파바디 분자는 표적 단백질에 특이적으로 결합하도록 변형된 하나 이상의 아미노산을 포함하는 단백질 스캐폴드를 포함할 수 있다. 분자의 결합 특이성에 관계없이, 본 개시내용의 알파바디 분자는 정확한 접힘 및 열안정성을 유지한다.

[0073] 본 개시내용의 안티칼린 분자는 단백질 또는 소분자 내의 표적 서열 또는 부위에 결합하는 인공 단백질을 포함한다. 본 개시내용의 안티칼린 분자는 인간 리포칼린으로부터 유래된 인공 단백질을 포함할 수 있다. 본 개시내용의 안티칼린 분자는, 예를 들어, 단클론 항체 또는 이의 단편 대신에 사용될 수 있다. 안티칼린 분자는 단클론 항체 또는 이의 단편보다 우수한 조직 침투 및 열안정성을 입증할 수 있다. 본 개시내용의 예시적인 안티칼린 분자는 대략 20 kDa의 질량을 갖는 약 180개의 아미노산을 포함할 수 있다. 구조적으로, 본 개시내용의 안티칼린 분자는 루프에 의해 쌍으로 연결된 역평행 베타-가닥 및 부착된 알파 나선을 포함하는 배럴(barrel) 구조를 포함한다. 일부 양태에서, 본 개시내용의 안티칼린 분자는 루프에 의해 쌍으로 연결된 8개의 역평행 베타-가닥 및 부착된 알파 나선을 포함하는 배럴 구조를 포함한다.

[0074] 본 개시내용의 아비머 분자는 표적 서열(항원일 수도 있음)에 특이적으로 결합하는 인공 단백질을 포함한다. 본 개시내용의 아비머는 동일한 표적 내 또는 구별되는 표적 내의 다수의 결합 부위를 인식할 수 있다. 본 개시내용의 아비머가 하나 초과의 표적을 인식하는 경우, 아비머는 이중특이적 항체의 기능을 모방한다. 인공 단백질 아비머는 각각 대략 30-35개의 아미노산의 2개 이상의 펩타이드 서열을 포함할 수 있다. 이들 펩타이드는 하나 이상의 링커 펩타이드를 통해 연결될 수 있다. 아비머의 하나 이상의 펩타이드의 아미노산 서열은 막 수용체의 A 도메인으로부터 유래될 수 있다. 아비머는 선택적으로 디설파이드 결합 및/또는 칼슘을 포함할 수 있는 견고한 구조를 갖는다. 본 개시내용의 아비머는 항체에 비해 더 큰 열 안정성을 입증할 수 있다.

[0075] 본 개시내용의 DARPin(Designed Ankyrin Repeat Protein)은 표적 서열에 대해 높은 특이성 및 높은 친화성을 갖는 유전적으로 조작된, 재조합, 또는 키메라 단백질을 포함한다. 일부 양태에서, 본 개시내용의 DARPin은 안키린 단백질로부터 유래되고, 선택적으로, 안키린 단백질의 적어도 3개의 반복 모티프(반복적인 구조 단위로도 지칭됨)를 포함한다. 안키린 단백질은 높은 친화성 단백질-단백질 상호작용을 매개한다. 본 개시내용의 DARPin은 큰 표적 상호작용 표면을 포함한다.

[0076] 본 개시내용의 피노머는 인간 Fyn SH3 도메인로부터 유래되고 항체와 동일한 친화성 및 동일한 특이성으로 표적 서열 및 분자에 결합하도록 조작된 작은 결합 단백질(약 7 kDa)을 포함한다.

[0077] 본 개시내용의 쿠니츠 도메인 펩타이드는 쿠니츠 도메인을 포함하는 단백질 스캐폴드를 포함한다. 쿠니츠 도메인은 프로테아제 활성을 억제하기 위한 활성 부위를 포함한다. 구조적으로, 본 개시내용의 쿠니츠 도메인은 디설파이드가 풍부한 알파+베타 폴드를 포함한다. 이 구조는 소 췌장 트립신 억제제에 의해 예시된다. 쿠니츠 도메인 펩타이드는 특이적 단백질 구조를 인식하고 경쟁적 프로테아제 억제제로서 작용한다. 본 개시내용의 쿠니츠 도메인은 에칼란티드(Ecallantide)(인간 지질단백질 연관 응고 억제제(LACI)로부터 유래됨)를 포함할 수 있다.

[0078] 본 개시내용의 모노바디는 단일 사슬 항체와 크기가 비슷한 작은 단백질(약 94개의 아미노산을 포함하고 약 10 kDa의 질량을 가짐)이다. 이러한 유전적으로 조작된 단백질은 항원을 포함하는 표적 서열에 특이적으로 결합한다. 본 개시내용의 모노바디는 하나 이상의 구별되는 단백질 또는 표적 서열을 특이적으로 표적화할 수 있다. 일부 양태에서, 본 개시내용의 모노바디는 인간 피브로넥틴의 구조를 모방하고, 보다 바람직하게는, 피브로넥틴의 10번째 세포외 유형 III 도메인의 구조를 모방하는 단백질 스캐폴드를 포함한다. 피브로넥틴의 10번째 세포외 유형 III 도메인뿐만 아니라 이의 모노바디 모방체는 배럴을 형성하는 7개의 베타 시트 및 항체의 3개의 상보성 결정 영역(CDR)에 상응하는 각 측면에 3개의 노출된 루프를 함유한다. 항체의 가변 도메인의 구조와 달리, 모노바디는 금속 이온에 대한 임의의 결합 부위뿐만 아니라 중심 디설파이드 결합이 없다. 다중특이적 모노바디는 루프 BC 및 FG를 변형시킴으로써 최적화될 수 있다. 본 개시내용의 모노바디는 아드넥틴(adnectin)을 포함할 수 있다.

[0079] VHH

[0080] 본 개시내용의 조성물(예컨대, 트랜스포존 또는 나노트랜스포존)은 CAR을 포함할 수 있고; 일부 양태에서, CAR의 항원 인식 영역은 특정 표적 단백질/항원을 인식하고 이에 결합하는 적어도 하나의 단일 도메인 항체(SdAb)를 포함할 수 있다. 일 양태에서, 단일 도메인 항체는 VHH이다. VHH는 낙타과에서 발견되는 중쇄 항체이다. 항원에 특이적으로 결합하는 VHH는 세포(예컨대, 세포독성 면역 세포)의 특이성을 특정 항원으로 향하게 하는 데 사용될 수 있다. 항원 인식 영역은 적어도 2개의 VHH를 포함할 수 있다. 항원 인식 영역은 적어도 3개의 VHH를 포함할 수 있다. 일 양태에서, 본 개시내용의 CAR은 2개의 구별되는 항원에 특이적으로 결합하는 적어도 2개의 VHH를 포함하는 이중특이적 CAR이다. VHH를 포함하는 CAR은 본원에서 VCAR로 지칭된다.

[0081] 본 개시내용의 적어도 하나의 VHH 단백질 또는 VCAR은 당업계에 잘 알려진 바와 같이 세포주, 혼합된 세포주, 불멸화된 세포 또는 불멸화된 세포의 클론 집단에 의해 선택적으로 생산될 수 있다. 예컨대, 문헌[Ausubel, et al., ed., Current Protocols in Molecular Biology, John Wiley & Sons, Inc., NY, N.Y.(1987-2001); Sambrook, et al., Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor, N.Y.(1989); Harlow and Lane, Antibodies, a Laboratory Manual, Cold Spring Harbor, N.Y.(1989); Colligan, et al., eds., Current Protocols in Immunology, John Wiley & Sons, Inc., NY(1994-2001); Colligan et al., Current Protocols in Protein Science, John Wiley & Sons, NY, N.Y.,(1997-2001)]을 참고한다.

[0082] VHH 단백질로부터의 아미노산은 면역원성을 감소시키거나 결합, 친화성, 온-속도, 오프-속도, 결합력, 특이성, 반감기, 안정성, 용해도 또는 당업계에 공지된 바와 같은 임의의 다른 적합한 특징을 감소, 향상 또는 변형시키기 위해 변경, 부가 및/또는 결실될 수 있다.

[0083] 선택적으로, VHH 단백질은 항원에 대한 높은 친화성 및 다른 유리한 생물학적 특성을 유지하면서 조작될 수 있다. 이 목표를 달성하기 위해, VHH 단백질은 부모 및 조작된 서열의 3차원 모델을 사용하여 부모 서열 및 다양한 개념적 조작된 생성물의 분석 과정에 의해 선택적으로 제조될 수 있다. 3차원 모델이 일반적으로 이용가능하며, 당업자에게 익숙하다. 선택된 후보 서열의 가능한 3차원 형태 구조를 예시하고 표시하며 가능한 면역원성을 측정할 수 있는 컴퓨터 프로그램(예컨대, 캘리포니아 몬로비아의 Xencor사의 Immunofilter 프로그램)을 이용할 수 있다. 이러한 디스플레이의 검사는 후보 서열의 기능에서 잔기의 가능한 역할의 분석, 즉 그의 항원에 결합

하는 후보 VHH 단백질의 능력에 영향을 미치는 잔기의 분석을 허용한다. 이러한 방식으로, 표적 항원(들)에 대한 친화성과 같은 원하는 특징이 달성되도록 부모 및 참조 서열로부터 잔기가 선택되고 조합될 수 있다. 대안적으로, 또는 상기 절차에 부가적으로, 다른 적합한 조작 방법이 사용될 수 있다. 유사한 단백질 또는 단편에의 특이적 결합에 대해 VHH를 스크리닝하는 것은 뉴클레오타이드(DNA 또는 RNA 디스플레이) 또는 펩타이드 디스플레이 라이브러리, 예를 들어, 시험관내 디스플레이를 사용하여 편리하게 달성될 수 있다. 어떤 단백질, 항체, 및 다른 길항제가 본 개시내용의 VHH 또는 VCAR과 표적 단백질에의 결합에 대해 경쟁하고/거나 에피토프 영역을 공유하는지 결정하기 위해 본 개시내용의 VHH 또는 VCAR을 이용하여 경쟁적 분석이 수행될 수 있다. 당업자에게 쉽게 알려진 바와 같은 이러한 분석은 단백질 상의 제한된 수의 결합 부위에 대한 길항제 또는 리간드 사이의 경쟁을 평가한다.

[0084] VH

[0085] 본 개시내용의 조성물(예컨대, 트랜스포존 또는 나노트랜스포존)은 CAR을 포함할 수 있고; 일부 양태에서, CAR의 항원 인식 영역은 특정 표적 단백질/항원을 인식하고 이에 결합하기 위해 적어도 하나의 단일 도메인 항체(SdAb)를 포함할 수 있다. 일 양태에서, 단일 도메인 항체는 VH이다. VH는 일반적인 IgG로부터 유래된 단일 도메인 결합제이다. 항원에 특이적으로 결합하는 VH는 세포(예컨대, 세포독성 면역 세포)의 특이성을 특이적 항원으로 향하게 하는 데 사용될 수 있다. 항원 인식 영역은 적어도 2개의 VH를 포함할 수 있다. 항원 인식 영역은 적어도 3개의 VH를 포함할 수 있다. 일 양태에서, 본 개시내용의 CAR은 2개의 구별되는 항원에 특이적으로 결합하는 적어도 2개의 VH를 포함하는 이중특이적 CAR이다.

[0086] VH는 인간 서열로부터 단리되거나 유래될 수 있다. VH는 인간 CDR 서열 및/또는 인간 프레임워크 서열 및 비인간 또는 인간화된 서열(예컨대, 랫트 Fc 도메인)을 포함할 수 있다. 일부 양태에서, VH는 완전히 인간화된 VH이다. 일부 양태에서, VH는 자연발생 항체도 아니고 자연발생 항체의 단편도 아니다. 일부 양태에서, VH는 단클론 항체의 단편이 아니다. 일부 양태에서, VH는 UniDab 항체(TeneoBio)이다. 일부 양태에서, VH는 Fc 도메인 또는 이의 일부를 제거하도록 변형된다. 일부 양태에서, VH의 프레임워크 서열은, 예를 들어, 발현을 개선하거나, 면역원성을 감소시키거나 기능을 개선하도록 변형된다.

[0087] VH는 VH를 생산하기 위해 UniRat(TeneoBio) 시스템 및 "NGS 기반 발견"을 사용하여 완전히 조작될 수 있다. 이 방법을 사용하면, 특정 VH는 자연 발생하지 않으며 완전 조작된 시스템을 사용하여 생성된다. VH는 숙주(예를 들어, 마우스, 랫트 또는 인간)로부터 직접 단리되거나 세포 또는 세포주의 단일 클론(하이브리도마)으로부터 직접 단리된 자연발생 단클론 항체(mAb)로부터 유래되지 않는다. 이러한 VH는 후속적으로 상기 세포주로부터 클로닝되지 않았다. 대신, VH 서열은 랫트 Fc 도메인을 갖는 인간 가변 영역(VH 도메인)을 포함하는 전이유전자(transgene)로서 UniRat 시스템을 사용하여 완전 조작되고, 따라서 경쇄가 없는 인간/랫트 키메라이며 표준 mAb 형식과 다르다. 천연 랫트 유전자는 넉아웃되고 랫트에서 발현되는 유일한 항체는 랫트 Fc(UniAbs)에 연결된 VH 도메인을 갖는 전이유전자로부터 유래된다. 이들은 UniRat에서 발현되는 독점적인 Ab이다. 차세대 시퀀싱(NGS) 및 생물정보학은 면역 후 UniRat에 의해 생성된 중쇄 항체의 전체 항원 특이적 레퍼토리를 확인하는 데 사용된다. 그리고 나서, 항체 레퍼토리 서열 정보를 다양한 기능에 대해 시험관내에서 스크리닝될 수 있는 완전 인간 중쇄 항체의 큰 집합으로 변환하기 위해 고유한 유전자 조립 방법이 사용된다. 일부 양태에서, 완전히 인간화된 VH는 인간 VH 도메인을 시험관내에서 인간 Fc와 융합함으로써(비자연발생 재조합 VH 항체를 생성하기 위해) 생성된다. 일부 양태에서, VH는 완전히 인간화되지만, 이들은 경쇄가 없는 인간/랫트 키메라(인간 VH, 랫트 Fc)로서 생체내에서 발현된다. 경쇄가 없는 인간/랫트 키메라(인간 VH, 랫트 Fc)로서 생체내에서 발현되는 완전히 인간화된 VH는 약 80kDa(대 150 kDa)이다.

[0088] 본 개시내용의 CAR은 $10^{-9}$ M 이하, $10^{-10}$ M 이하, $10^{-11}$ M 이하, $10^{-12}$ M 이하, $110^{-13}$ M 이하, $10^{-14}$ M 이하, 및 $10^{-15}$ M 이하의 $K_D$로부터 선택된 적어도 하나의 친화성으로 인간 항원에 결합할 수 있다. $K_D$는, 비제한적으로, 표면 플라즈몬 공명을 포함하는 임의의 수단에 의해 결정될 수 있다.

[0089] 일 양태에서, 개시된 CAR의 항원 인식 영역은 적어도 하나의 항-BCMA 센티린을 포함한다. 항-BCMA 센티린은 서열번호: 29와 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. 항-BCMA 센티린은 서열번호: 28과 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0090] 항-BCMA 센티린을 포함하는 CAR은 본원에서 BCMA CARTyrin으로 지칭된다. 바람직한 양태에서, BCMA CARTyrin은

서열번호: 30과 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. BCMA CARTyrin은 서열번호: 9와 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0091] BCMA CARTyrin을 포함하는 본 개시내용의 조성물(예컨대, 나노트랜스포존)은 서열번호: 17과 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. BCMA CARTyrin을 포함하는 본 개시내용의 조성물(예컨대, 나노트랜스포존)은 서열번호: 1과 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다. BCMA CARTyrin을 포함하는 본 개시내용의 조성물(예컨대, 나노트랜스포존)은 본원에서 P-BCMA-101-트랜스포존(도 13에 예시된 바와 같음)으로 지칭된다.

[0092] 일 양태에서, 개시된 CAR의 항원 인식 영역은 적어도 하나의 항-PSMA 센티린을 포함한다. 항-PSMA 센티린은 서열번호: 94와 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. 항-PSMA 센티린은 서열번호: 93과 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0093] 항-PSMA 센티린을 포함하는 CAR은 본원에서 PSMA CARTyrin으로 지칭된다. 바람직한 양태에서, PSMA CARTyrin은 서열번호: 95와 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. PSMA CARTyrin은 서열번호: 19와 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0094] PSMA CARTyrin을 포함하는 본 개시내용의 조성물(예컨대, 나노트랜스포존)은 서열번호: 20과 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. PSMA CARTyrin을 포함하는 본 개시내용의 조성물(예컨대, 나노트랜스포존)은 서열번호: 18과 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다. PSMA CARTyrin을 포함하는 본 개시내용의 조성물(예컨대, 나노트랜스포존)은 본원에서 P-PSMA-101 트랜스포존(도 14에 예시된 바와 같음)으로도 지칭된다.

[0095] 일 양태에서, 개시된 CAR의 항원 인식 영역은 적어도 하나의 항-BCMA VH를 포함한다. 항-BCMA VH는 서열번호: 97과 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. 항-BCMA VH는 서열번호: 96과 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0096] 항-BCMA VH를 포함하는 CAR은 본원에서 BCMA VCAR로 지칭된다. 바람직한 양태에서, BCMA VCAR은 서열번호: 98과 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. BCMA VCAR은 서열번호: 22와 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0097] BCMA VCAR을 포함하는 본 개시내용의 조성물(예컨대, 나노트랜스포존)은 서열번호: 23과 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. BCMA VCAR을 포함하는 본 개시내용의 조성물(예컨대, 나노트랜스포존)은 서열번호: 21과 적어도 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다. BCMA VCAR을 포함하는 본 개시내용의 조성물(예컨대, 나노트랜스포존)은 본원에서 P-BCMA-ALLO1-트랜스포존(도 15에 예시된 바와 같음)으로도 지칭된다.

[0098] 엑토도메인은 신호 펩타이드를 포함할 수 있다. 신호 펩타이드는 인간 CD2, CD3δ, CD3ε, CD3γ, CD3ζ, CD4, CD8α, CD19, CD28, 4-1BB 또는 GM-CSFR 신호 펩타이드를 코딩하는 서열을 포함할 수 있다. 바람직한 양태에서,

신호 펩타이드는 인간 CD8 알파(CD8α) 신호 펩타이드(SP) 또는 이의 일부를 포함하거나, 본질적으로 구성되거나, 또는 구성된다. 인간 CD8α SP는 서열번호: 31과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. 바람직하게는, 인간 CD8α SP는 서열번호: 31의 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다.

[0099] 인간 CD8α SP는 서열번호: 32와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다. 바람직하게는, 인간 CD8α SP는 서열번호: 32의 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0100] 힌지 도메인 또는 힌지 영역은 인간 CD8α, IgG4, CD4 서열, 또는 이의 조합을 포함할 수 있다. 바람직한 양태에서, 힌지는 인간 CD8 알파(CD8α) 힌지 또는 이의 일부를 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있다. 인간 CD8a 힌지는 서열번호: 35와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. 바람직하게는, 인간 CD8α 힌지 도메인은 서열번호: 35의 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다.

[0101] 인간 CD8α 힌지는 서열번호: 36과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다. 바람직하게는, 인간 CD8α 힌지 도메인은 서열번호: 36의 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0102] 막관통 도메인은 인간 CD2, CD3δ, CD3ε, CD3γ, CD3ζ, CD4, CD8α, CD19, CD28, 4-1BB 또는 GM-CSFR 막관통 도메인을 코딩하는 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있다. 바람직하게는, 막관통 도메인은 인간 CD8 알파(CD8α) 막관통 도메인, 또는 이의 일부를 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있다. CD8a 막관통 도메인은 서열번호: 37과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. 바람직하게는, 인간 CD8α 막관통 도메인은 서열번호: 37의 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다.

[0103] CD8α 막관통 도메인은 서열번호: 38과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다. 바람직하게는, CD8α 막관통 도메인은 서열번호: 38의 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0104] 적어도 하나의 공자극 도메인은 인간 4-1BB, CD28, CD3 제타(CD3ζ), CD40, ICOS, MyD88, OX-40 세포내 도메인, 또는 임의의 이의 조합을 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있다. 바람직하게는, 적어도 하나의 공자극 도메인은 CD3ζ, 4-1BB 공자극 도메인, 또는 이의 조합을 포함한다.

[0105] 4-1BB 세포내 도메인은 서열번호: 39와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. 바람직하게는, 4-1BB 세포내 도메인은 서열번호: 39의 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다.

[0106] 4-1BB 세포내 도메인은 서열번호: 40과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다. 바람직하게는, 4-1BB 세포내 도메인은 서열번호: 40의 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0107] CD3ζ 세포내 도메인은 서열번호: 41과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다. 바람직하게는, CD3ζ 세포내 도메인은 서열번호: 41의 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성된다.

[0108] CD3ζ 세포내 도메인은 서열번호: 42와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클

레오타이드에 의해 코딩된다. 바람직하게는, CD3ζ 세포내 도메인은 서열번호: 42의 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0109] **트랜스포존 및 벡터 조성물**

[0110] 전위 시스템

[0111] 본 개시내용은 (a) 제1 역위 말단 반복(ITR) 또는 제1 ITR을 코딩하는 서열, (b) 제2 ITR 또는 제2 ITR을 코딩하는 서열, 및 (c) 인트라-ITR 서열 또는 인트라-ITR을 코딩하는 서열을 포함하는 제1 핵산 서열로서, 인트라-ITR 서열은 트랜스포존 서열 또는 트랜스포존을 코딩하는 서열을 포함하는 것인 제1 핵산 서열; 및 인터-ITR 서열 또는 인터-ITR을 코딩하는 서열을 포함하는 제2 핵산 서열로서, 인터-ITR 서열의 길이는 700개 이하의 뉴클레오타이드인 것인 제2 핵산 서열을 포함하는 트랜스포존 또는 나노트랜스포존을 제공한다.

[0112] 본 개시내용의 트랜스포존 또는 나노트랜스포존은 단백질 스캐폴드(예컨대, 적어도 하나의 scFv, 단일 도메인 항체 또는 센티린을 포함하는 CAR)를 포함한다. 트랜스포존 또는 나노트랜스포존은 2개의 시스-조절성 인슐레이터 요소 측면에 있는 단백질 스캐폴드(예컨대, 적어도 하나의 scFv, 단일 도메인 항체 또는 센티린을 포함하는 CAR)를 코딩하는 서열을 포함하는 플라스미드 DNA 트랜스포존일 수 있다. 트랜스포존 또는 나노트랜스포존은 트랜스포사제를 코딩하는 서열을 포함하는 플라스미드를 추가로 포함할 수 있다. 트랜스포사제를 코딩하는 서열은 DNA 서열 또는 RNA 서열일 수 있다. 바람직하게는, 트랜스포사제를 코딩하는 서열은 mRNA 서열이다.

[0113] 본 개시내용의 트랜스포존 또는 나노트랜스포존은 piggyBac™(PB) 트랜스포존일 수 있다. 일부 양태에서, 트랜스포존이 PB 트랜스포존인 경우, 트랜스포사제는 piggyBac™(PB) 트랜스포사제 piggyBac-유사(PBL) 트랜스포사제 또는 Super piggyBac™(SPB) 트랜스포사제이다. 바람직하게는, SPB 트랜스포사제를 코딩하는 서열은 mRNA 서열이다.

[0114] PB 트랜스포존 및 PB, PBL 및 SPB 트랜스포사제의 비제한적인 예는 미국 특허 제6,218,182호; 미국 특허 제6,962,810호; 미국 특허 제8,399,643호 및 PCT 공개 WO 제2010/099296호에 상세히 기재되어 있다.

[0115] PB, PBL 및 SPB 트랜스포사제는 트랜스포존의 말단에 있는 트랜스포존 특이적 역위 말단 반복 서열(ITR)을 인식하고, 염색체 부위 내의 서열 5'-TTAT-3'(TTAT 표적 서열) 또는 염색체 부위 내의 서열 5'-TTAA-3'(TTAA 표적 서열)에 있는 ITR 사이에 내용물을 삽입한다. PB 또는 PBL 트랜스포존의 표적 서열은 5'-CTAA-3', 5'-TTAG-3', 5'-ATAA-3', 5'-TCAA-3', 5'AGTT-3', 5'-ATTA-3', 5'-GTTA-3', 5'-TTGA-3', 5'-TTTA-3', 5'-TTAC-3', 5'-ACTA-3', 5'-AGGG-3', 5'-CTAG-3', 5'-TGAA-3', 5'-AGGT-3', 5'-ATCA-3', 5'-CTCC-3', 5'-TAAA-3', 5'-TCTC-3', 5'TGAA-3', 5'-AAAT-3', 5'-AATC-3', 5'-ACAA-3', 5'-ACAT-3', 5'-ACTC-3', 5'-AGTG-3', 5'-ATAG-3', 5'-CAAA-3', 5'-CACA-3', 5'-CATA-3', 5'-CCAG-3', 5'-CCCA-3', 5'-CGTA-3', 5'-GTCC-3', 5'-TAAG-3', 5'-TCTA-3', 5'-TGAG-3', 5'-TGTT-3', 5'-TTCA-3'5'-TTCT-3' 및 5'-TTTT-3'를 포함하거나 구성될 수 있다. PB 또는 PBL 트랜스포존 시스템은 ITR 사이에 포함될 수 있는 관심 유전자에 대한 탑재력(payload) 제한이 없다.

[0116] 하나 이상의 PB, PBL 및 SPB 트랜스포사제에 대한 예시적인 아미노산 서열이 미국 특허 제6,218,185호; 미국 특허 제6,962,810호 및 미국 특허 제8,399,643호에 개시되어 있다. 바람직한 양태에서, PB 트랜스포사제는 서열번호: 63과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나 구성된다.

[0117] PB 또는 PBL 트랜스포사제는 서열번호: 63의 서열의 위치 30, 165, 282, 또는 538 중 2개 이상, 3개 이상 또는 각각에 아미노산 치환을 갖는 아미노산 서열을 포함하거나 구성될 수 있다. 트랜스포사제는 서열번호: 63의 서열의 아미노산 서열을 포함하거나 구성되는 SPB 트랜스포사제일 수 있고, 위치 30에서의 아미노산 치환은 이소류신(I)에 대한 발린(V)의 치환일 수 있고, 위치 165에서의 아미노산 치환은 글리신(G)에 대한 세린(S)의 치환일 수 있고, 위치 282에서의 아미노산 치환은 메티오닌(M)에 대한 발린(V)의 치환일 수 있고, 위치 538에서의 아미노산 치환은 아스파라긴(N)에 대한 리신(K)의 치환일 수 있다. 바람직한 양태에서, SPB 트랜스포사제는 서열번호: 64와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나 구성된다.

[0118] 트랜스포사제가 위치 30, 165, 282 및/또는 538에 상기 기재된 돌연변이를 포함하는 특정 양태에서, PB, PBL 및 SPB 트랜스포사제는 서열번호: 63 또는 서열번호: 64의 서열의 위치 3, 46, 82, 103, 119, 125, 177, 180, 185, 187, 200, 207, 209, 226, 235, 240, 241, 243, 258, 296, 298, 311, 315, 319, 327, 328, 340, 421, 436, 456, 470, 486, 503, 552, 570 및 591 중 하나 이상에 아미노산 치환을 추가로 포함할 수 있고, 이는 PCT

공개 WO 제2019/173636호 및 제PCT/US2019/049816호에 더 상세히 기재되어 있다.

[0119] PB, PBL 또는 SPB 트랜스포사제는 PCT 공개 WO 제2019/173636호 및 제PCT/US2019/049816호에 더 상세히 기재된 바와 같이 곤충, 척추동물, 갑각류 또는 미삭류(urochordate)로부터 단리되거나 유래될 수 있다. 바람직한 양태에서, PB, PBL 또는 SPB 트랜스포사제는 곤충 트리코플루시아 니(*Trichoplusia ni*)(GenBank 등록번호 AAA87375) 또는 봄빅스 모리(*Bombyx mori*)(GenBank 등록번호 BAD11135)로부터 단리되거나 유래된다.

[0120] 과활성 PB 또는 PBL 트랜스포사제는 그것이 유래된 자연발생 변이체보다 더 활성인 트랜스포사제이다. 바람직한 양태에서, 과활성 PB 또는 PBL 트랜스포사제는 봄빅스 모리 또는 제노푸스 트로피칼리스(*Xenopus tropicalis*)로부터 단리되거나 유래된다. 과활성 PB 또는 PBL 트랜스포사제의 예가 미국 특허 제6,218,185호; 미국 특허 제6,962,810호, 미국 특허 제8,399,643호 및 WO 제2019/173636호에 개시되어 있다. 과활성 아미노산 치환의 목록이 미국 특허 제10,041,077호에 개시되어 있다.

[0121] 일부 양태에서, PB 또는 PBL 트랜스포사제는 통합 부족(integration deficient)이다. 통합 부족 PB 또는 PBL 트랜스포사제는 그의 상응하는 트랜스포존을 절제할 수 있지만 상응하는 야생형 트랜스포사제보다 낮은 빈도로 절제된 트랜스포존을 통합하는 트랜스포사제이다. 통합 부족 PB 또는 PBL 트랜스포사제의 예는 미국 특허 제6,218,185호; 미국 특허 제6,962,810호, 미국 특허 제8,399,643호 및 WO 제2019/173636호에 개시되어 있다. 통합 부족 아미노산 치환의 목록은 미국 특허 제10,041,077호에 개시되어 있다.

[0122] 일부 양태에서, PB 또는 PBL 트랜스포사제는 핵 국소화 신호에 융합된다. 핵 국소화 신호에 융합된 PB 또는 PBL 트랜스포사제의 예는 미국 특허 제6,218,185호; 미국 특허 제6,962,810호, 미국 특허 제8,399,643호 및 WO 제2019/173636호에 개시되어 있다.

[0123] 본 개시내용의 트랜스포존 또는 나노트랜스포존은 슬리핑 뷰티(Sleeping Beauty) 트랜스포존일 수 있다. 일부 양태에서, 트랜스포존이 슬리핑 뷰티 트랜스포존인 경우, 트랜스포사제는 슬리핑 뷰티 트랜스포사제(예를 들어, 미국 특허 제9,228,180호에 개시된 바와 같음) 또는 과활성 슬리핑 뷰티(SB100X) 트랜스포사제이다. 바람직한 양태에서, 슬리핑 뷰티 트랜스포사제는 서열번호: 65와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나 구성된다. 바람직한 양태에서, 과활성 슬리핑 뷰티(SB100X) 트랜스포사제는 서열번호: 66과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나 구성된다.

[0124] 본 개시내용의 트랜스포존 또는 나노트랜스포존은 Helraiser 트랜스포존일 수 있다. 예시적인 Helraiser 트랜스포존은 서열번호: 67과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나 구성되는 Helibat1을 포함한다. 일부 양태에서, 트랜스포존이 Helraiser 트랜스포존인 경우, 트랜스포사제는 Helitron 트랜스포사제(예를 들어, WO 제2019/173636호에 개시된 바와 같음)이다. 바람직한 양태에서, Helitron 트랜스포사제는 서열번호: 68과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나 구성된다.

[0125] 본 개시내용의 트랜스포존 또는 나노트랜스포존은 Tol2 트랜스포존일 수 있다. 역위 반복, 서브말단 서열 및 Tol2 트랜스포사제를 포함하는 예시적인 Tol2 트랜스포존은 서열번호: 69와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산을 포함하거나 구성된다. 일부 양태에서, 트랜스포존이 Tol2 트랜스포존인 경우, 트랜스포사제는 Tol2 트랜스포사제(예를 들어, WO 제2019/173636호에 개시된 바와 같음)이다. 바람직한 양태에서, Tol2 트랜스포사제는 서열번호: 70과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나 구성된다.

[0126] 본 개시내용의 트랜스포존 또는 나노트랜스포존은 TcBuster 트랜스포존일 수 있다. 일부 양태에서, 트랜스포존이 TcBuster 트랜스포존인 경우, 트랜스포사제는 TcBuster 트랜스포사제 또는 과활성 TcBuster 트랜스포사제(예를 들어, WO 제2019/173636호에 개시된 바와 같음)이다. TcBuster 트랜스포사제는 자연발생 아미노산 서열 또는 비자연발생 아미노산 서열을 포함하거나 구성될 수 있다. 바람직한 양태에서, TcBuster 트랜스포사제는 서열번호: 71과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나 구성된다. TcBuster 트랜스포사제를 코딩하는 폴리뉴클레오타이드는 자연발생 핵산 서열 또는 비자연발생 핵산 서열을 포함하거나 구성될 수 있다. 바람직한 양태에서, TcBuster 트랜스포사제는 서열번호: 72와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임

의의 백분율) 동일한 핵산을 포함하거나 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0127] 일부 양태에서, 돌연변이체 TcBuster 트랜스포사제는 PCT 공개 WO 제2019/173636호 및 제PCT/US2019/049816호에 더 상세히 기재된 바와 같이 야생형 TcBuster 트랜스포사제와 비교하여 하나 이상의 서열 변이를 포함한다.

[0128] 본원에 개시된 세포 전달 조성물(예컨대, 트랜스포존)은 치료적 단백질 또는 치료제를 코딩하는 핵산을 포함할 수 있다. 치료적 단백질의 예는 PCT 공개 WO 제2019/173636호 및 제PCT/US2019/049816호에 개시된 것을 포함한다.

[0129] 벡터 시스템

[0130] 일부 양태에서, 본 개시내용의 조성물(예컨대, 나노트랜스포존)은 또 다른 트랜스포존 또는 나노트랜스포존과 조합되어 또는 벡터와 조합되어 이용될 수 있다. 본 개시내용의 벡터는 바이러스 벡터 또는 재조합 벡터일 수 있다. 바이러스 벡터는 레트로바이러스, 렌티바이러스, 아데노바이러스, 아데노 관련 바이러스 또는 임의의 이의 조합으로부터 단리되거나 유래된 서열을 포함할 수 있다. 바이러스 벡터는 아데노 관련 바이러스(AAV)로부터 단리되거나 유래된 서열을 포함할 수 있다. 바이러스 벡터는 재조합 AAV(rAAV)를 포함할 수 있다. 예시적인 아데노 관련 바이러스 및 재조합 아데노 관련 바이러스는 본 개시내용의 scFv 또는 CAR을 코딩하는 서열 옆에 cis로 위치한 2개 이상의 역위 말단 반복(ITR) 서열을 포함한다. 예시적인 아데노 관련 바이러스 및 재조합 아데노 관련 바이러스는, 비제한적으로 모든 혈청형(예컨대, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, 및 AAV9)을 포함한다. 예시적인 아데노 관련 바이러스 및 재조합 아데노 관련 바이러스는, 비제한적으로, 자가 상보적 AAV(scAAV) 및 하나의 혈청형의 게놈 및 또 다른 혈청형(예컨대, AAV2/5, AAV-DJ 및 AAV-DJ8)의 캡시드를 함유하는 AAV 하이브리드를 포함한다. 예시적인 아데노 관련 바이러스 및 재조합 아데노 관련 바이러스는, 비제한적으로, rAAV-LK03을 포함한다.

[0131] 본 개시내용의 벡터는 나노입자일 수 있다. 나노입자 벡터의 비제한적인 예는 핵산(예컨대, RNA, DNA, 합성 뉴클레오타이드, 변형된 뉴클레오타이드 또는 이의 임의의 조합), 아미노산(L-아미노산, D-아미노산, 합성 아미노산, 변형된 아미노산, 또는 이의 임의의 조합), 중합체(예컨대, 폴리머좀), 마이셀, 지질(예컨대, 리포좀), 유기 분자(예컨대, 탄소 원자, 시트, 섬유, 튜브), 무기 분자(예컨대, 인산 칼슘 또는 금) 또는 이의 임의의 조합을 포함한다. 나노입자 벡터는 세포막을 통해 수동적으로 또는 능동적으로 수송될 수 있다.

[0132] 본원에 개시된 세포 전달 조성물(예컨대, 트랜스포존, 벡터)은 치료적 단백질 또는 치료제를 코딩하는 핵산을 포함할 수 있다. 치료적 단백질의 예는 PCT 공개 WO 제2019/173636호 및 제PCT/US2019/049816호에 개시된 것을 포함한다.

[0133] **본 개시내용의 세포 및 변형된 세포**

[0134] 본 개시내용의 세포 및 변형된 세포는 포유동물 세포일 수 있다. 바람직하게는, 세포 및 변형된 세포는 인간 세포이다. 본 개시내용의 세포 및 변형된 세포는 면역 세포일 수 있다. 본 개시내용의 면역 세포는 림프성 선조 세포(lymphoid progenitor cell), 자연 살해(NK) 세포, T 림프구(T-세포), 줄기 기억 T 세포($T_{SCM}$ 세포), 중심 기억 T 세포($T_{CM}$), 줄기 세포 유사 T 세포, B 림프구(B-세포), 항원 제시 세포(APC), 사이토카인 유도 사멸(CIK) 세포, 골수성 선조 세포, 호중구, 호염기구, 호산구, 단핵구, 대식세포, 혈소판, 적혈구(erythrocyte), 적혈구(RBC), 거핵구 또는 파골세포를 포함할 수 있다.

[0135] 면역 전구체 세포는 하나 이상의 유형의 면역 세포로 분화할 수 있는 임의의 세포를 포함할 수 있다. 면역 전구체 세포는 자가 재생하여 면역 세포로 발달할 수 있는 다능성 줄기 세포를 포함할 수 있다. 면역 전구체 세포는 조혈 줄기 세포(HSC) 또는 이의 후손을 포함할 수 있다. 면역 전구체 세포는 면역 세포로 발달할 수 있는 전구체 세포를 포함할 수 있다. 면역 전구체 세포는 조혈 선조 세포(HPC)를 포함할 수 있다.

[0136] 조혈 줄기 세포(HSC)는 다능성, 자가 재생 세포이다. 림프성 및 골수성 계통의 모든 분화된 혈액 세포는 HSC로부터 발생한다. HSC는 성인 골수, 말초 혈액, 동원된 말초 혈액, 복막 투석 유출물 및 제대혈에서 발견될 수 있다.

[0137] HSC는 1차 또는 배양된 줄기 세포로부터 단리되거나 유래될 수 있다. HSC는 배아 줄기 세포, 다능성(multipotent) 줄기 세포, 다능성(pluripotent) 줄기 세포, 성체 줄기 세포, 또는 유도된 다능성 줄기 세포(iPSC)로부터 단리되거나 유래될 수 있다.

[0138] 면역 전구체 세포는 HSC 또는 HSC 후손 세포를 포함할 수 있다. HSC 후손 세포의 비제한적인 예는 다능성 줄기

세포, 림프성 선조 세포, 자연 살해(NK) 세포, T 림프구 세포(T-세포), B 림프구 세포(B-세포), 골수성 선조 세포, 호중구, 호염기구, 호산구, 단핵구 및 대식세포를 포함한다.

[0139] 개시된 방법에 의해 생산된 HSC는, 성체 줄기 세포로부터 단리되거나 유래되는 동안 그리고 단일 계통에 수임되는 동안, 배아 줄기 세포의 특징을 공유하는 "원시(primitive)" 줄기 세포의 특징을 보유할 수 있다. 예를 들어, 개시된 방법에 의해 생산된 "원시" HSC는 분열 후에 이들의 "줄기성(stemness)"을 유지하고 분화하지 않는다. 결과적으로, 입양 세포 요법으로서, 개시된 방법에 의해 생산된 "원시" HSC는 그들의 수를 보충할뿐만 아니라 생체내에서 확장된다. 개시된 방법에 의해 생산된 "원시" HSC는 단일 용량으로서 투여될 때 치료적으로 효과적일 수 있다.

[0140] 원시 HSC는 CD34+일 수 있다. 원시 HSC는 CD34+ 및 CD38-일 수 있다. 원시 HSC는 CD34+, CD38- 및 CD90+일 수 있다. 원시 HSC는 CD34+, CD38-, CD90+ 및 CD45RA-일 수 있다. 원시 HSC는 CD34+, CD38-, CD90+, CD45RA-, 및 CD49f+일 수 있다. 원시 HSC는 CD34+, CD38-, CD90+, CD45RA-, 및 CD49f+일 수 있다

[0141] 원시 HSC, HSC, 및/또는 HSC 후손 세포는 외인성 서열(예컨대, 키메라 항원 수용체 또는 치료적 단백질)을 발현하기 위해 개시된 방법에 따라 변형될 수 있다. 변형된 원시 HSC, 변형된 HSC, 및/또는 변형된 HSC 후손 세포는, 비제한적으로, 변형된 T 세포, 변형된 자연 살해 세포 및/또는 변형된 B-세포를 포함하는 변형된 면역 세포를 생산하기 위해 전방 분화될 수 있다.

[0142] 변형된 면역 또는 면역 전구체 세포는 NK 세포일 수 있다. NK 세포는 림프성 선조 세포로부터 분화하는 세포독성 림프구일 수 있다. 변형된 NK 세포는 변형된 조혈 줄기 및 선조 세포(HSPC) 또는 변형된 HSC로부터 유래될 수 있다. 일부 양태에서, 활성화되지 않은 NK 세포는 CD3-고갈된 백혈구채집술(CD14/CD19/CD56+ 세포 함유)로부터 유래된다.

[0143] 변형된 면역 또는 면역 전구체 세포는 B 세포일 수 있다. B 세포는 세포 표면 상에 B 세포 수용체를 발현하는 림프구의 유형이다. B 세포 수용체는 특정 항원에 결합한다. 변형된 B 세포는 변형된 조혈 줄기 및 선조 세포(HSPC) 또는 변형된 HSC로부터 유래될 수 있다.

[0144] 본 개시내용의 변형된 T 세포는 변형된 조혈 줄기 및 선조 세포(HSPC) 또는 변형된 HSC로부터 유래될 수 있다. 전통적인 생물학적 제제 및 화학치료제와 달리, 개시된 변형된-T 세포는 항원 인식시 빠르게 재생하는 능력을 가지고 있어, 잠재적으로 반복 치료의 필요를 제거한다. 이를 달성하기 위해, 일부 양태에서, 변형된-T 세포는 초기 반응을 유도할뿐만 아니라 생존가능한 기억 T 세포의 안정적인 집단으로서 환자에서 지속되어 잠재적인 재발을 방지한다. 대안적으로, 일부 양태에서, 원하지 않는 경우, 변형된-T 세포는 환자에서 지속되지 않는다.

[0145] 항원 독립적(tonic) 신호전달을 통해 T 세포 고갈을 유발하지 않는 항원 수용체 분자의 개발뿐만 아니라 초기 기억 T 세포, 특히 줄기 세포 기억($T_{SCM}$) 또는 줄기 세포-유사 T 세포를 함유하는 변형된 T 세포 생성물의 개발에 집중적인 노력이 이뤄졌다. 본 개시내용의 줄기 세포-유사 변형된 T 세포는 중심 기억($T_{CM}$) T 세포 또는 $T_{CM}$ 유사 세포, 효과기 기억($T_{EM}$) 및 효과기 T 세포($T_E$)를 유도하는 자가 재생을 위한 가장 큰 능력 및 다능성 능력을 나타내어 더 나은 종양 박멸 및 장기간 변형된 T 세포 생착(engraftment)을 생성한다. 선형 분화 경로는 이러한 세포인 나이브 T 세포($T_N$) > $T_{SCM}$ > $T_{CM}$ > $T_{EM}$ > $T_E$ > $T_{TE}$를 생성하는 것을 담당할 수 있으며, 여기서 $T_N$은 $T_{SCM}$을 직접 발생시키는 부모 전구체 세포이며, 이는 이후에 차례로 $T_{CM}$ 등을 발생시킨다. 본 개시내용의 T 세포의 조성물은 각각의 부모 T 세포 서브세트 중 하나 이상을 포함할 수 있고, $T_{SCM}$ 세포가 가장 풍부하다(예컨대, $T_{SCM}$ > $T_{CM}$ > $T_{EM}$ > $T_E$ > $T_{TE}$).

[0146] 면역 세포 전구체는 초기 기억 T 세포, 줄기 세포 유사 T-세포, 나이브 T 세포($T_N$), $T_{SCM}$, $T_{CM}$, $T_{EM}$, $T_E$, 또는 $T_{TE}$로 분화될 수 있거나 분화할 수 있다. 면역 세포 전구체는 본 개시내용의 원시 HSC, HSC, 또는 HSC 후손 세포일 수 있다. 면역 세포는 초기 기억 T 세포, 줄기 세포 유사 T-세포, 나이브 T 세포($T_N$), $T_{SCM}$, $T_{CM}$, $T_{EM}$, $T_E$, 또는 $T_{TE}$일 수 있다.

[0147] 본 개시내용의 방법은 변형된 T 세포 집단을 변형 및/또는 생산할 수 있으며, 집단 내의 복수의 변형된 T 세포 중 적어도 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% 또는 이들 사이의 임의의 백분율은 초기 기억 T 세포의 하나 이상의 세포 표면 마커(들)를 발현한다. 변형된 초기 기억 T 세포 집단은 복수의 변형된 줄기 세포-유사 T 세포를 포함한다. 변형된 초기 기억 T 세포 집단

은 복수의 변형된 $T_{SCM}$ 세포를 포함한다. 변형된 초기 기억 T 세포 집단은 복수의 변형된 $T_{CM}$ 세포를 포함한다.

[0148] 본 개시내용의 방법은 변형된 T 세포 집단을 변형 및/또는 생산할 수 있으며, 집단 내의 복수의 변형된 T 세포 중 적어도 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% 또는 이들 사이의 임의의 백분율은 줄기 세포-유사 T 세포의 하나 이상의 세포 표면 마커(들)를 발현한다. 변형된 줄기 세포-유사 T 세포 집단은 복수의 변형된 $T_{SCM}$ 세포를 포함한다. 변형된 줄기 세포-유사 T 세포 집단은 복수의 변형된 $T_{CM}$ 세포를 포함한다.

[0149] 일부 양태에서, 집단 내의 복수의 변형된 T 세포 중 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적 포ㅎ어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100% 또는 이들 사이의 임의의 백분율은 줄기 기억 T 세포($T_{SCM}$) 또는 $T_{SCM}$-유사 세포의 하나 이상의 세포 표면 마커(들)를 발현하며; 하나 이상의 세포 표면 마커(들)는 CD45RA 및 CD62L을 포함한다. 세포 표면 마커는 CD62L, CD45RA, CD28, CCR7, CD127, CD45RO, CD95, CD95 및 IL-2Rβ 중 하나 이상을 포함할 수 있다. 세포 표면 마커는 CD45RA, CD95, IL-2Rβ, CCR7, 및 CD62L 중 하나 이상을 포함할 수 있다.

[0150] 일부 양태에서, 집단 내의 복수의 변형된 T 세포 중 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 또는 100%는 중심 기억 T 세포($T_{CM}$) 또는 $T_{CM}$-유사 세포의 하나 이상의 세포 표면 마커(들)를 발현하며; 하나 이상의 세포 표면 마커(들)는 CD45RO 및 CD62L을 포함한다. 세포 표면 마커는 CD45RO, CD95, IL-2Rβ, CCR7, 및 CD62L 중 하나 이상을 포함할 수 있다.

[0151] 본 개시내용의 방법은 변형된 T 세포 집단을 변형 및/또는 생산할 수 있으며, 집단 내의 복수의 변형된 T 세포 중 적어도 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% 또는 이들 사이의 임의의 백분율은 나이브 T 세포($T_N$)의 하나 이상의 세포 표면 마커(들)를 발현한다. 세포 표면 마커는 CD45RA, CCR7 및 CD62L 중 하나 이상을 포함할 수 있다.

[0152] 본 개시내용의 방법은 변형된 T 세포 집단을 변형 및/또는 생산할 수 있으며, 집단 내의 복수의 변형된 T 세포 중 적어도 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% 또는 이들 사이의 임의의 백분율은 효과기 T-세포(변형된 $T_{EFF}$)의 하나 이상의 세포 표면 마커(들)를 발현한다. 세포 표면 마커는 CD45RA, CD95, 및 IL-2Rβ 중 하나 이상을 포함할 수 있다.

[0153] 본 개시내용의 방법은 변형된 T 세포 집단을 변형 및/또는 생산할 수 있으며, 집단 내의 복수의 변형된 T 세포 중 적어도 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 99% 또는 이들 사이의 임의의 백분율은 줄기 세포-유사 T 세포, 줄기 기억 T 세포($T_{SCM}$) 또는 중심 기억 T 세포($T_{CM}$)의 하나 이상의 세포 표면 마커(들)를 발현한다.

[0154] 집단의 복수의 변형된 세포는 전이유전자 또는 전이유전자를 코딩하는 서열(예컨대, CAR)을 포함하고, 집단의 복수의 세포 중 적어도 75%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 적어도 99.9% 또는 100%는 전이유전자 또는 전이유전자를 코딩하는 서열을 포함하며, 변형된 세포 집단 중 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 적어도 99.9% 또는 100%는 CD34를 포함하는 하나 이상의 세포 표면 마커(들)를 발현하거나 또는 변형된 세포 집단 중 적어도 약 70% 내지 약 99%, 약 75% 내지 약 95% 또는 약 85% 내지 약 95%는 CD34를 포함하는 하나 이상의 세포 표면 마커(들)를 발현한다(예컨대, 세포 표면 마커 표현형 CD34+를 발현한다).

[0155] 집단의 복수의 변형된 세포는 전이유전자 또는 전이유전자를 코딩하는 서열(예컨대, CAR)을 포함하고, 집단의 복수의 세포 중 적어도 75%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 적어도 99.9% 또는 100%는 전이유전자 또는 전이유전자를 코딩하는 서열을 포함하며, 변형된 세포 집단 중 적어도 40%, 적어도 45%, 적어도 50%,

적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 적어도 99.9% 또는 100%는 CD34를 포함하는 하나 이상의 세포 표면 마커(들)를 포함하고 CD38을 포함하는 하나 이상의 세포 표면 마커(들)를 발현하지 않거나, 또는 변형된 세포 집단 중 적어도 약 45% 내지 약 90%, 약 50% 내지 약 80% 또는 약 65% 내지 약 75%는 CD34를 포함하는 하나 이상의 세포 표면 마커(들)를 발현하고 CD38을 포함하는 하나 이상의 세포 표면 마커(들)를 발현하지 않는다(예컨대, 세포 표면 마커 표현형 CD34+ 및 CD38-를 포함한다).

[0156] 집단의 복수의 변형된 세포는 전이유전자 또는 전이유전자를 코딩하는 서열(예컨대, CAR)을 포함하고, 집단의 복수의 세포 중 적어도 75%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 적어도 99.9% 또는 100%는 전이유전자 또는 전이유전자를 코딩하는 서열을 포함하며, 변형된 세포 집단중 적어도 0.1%, 적어도 0.2%, 적어도 0.3%, 적어도 0.4%, 적어도 0.5%, 적어도 0.6%, 적어도 0.7%, 적어도 0.8%, 적어도 0.9%, 적어도 1%, 적어도 1.5%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 적어도 99.9% 또는 100%는 CD34 및 CD90을 포함하는 하나 이상의 세포 표면 마커(들)를 발현하고 CD38을 포함하는 하나 이상의 세포 표면 마커(들)를 발현하지 않거나, 또는 변형된 세포 집단 중 적어도 약 0.2% 내지 약 40%, 약 0.2% 내지 약 30%, 약 0.2% 내지 약 2% 또는 0.5% 내지 약 1.5%는 CD34 및 CD90을 포함하는 하나 이상의 세포 표면 마커(들)를 발현하고 CD38을 포함하는 하나 이상의 세포 표면 마커(들)를 발현하지 않는다(예컨대, 세포 표면 마커 표현형 CD34+, CD38- 및 CD90+를 포함한다).

[0157] 집단의 복수의 변형된 세포는 전이유전자 또는 전이유전자를 코딩하는 서열(예컨대, CAR)을 포함하고, 집단의 복수의 세포 중 적어도 75%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 적어도 99.9% 또는 100%는 전이유전자 또는 전이유전자를 코딩하는 서열을 포함하며, 변형된 세포 집단 중 적어도 0.1%, 적어도 0.2%, 적어도 0.3%, 적어도 0.4%, 적어도 0.5%, 적어도 0.6%, 적어도 0.7%, 적어도 0.8%, 적어도 0.9%, 적어도 1%, 적어도 1.5%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 적어도 99.9% 또는 100%는 CD34 및 CD90을 포함하는 하나 이상의 세포 표면 마커(들)를 발현하고 CD38 및 CD45RA를 포함하는 하나 이상의 세포 표면 마커(들)를 발현하지 않거나, 또는 변형된 세포 집단 중 적어도 약 0.2% 내지 약 40%, 약 0.2% 내지 약 30%, 약 0.2% 내지 약 2% 또는 0.5% 내지 약 1.5%는 CD34 및 CD90을 포함하는 하나 이상의 세포 표면 마커(들)를 발현하고 CD38 및 CD45RA를 포함하는 하나 이상의 세포 표면 마커(들)를 발현하지 않는다(예컨대, 세포 표면 마커 표현형 CD34+, CD38-, CD90+, CD45RA-를 포함한다).

[0158] 집단의 복수의 변형된 세포는 전이유전자 또는 전이유전자를 코딩하는 서열(예컨대, CAR)을 포함하고, 집단의 복수의 세포 중 적어도 75%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 적어도 99.9% 또는 100%는 전이유전자 또는 전이유전자를 코딩하는 서열을 포함하며, 변형된 세포 집단 중 적어도 0.01%, 적어도 0.02%, 적어도 0.03%, 적어도 0.04%, 적어도 0.05%, 적어도 0.06%, 적어도 0.07%, 적어도 0.08%, 적어도 0.09%, 적어도 0.1%, 적어도 0.2%, 적어도 0.3%, 적어도 0.4%, 적어도 0.5%, 적어도 0.6%, 적어도 0.7%, 적어도 0.8%, 적어도 0.9%, 적어도 1%, 적어도 1.5%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 적어도 99.9% 또는 100%는 CD34, CD90 및 CD49f를 포함하는 하나 이상의 세포 표면 마커(들)를 발현하고 CD38 및 CD45RA를 포함하는 하나 이상의 세포 표면 마커(들)를 발현하지 않거나, 또는 변형된 세포 집단 중 적어도 약 0.02% 내지 약 30%, 약 0.02% 내지 약 2%, 약 0.04% 내지 약 2% 또는 약 0.04% 내지 약 1%는 CD34, CD90 및 CD49f를 포함하는 하나 이상의 세포 표면 마커(들)를 발현하고 CD38 및 CD45RA를 포함하는 하나 이상의 세포 표면 마커(들)를 발

현하지 않는다(예컨대, 세포 표면 마커 표현형 CD34+, CD38-, CD90+, CD45RA- 및 CD49f+를 포함한다).

[0159] 집단의 복수의 변형된 세포는 전이유전자 또는 전이유전자를 코딩하는 서열(예컨대, CAR)을 포함하고, 집단의 복수의 세포 중 적어도 75%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 적어도 99.9% 또는 100%는 전이유전자 또는 전이유전자를 코딩하는 서열을 포함하며, 변형된 세포 집단 중 적어도 0.01%, 적어도 0.02%, 적어도 0.03%, 적어도 0.04%, 적어도 0.05%, 적어도 0.06%, 적어도 0.07%, 적어도 0.08%, 적어도 0.09%, 적어도 0.1%, 적어도 0.2%, 적어도 0.3%, 적어도 0.4%, 적어도 0.5%, 적어도 0.6%, 적어도 0.7%, 적어도 0.8%, 적어도 0.9%, 적어도 1%, 적어도 1.5%, 적어도 2%, 적어도 3%, 적어도 4%, 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 91%, 적어도 92%, 적어도 93%, 적어도 94%, 적어도 95%, 적어도 96%, 적어도 97%, 적어도 98%, 적어도 99%, 적어도 99.5%, 적어도 99.9% 또는 100%는 CD34 및 CD90을 포함하는 하나 이상의 세포 표면 마커(들)를 발현하고 CD45RA를 포함하는 하나 이상의 세포 표면 마커(들)를 발현하지 않거나, 또는 변형된 세포 집단 중 적어도 약 0.2% 내지 약 5%, 약 0.2% 내지 약 3% 또는 약 0.4% 내지 약 3%는 CD34 및 CD90을 포함하는 하나 이상의 세포 표면 마커(들)를 발현하고 CD45RA를 포함하는 하나 이상의 세포 표면 마커(들)를 발현하지 않는다(예컨대, 세포 표면 마커 표현형 CD34+, CD90+ 및 CD45RA-를 포함한다).

[0160] 면역 세포 또는 면역 전구체 세포(예컨대, 개시된 변형된 T-세포)를 생산 및/또는 확장하는 조성물 및 방법 및 면역 세포 또는 면역 전구체 세포(예컨대, 개시된 변형된 T-세포)의 세포 생존력 및/또는 줄기 유사 표현형의 수준을 유지 또는 향상시키기 위한 완충제가 본원의 다른 곳에 개시되어 있으며 미국 특허 제10,329,543호 및 PCT 공개 WO 제2019/173636호에 더 상세히 개시되어 있다.

[0161] 본 개시내용의 세포 및 변형된 세포는 체세포일 수 있다. 본 개시내용의 세포 및 변형된 세포는 분화된 세포일 수 있다. 본 개시내용의 세포 및 변형된 세포는 자가조직 세포 또는 동종이계 세포일 수 있다. 동종이계 세포는 대상체에게 투여 후 생착에 대한 부작용을 방지하도록 조작된다. 동종이계 세포는 임의의 유형의 세포일 수 있다. 동종이계 세포는 줄기 세포일 수 있거나 또는 줄기 세포로부터 유래될 수 있다. 동종이계 세포는 분화된 체세포일 수 있다.

[0162] **키메라 항원 수용체를 발현하는 방법**

[0163] 본 개시내용은 세포의 표면 상에 CAR을 발현시키는 방법을 제공한다. 방법은 (a) 세포 집단을 수득하는 단계; (b) 세포 집단에서 적어도 하나의 세포의 세포막을 가로질러 CAR을 전달하기에 충분한 조건 하에 세포 집단을 CAR 또는 CAR을 코딩하는 서열을 포함하는 조성물에 접촉시켜, 변형된 세포 집단을 생성하는 단계; (c) CAR을 코딩하는 서열의 통합에 적합한 조건 하에 배양하는 단계; 및 (d) 세포 표면 상에 CAR을 발현하는 변형된 세포 집단으로부터 적어도 하나의 세포를 확장 및/또는 선택하는 단계를 포함한다.

[0164] 일부 양태에서, 세포 집단은 백혈구 및/또는 CD4+ 및 CD8+ 백혈구를 포함할 수 있다. 세포 집단은 최적화된 비율로 CD4+ 및 CD8+ 백혈구를 포함할 수 있다. CD4+ 대 CD8+ 백혈구의 최적화된 비율은 생체내에서 자연적으로 발생하지 않는다. 세포 집단은 종양 세포를 포함할 수 있다.

[0165] 일부 양태에서, 세포 집단에서 적어도 하나의 세포의 세포막을 가로질러 CAR 또는 CAR을 코딩하는 서열, 트랜스포존, 또는 벡터를 전달하기에 충분한 조건은 지정된 전압에서 전기의 하나 이상의 펄스의 적용, 완충제, 및 하나 이상의 추가 인자(들) 중 적어도 하나를 포함한다. 일부 양태에서, CAR을 코딩하는 서열의 통합에 적합한 조건은 완충제 및 하나 이상의 추가 인자(들) 중 적어도 하나를 포함한다.

[0166] 완충제는 PBS, HBSS, OptiMEM, BTXpress, Amaxa Nucleofector, 인간 T 세포 핵감염(nucleofection) 완충제 또는 임의의 이의 조합을 포함할 수 있다. 하나 이상의 추가 인자(들)는 (a) 재조합 인간 사이토카인, 케모카인, 인터루킨 또는 임의의 이의 조합; (b) 염, 미네랄, 대사산물 또는 임의의 이의 조합; (c) 세포 배지; (d) 세포 DNA 감지, 대사, 분화, 신호 전달, 하나 이상의 아폽토시스 경로(들) 또는 이의 조합의 억제제; 및 (e) 하나 이상의 핵산을 변형 또는 안정화시키는 시약을 포함할 수 있다. 재조합 인간 사이토카인, 케모카인, 인터루킨 또는 이의 임의의 조합은 IL2, IL7, IL12, IL15, IL21, IL1, IL3, IL4, IL5, IL6, IL8, CXCL8, IL9, IL10, IL11, IL13, IL14, IL16, IL17, IL18, IL19, IL20, IL22, IL23, IL25, IL26, IL27, IL28, IL29, IL30, IL31, IL32, IL33, IL35, IL36, GM-CSF, IFN-감마, IL-1 알파/IL-1F1, IL-1 베타/IL-1F2, IL-12 p70, IL-12/IL-35 p35, IL-13, IL-17/IL-17A, IL-17A/F 이종이량체, IL-17F, IL-18/IL-1F4, IL-23, IL-24, IL-32, IL-32 베타,

IL-32 감마, IL-33, LAP(TGF-베타 1), 림포톡신-알파/TNF-베타, TGF-베타, TNF-알파, TRANCE/TNFSF11/RANK L 또는 이의 임의의 조합을 포함할 수 있다. 염, 미네랄, 대사산물 또는 이의 임의의 조합은 HEPES, 니코틴아미드, 헤파린, 피루브산 나트륨, L-글루타민, MEM 비필수 아미노산 용액, 아스코르브산, 뉴클레오사이드, FBS/FCS, 인간 혈청, 혈청-대체물, 항생제, pH 조절제, Earle의 염, 2-머캅토에탄올, 인간 트랜스페린, 재조합 인간 인슐린, 인간 혈청 알부민, 뉴클레오펙터 PLUS 보충제, KCL, $MgCl_2$, $Na_2HPO_4$, $NAH_2PO_4$, 나트륨 락토비오네이트(Sodium lactobionate), 만니톨, 나트륨 석시네이트, 염화 나트륨, Na, 글루코스, $Ca(NO_3)_2$, Tris/HCl, $K_2HPO_4$, $KH_2PO_4$, 폴리에틸렌이민, 폴리-에틸렌-글리콜, 폴록사머 188, 폴록사머 181, 폴록사머 407, 폴리-비닐피롤리돈, Pop313, Crown-5, 또는 임의의 이의 조합을 포함할 수 있다. 세포 배지는 PBS, HBSS, OptiMEM, DMEM, RPMI 1640, AIM-V, X-VIVO 15, CellGro DC 배지, CTS OpTimizer T 세포 확장 SFM, TexMACS 배지, PRIME-XV T 세포 확장 배지, ImmunoCult-XF T 세포 확장 배지 또는 임의의 이의 조합을 포함할 수 있다. 세포 DNA 감지, 대사, 분화, 신호 전달, 하나 이상의 아폽토시스 경로(들) 또는 이의 조합의 억제제는 TLR9, MyD88, IRAK, TRAF6, TRAF3, IRF-7, NF-KB, 유형 1 인터페론, 전염증성 사이토카인, cGAS, STING, Sec5, TBK1, IRF-3, RNA pol III, RIG-1, IPS-1, FADD, RIP1, TRAF3, AIM2, ASC, Caspase1, Pro-IL1B, PI3K, Akt, Wnt3A의 억제제, 글리코겐 신타아제 키나아제-3β(GSK-3β)의 억제제(예컨대 TWS119), 또는 이의 임의의 조합을 포함한다. 이러한 억제제의 예는 바필로마이신(Bafilomycin), 클로로퀸(Chloroquine), 퀴나크린(Quinacrine), AC-YVAD-CMK, Z-VAD-FMK, Z-IETD-FMK 또는 이의 임의의 조합을 포함할 수 있다. 하나 이상의 핵산을 변형 또는 안정화시키는 시약은 pH 변형제, DNA 결합 단백질, 지질, 인지질, CaPO4, NLS 서열을 갖거나 갖지 않는 순 중성 전하 DNA 결합 펩타이드, TREX1 효소 또는 이의 임의의 조합을 포함한다.

[0167] 확장 및 선택 단계는 동시에 또는 순차적으로 발생할 수 있다. 확장은 선택하기 전에 발생할 수 있다. 확장은 선택 후에 발생할 수 있고, 선택적으로, 추가(즉, 두 번째) 선택이 확장 후에 발생할 수 있다. 동시 확장 및 선택은 동시에 발생할 수 있다. 확장 및/또는 선택 단계는 종점을 포함하여 10 내지 14일 동안 진행될 수 있다.

[0168] 확장은 변형된 세포 집단의 적어도 하나의 세포를 항원과 접촉시켜 CAR을 통해 적어도 하나의 세포를 자극함으로써 확장된 세포 집단을 생성하는 것을 포함할 수 있다. 항원은 기질의 표면 상에 제시될 수 있다. 기질은, 비제한적으로 표면, 웰, 비드 또는 이들의 복수, 및 매트릭스를 포함하는 임의의 형태를 가질 수 있다. 기질은 상자성 또는 자성 구성요소를 추가로 포함할 수 있다. 항원은 기질의 표면 상에 제시될 수 있고, 기질은 자성 비드이며, 변형되고 확장된 세포 집단으로부터 자성 비드를 제거하거나 분리하기 위해 자석이 사용될 수 있다. 항원은 세포 또는 인공 항원 제시 세포의 표면 상에 제시될 수 있다. 인공 항원 제시 세포는, 비제한적으로, 종양 세포 및 줄기 세포를 포함할 수 있다.

[0169] 트랜스포존 또는 벡터가 선택 유전자를 포함하는 일부 양태에서, 선택 단계는 변형된 세포 집단의 적어도 하나의 세포를 선택 유전자가 내성을 부여하는 화합물과 접촉시켜, 선택 유전자를 발현하는 세포를 선택에서 생존하는 것으로서 확인하고 선택 유전자를 발현하지 못하는 세포를 선택 단계에서 생존하지 못하는 것으로 확인하는 것을 포함한다.

[0170] 본 개시내용은 본원에 기재된 방법의 변형된, 확장된 및 선택된 세포 집단을 포함하는 조성물을 제공한다.

[0171] 세포의 표면 상에 CAR을 발현하기 위한 방법의 보다 상세한 설명은 PCT 공개 WO 제2019/049816호 및 제PCT/US2019/049816호에 개시되어 있다.

[0172] 본 개시내용은 세포 또는 세포 집단으로서, 세포는 (a) 유도성 프로모터를 코딩하는 서열 및 전이유전자를 코딩하는 서열을 포함하는 유도성 전이유전자 구조체, 및 (b) 항시성 프로모터를 코딩하는 서열 및 외인성 수용체를 코딩하는 서열, 예컨대 CAR을 포함하는 수용체 구조체를 포함하는 조성물을 포함하고, (a)의 구조체 및 (b)의 구조체를 세포의 게놈 서열 내로 통합시에 외인성 수용체는 발현되며, 외인성 수용체는 리간드 또는 항원의 결합시에 유도성 전이유전자 (a)의 발현을 조절하는 유도성 프로모터를 직접 또는 간접적으로 표적화하는 세포내 신호를 전달하여 유전자 발현을 변형시킨다.

[0173] 조성물은 유전자 발현을 감소시킴으로써 유전자 발현을 변형시킬 수 있다. 조성물은 유전자 발현을 일시적으로 변형시킴으로써(예컨대, 리간드가 외인성 수용체에 결합하는 동안) 유전자 발현을 변형시킬 수 있다. 조성물은 유전자 발현을 급성으로 변형시킬 수 있다(예컨대, 리간드가 외인성 수용체에 가역적으로 결합함). 조성물은 유전자 발현을 만성적으로 변형시킬 수 있다(예컨대, 리간드가 외인성 수용체에 비가역적으로 결합함).

[0174] 외인성 수용체는 세포의 게놈 서열과 관련하여 내인성 수용체를 포함할 수 있다. 예시적인 수용체는, 비제한적으로, 세포내 수용체, 세포 표면 수용체, 막관통 수용체, 리간드 게이트된 이온 채널, 및 G-단백질 결합된 수용

체를 포함한다.

[0175] 외인성 수용체는 비자연발생 수용체를 포함할 수 있다. 비자연발생 수용체는 합성, 변형된, 재조합, 돌연변이체 또는 키메라 수용체일 수 있다. 비자연발생 수용체는 T-세포 수용체(TCR)로부터 단리되거나 유래된 하나 이상의 서열을 포함할 수 있다. 비자연발생 수용체는 스캐폴드 단백질로부터 단리되거나 유래된 하나 이상의 서열을 포함할 수 있다. 비자연발생 수용체가 막관통 도메인을 포함하지 않는 것을 포함하는 일부 양태에서, 비자연발생 수용체는 비자연발생 수용체와 접촉 후 세포내 신호를 전달하는 제2 막관통, 막 결합된 및/또는 세포내 수용체와 상호작용한다. 비자연발생 수용체는 막관통 도메인을 포함할 수 있다. 비자연발생 수용체는 세포내 신호를 전달하는 세포내 수용체와 상호작용할 수 있다. 비자연발생 수용체는 세포내 신호전달 도메인을 포함할 수 있다. 비자연발생 수용체는 키메라 리간드 수용체(CLR)일 수 있다. CLR은 키메라 항원 수용체(CAR)일 수 있다.

[0176] 유도성 프로모터를 코딩하는 서열은 NFkB 프로모터를 코딩하는 서열, 인터페론(IFN) 프로모터를 코딩하는 서열 또는 인터루킨-2 프로모터를 코딩하는 서열을 포함한다. 일부 양태에서, IFN 프로모터는 IFNɣ 프로모터이다. 유도성 프로모터는 사이토카인 또는 케모카인의 프로모터로부터 단리되거나 유래될 수 있다. 사이토카인 또는 케모카인은 IL2, IL3, IL4, IL5, IL6, IL10, IL12, IL13, IL17A/F, IL21, IL22, IL23, 형질전환 성장 인자 베타(TGFβ), 콜로니 자극 인자 2(GM-CSF), 인터페론 감마(IFNɣ), 종양 괴사 인자 알파(TNFα), LTα, 퍼포린(perforin), 그랜자임 C(Granzyme C, Gzmc), 그랜자임 B(Gzmb), C-C 모티프 케모카인 리간드 5(CCL5), C-C 모티프 케모카인 리간드 4(Ccl4), C-C 모티프 케모카인 리간드 3(Ccl3), X-C 모티프 케모카인 리간드 1(Xcl1) 또는 LIF 인터루킨 6 패밀리 사이토카인(Lif)을 포함할 수 있다.

[0177] 유도성 프로모터는 세포 분화, 활성화, 고갈 및 기능에 관여하는 표면 단백질을 포함하는 유전자의 프로모터로부터 단리되거나 유래될 수 있다. 일부 양태에서, 유전자는 CD69, CD71, CTLA4, PD-1, TIGIT, LAG3, TIM-3, GITR, MHCII, COX-2, FASL 또는 4-1BB를 포함한다.

[0178] 유도성 프로모터는 CD 대사 및 분화에 관여하는 유전자의 프로모터로부터 단리되거나 유래될 수 있다. 유도성 프로모터는 Nr4a1, Nr4a3, Tnfrsf9(4-1BB), Sema7a, Zfp3612, Gadd45b, Dusp5, Dusp6 및 Neto2의 프로모터로부터 단리되거나 유래될 수 있다.

[0179] 일부 양태에서, 유도성 전이유전자 구조체는 억제성 체크포인트 신호, 전사 인자, 사이토카인 또는 사이토카인 수용체, 케모카인 또는 케모카인 수용체, 세포 사멸 또는 아폽토시스 수용체/리간드, 대사 감지 분자, 암 요법에 민감성을 부여하는 단백질, 및 종양유전자 또는 종양 억제 유전자의 하류에 신호전달 구성요소를 포함하거나 이의 발현을 유도한다. 이의 비제한적인 예가 PCT 공개 WO 제2019/173636호 및 PCT 출원 제PCT/US2019/049816호에 개시되어 있다.

[0180] **무장화된(Armored) 세포**

[0181] 본 개시내용의 변형된 세포(예컨대, CAR T-세포)는 그들의 치료 잠재력을 향상시키기 위해 추가로 변형될 수 있다. 대안적으로, 또는 추가로, 변형된 세포는 이들을 면역학적 및/또는 대사 체크포인트에 덜 민감하게 만들도록 추가로 변형될 수 있다. 이러한 유형의 변형은 세포를 "무장화"시키고, 이는 변형 후에 본원에서 "무장화된" 세포(예컨대, 무장화된 T-세포)로 지칭될 수 있다. 무장화된 세포는, 예를 들어 종양 면역억제 미세환경 내에서 자연적으로 세포에 전달되는 특정 체크포인트 신호를 차단 및/또는 희석함으로써(예컨대, 체크포인트 억제) 생산될 수 있다.

[0182] 본 개시내용의 무장화된 세포는 임의의 세포, 예를 들어, T 세포, NK 세포, 조혈 선조 세포, 말초 혈액(PB) 유래 T 세포(G-CSF-동원된 말초 혈액으로부터 단리되거나 유래된 T 세포 포함), 또는 제대혈(UCB) 유래 T 세포로부터 유래될 수 있다. 무장화된 세포(예컨대, 무장화된 T-세포)는 키메라 리간드 수용체(단백질 스캐폴드, 항체, ScFv, 또는 항체 모방체를 포함하는 CLR)/키메라 항원 수용체(단백질 스캐폴드, 항체, ScFv, 또는 항체 모방체를 포함하는 CAR), CARTyrin(센티린을 포함하는 CAR), 및/또는 VCAR(낙타과 VHH 또는 단일 도메인 VH를 포함하는 CAR) 중 하나 이상을 포함할 수 있다. 무장화된 세포(예컨대, 무장화된 T-세포)는 본원에 개시된 바와 같은 유도성 아폽토시스 촉진 폴리펩타이드를 포함할 수 있다. 무장화된 세포(예컨대, 무장화된 T-세포)는 외인성 서열을 포함할 수 있다. 외인성 서열은 치료적 단백질을 코딩하는 서열을 포함할 수 있다. 예시적인 치료적 단백질은 핵, 세포질, 세포내, 막관통, 세포 표면 결합된, 또는 분비된 단백질일 수 있다. 무장화된 세포(예컨대, 무장화된 T-세포)에 의해 발현된 예시적인 치료적 단백질은 무장화된 세포의 활성을 변형시키거나 제2 세포의 활성을 변형시킬 수 있다. 무장화된 세포(예컨대, 무장화된 T-세포)는 선택 유전자 또는 선택 마커를 포함할 수 있다. 무장화된 세포(예컨대, 무장화된 T-세포)는 합성 유전자 발현 카세트(본원에서 유도성 전이유전자 구

조체로도 지칭됨)를 포함할 수 있다.

[0183] 본 개시내용의 변형된 세포(예컨대, CAR T-세포)는 무장화된 세포(예컨대, 무장화된 CAR T-세포)를 생산하기 위해 억제성 체크포인트 신호의 수용체(들)를 코딩하는 하나 이상의 유전자(들)의 발현을 침묵시키거나 감소시키도록 추가로 변형될 수 있다. 억제성 체크포인트 신호의 수용체는 세포 표면 상에 또는 세포의 세포질 내에서 발현된다. 억제성 체크포인트 신호의 수용체를 코딩하는 유전자의 발현을 침묵시키거나 감소시키는 것은 무장화된 세포의 표면 상에 또는 세포질 내에 억제성 체크포인트 수용체의 단백질 발현의 손실을 초래한다. 따라서, 억제성 체크포인트 수용체를 코딩하는 하나 이상의 유전자의 발현이 침묵되거나 감소된 무장화된 세포는 체크포인트 신호에 내성, 비수용성 또는 둔감하다. 억제성 체크포인트 신호에 대한 무장화된 세포의 내성 또는 감소된 민감성은 이러한 억제성 체크포인트 신호의 존재하에 무장화된 세포의 치료 잠재력을 향상시킨다. 억제성 체크포인트 신호(및 면역억제를 유도하는 단백질)의 비제한적인 예는 PCT 공개 WO 제2019/173636호에 개시되어 있다. 침묵될 수 있는 억제성 체크포인트 신호의 바람직한 예는, 비제한적으로, PD-1 및 TGFβRII를 포함한다.

[0184] 본 개시내용의 변형된 세포(예컨대, CAR T-세포)는 무장화된 세포(예컨대, 무장화된 CAR T-세포)를 생산하기 위해 체크포인트 신호전달에 관여하는 세포내 단백질을 코딩하는 하나 이상의 유전자(들)의 발현을 침묵시키거나 감소시키도록 추가로 변형될 수 있다. 변형된 세포의 활성은 체크포인트 신호전달 경로에 관여하는 임의의 세포내 신호전달 단백질을 표적화함으로써 향상될 수 있으며, 이에 의해 하나 이상의 체크포인트 경로에 대한 체크포인트 억제 또는 간섭을 달성할 수 있다. 체크포인트 신호전달에 관여하는 세포내 신호전달 단백질의 비제한적인 예는 PCT 공개 WO 제2019/173636호에 개시되어 있다.

[0185] 본 개시내용의 변형된 세포(예컨대, CAR T-세포)는 무장화된 세포(예컨대, 무장화된 CAR T-세포)를 생산하기 위해 요법의 효능을 방해하는 전사 인자를 코딩하는 하나 이상의 유전자(들)의 발현을 침묵시키거나 감소시키도록 추가로 변형될 수 있다. 변형된 세포의 활성은 요법의 효능을 방해하는 전사 인자의 발현을 침묵시키거나 감소시킴으로써(또는 기능을 억제함으로써) 향상되거나 조절될 수 있다. 발현을 침묵시키거나 감소시키기 위해 또는 이의 기능을 억제하기 위해 변형될 수 있는 전사 인자의 비제한적인 예는, 비제한적으로, PCT 공개 WO 제2019/173636호에 개시된 예시적인 전사 인자를 포함한다.

[0186] 본 개시내용의 변형된 세포(예컨대, CAR T-세포)는 무장화된 세포(예컨대, 무장화된 CAR T-세포)를 생산하기 위해 세포 사멸 또는 세포 아폽토시스 수용체를 코딩하는 하나 이상의 유전자(들)의 발현을 침묵시키거나 감소시키도록 추가로 변형될 수 있다. 사멸 수용체 및 그의 내인성 리간드의 상호작용은 아폽토시스의 개시를 초래한다. 세포 사멸 및/또는 세포 아폽토시스 수용체 및/또는 리간드의 발현, 활성, 또는 상호작용의 방해는 변형된 세포를 사멸 신호에 대해 덜 수용성이게 만들어, 결과적으로 무장화된 세포를 종양 환경에 더 효과적이게 만든다. 세포 사멸 및/또는 세포 아폽토시스 수용체 및 리간드의 비제한적인 예는 PCT 공개 WO 제2019/173636호에 개시되어 있다. 변형될 수 있는 세포 사멸 수용체의 바람직한 예는 Fas(CD95)이다.

[0187] 본 개시내용의 변형된 세포(예컨대, CAR T-세포)는 무장화된 세포(예컨대, 무장화된 CAR T-세포)를 생산하기 위해 대사 감지 단백질을 코딩하는 하나 이상의 유전자(들)의 발현을 침묵시키거나 감소시키도록 추가로 변형될 수 있다. 변형된 세포에 의한 면역억제 종양 미세환경(산소, pH, 글루코스 및 다른 분자의 낮은 수준을 특징으로 함)의 대사 감지의 방해는 T-세포 기능의 연장된 보유를 초래하여, 결과적으로 세포당 더 많은 종양 세포가 사멸되게 한다. 대사 감지 유전자 및 단백질의 비제한적인 예는 PCT 공개 WO 제2019/173636호에 개시되어 있다. 바람직한 예인 HIF1a 및 VHL은 저산소 환경에 있는 동안 T-세포 기능에서 역할을 한다. 무장화된 T-세포는 HIF1a 또는 VHL을 코딩하는 하나 이상의 유전자의 발현이 침묵되거나 감소될 수 있다.

[0188] 본 개시내용의 변형된 세포(예컨대, CAR T-세포)는 무장화된 세포(예컨대, 무장화된 CAR T-세포)를 생산하기 위해, 단클론 항체를 포함하는 암 요법에 대한 민감성을 부여하는 단백질을 코딩하는 하나 이상의 유전자(들)의 발현을 침묵시키거나 감소시키도록 추가로 변형될 수 있다. 따라서, 무장화된 세포는 암 요법(예컨대, 화학요법, 단클론 항체 요법, 또는 또 다른 항종양 치료)의 존재 하에 있는 동안 기능할 수 있고 우수한 기능 또는 효능을 입증할 수 있다. 암 요법에 대한 민감성을 부여하는 데 관여하는 단백질의 비제한적인 예가 PCT 공개 WO 제2019/173636호에 개시되어 있다.

[0189] 본 개시내용의 변형된 세포(예컨대, CAR T-세포)는 무장화된 세포(예컨대, 무장화된 CAR T-세포)를 생산하기 위해 성장 이점 인자를 코딩하는 하나 이상의 유전자(들)의 발현을 침묵시키거나 감소시키도록 추가로 변형될 수 있다. 종양유전자의 발현을 침묵시키거나 감소시키는 것은 세포에 성장 이점을 부여할 수 있다. 예를 들어, CAR T-세포 제조 공정 동안 TET2 유전자의 발현을 침묵시키거나 감소시키는 것(예컨대, 발현을 방해하는 것)은 이러한 확장 능력이 부족한 비무장화된 CAR T-세포와 비교하여 유의한 확장 능력 및 후속적인 종양 제거를 갖는 무

장화된 CAR T-세포의 생성을 초래한다. 이 전략은 안전 스위치(예컨대, 본원에 기재된 iC9 안전 스위치)에 결합될 수 있으며, 이는 대상체로부터의 역반응 또는 무장화된 CAR T-세포의 통제되지 않은 성장의 사건에서 무장화된 CAR T-세포의 표적화된 방해를 허용한다. 성장 이점 인자의 비제한적인 예는 PCT 공개 WO 제2019/173636호에 개시되어 있다.

[0190] 본 개시내용의 변형된 세포(예컨대, CAR T-세포)는 본 개시내용의 무장화된 T-세포를 생산하기 위해 변형된/키메라 체크포인트 수용체를 발현하도록 추가로 변형될 수 있다.

[0191] 변형된/키메라 체크포인트 수용체는 null 수용체, 유인(decoy) 수용체 또는 우성 음성 수용체를 포함할 수 있다. null 수용체, 유인 수용체 또는 우성 음성 수용체는 변형된/키메라 수용체/단백질일 수 있다. null 수용체, 유인 수용체 또는 우성 음성 수용체는 세포내 신호전달 도메인의 발현을 위해 절단될 수 있다. 대안적으로, 또는 추가로, null 수용체, 유인 수용체 또는 우성 음성 수용체는 효과적인 신호전달에 결정적이거나 필요한 하나 이상의 아미노산 위치에서 세포내 신호전달 도메인 내에서 돌연변이될 수 있다. null 수용체, 유인 수용체 또는 우성 음성 수용체의 절단 또는 돌연변이는 체크포인트 신호를 세포 또는 세포 내로 운반 또는 전달하는 수용체의 능력을 상실시킬 수 있다.

[0192] 예를 들어, 종양 세포의 표면 상에 발현된 PD-L1 수용체로부터의 면역억제 체크포인트 신호의 희석 또는 차단은 무장화된 세포(예컨대, 무장화된 CAR T-세포)의 표면 상에 변형된/키메라 PD-1 null 수용체를 발현시킴으로써 달성될 수 있으며, 이는 무장화된 세포의 내인성 PD-1 수용체를 통해 면역억제 체크포인트 신호의 전달을 감소시키거나 억제하기 위해 또한 무장화된 세포의 표면 상에 발현된 내인성(비변형된) PD-1 수용체과 효과적으로 경쟁한다. 이러한 비제한적인 예에서, 종양 세포 상에서 발현되는 PD-L1에의 결합에 대해 2개의 상이한 수용체 사이의 경쟁은 효과적인 체크포인트 신호전달의 수준을 감소시키거나 약화시켜 PD-1 null 수용체를 발현하는 무장화된 세포의 치료 잠재력을 향상시킨다.

[0193] 변형된/키메라 체크포인트 수용체는 막관통 수용체, 막 연관된 또는 막 연결된 수용체/단백질 또는 세포내 수용체/단백질인 null 수용체, 유인 수용체 또는 우성 음성 수용체를 포함할 수 있다. 예시적인 null, 유인, 또는 우성 음성 세포내 수용체/단백질은, 비제한적으로, 억제성 체크포인트 신호, 전사 인자, 사이토카인 또는 사이토카인 수용체, 케모카인 또는 케모카인 수용체, 세포 사멸 또는 아폽토시스 수용체/리간드, 대사 감지 분자, 암 요법에 대해 민감성을 부여하는 단백질, 및 종양유전자 또는 종양 억제 유전자의 하류에 신호전달 구성요소를 포함한다. 사이토카인, 사이토카인 수용체, 케모카인 및 케모카인 수용체의 비제한적인 예는 PCT 공개 WO 제2019/173636호에 개시되어 있다.

[0194] 변형된/키메라 체크포인트 수용체는 스위치 수용체를 포함할 수 있다. 예시적인 스위치 수용체는 변형된/키메라 수용체/단백질을 포함하며, 여기서 천연 또는 야생형 세포내 신호전달 도메인은 단백질에 대해 비천연이고/거나 야생형 도메인이 아닌 상이한 세포내 신호전달 도메인로 전환되거나 대체된다. 예를 들어, 억제성 신호전달 도메인을 자극 신호전달 도메인으로 대체하는 것은 면역억제 신호를 면역자극 신호로 전환할 것이다. 대안적으로, 억제성 신호전달 도메인을 상이한 억제성 도메인으로 대체하는 것은 억제성 신호전달의 수준을 감소시키거나 향상시킬 수 있다. 스위치 수용체의 발현 또는 과발현은 면역억제 종양 미세환경 내에서 발현된 동족 체크포인트 수용체에의 결합에 대해 내인성 야생형 체크포인트 수용체(스위치 수용체가 아님)와의 경쟁을 통해 동족 체크포인트 신호의 희석 및/또는 차단을 초래할 수 있다. 무장화된 세포(예컨대, 무장화된 CAR T-세포)는 하나 이상의 스위치 수용체의 발현을 초래하고, 결과적으로 무장화된 세포의 활성을 변경시키는 스위치 수용체를 코딩하는 서열을 포함할 수 있다. 무장화된 세포(예컨대, 무장화된 CAR T-세포)는 체크포인트 수용체, 전사 인자, 사이토카인 수용체, 사멸 수용체, 대사 감지 분자, 암 요법, 종양유전자, 및/또는 종양 억제 단백질 또는 유전자의 하류에 있는 세포내에서 발현된 단백질을 표적화하는 스위치 수용체를 발현할 수 있다.

[0195] 예시적인 스위치 수용체는, 비제한적으로, 억제성 체크포인트 신호, 전사 인자, 사이토카인 또는 사이토카인 수용체, 케모카인 또는 케모카인 수용체, 세포 사멸 또는 아폽토시스 수용체/리간드, 대사 감지 분자, 암 요법에 대해 민감성을 부여하는 단백질, 및 종양유전자 또는 종양 억제 유전자의 하류에 있는 신호전달 구성요소를 포함하는, 단백질을 포함할 수 있거나 또는 이로부터 유래될 수 있다.

[0196] 본 개시내용의 변형된 세포(예컨대, CAR T-세포)는 무장화된 T-세포를 생산하기 위해 조건부 유전자 발현을 매개하는 CLR/CAR을 발현하도록 추가로 변형될 수 있다. 무장화된 T-세포의 핵에서 CLR/CAR 및 조건부 유전자 발현 시스템의 조합은 동족 리간드(들)와 CLR의 조합 또는 동족 항원(들)과 CAR의 조합시 조건부로 활성화되는 합성 유전자 발현 시스템을 구성한다. 이 시스템은 예를 들어 종양 환경에서 또는 내부에서 리간드 또는 항원 결합의 부위에서 합성 유전자 발현을 감소시키거나 제한함으로써 변형된 T-세포의 치료 잠재력을 '무장화' 또는

향상시키는 데 도움을 줄 수 있다.

[0197] **유전자 편집 조성물 및 방법**

[0198] 변형된 세포는 전이유전자를 세포 내로 도입함으로써 생산된다. 도입 단계는 비전위 전달 시스템을 통한 핵산 서열, 전이유전자, 및/또는 게놈 편집 구조체의 전달을 포함할 수 있다.

[0199] 핵산 서열, 전이유전자 및/또는 게놈 편집 구조체를 생체외에서, 생체내에서, 시험관내에서 또는 원위치에서 세포 내로 도입하는 것은 국소 전달, 흡착, 흡수, 전기천공, 회전감염, 공동배양, 형질감염, 기계적 전달, 음파 전달, 진동 전달, 자기감염 또는 나노입자 매개 전달 중 하나 이상을 포함할 수 있다. 핵산 서열, 전이유전자 및/또는 게놈 편집 구조체를 생체외에서, 생체내에서, 시험관내에서 또는 원위치에서 세포 내로 도입하는 것은 리포솜 형질감염, 인산 칼슘 형질감염, 퓨진(fugene) 형질감염, 및 덴드리머(dendrimer) 매개 형질감염을 포함할 수 있다. 기계적 형질감염에 의해 핵산 서열, 전이유전자, 및/또는 게놈 편집 구조체를 생체외에서, 생체내에서, 시험관내에서 또는 원위치에서 세포 내로 도입하는 것은 세포 압착(squeezing), 세포 충격(bombardment), 또는 유전자 총 기술을 포함할 수 있다. 나노입자 매개 형질감염에 의해 핵산 서열, 전이유전자 및/또는 게놈 편집 구조체를 세포 내로 도입하는 것은 리포솜 전달, 마이셀에 의한 전달, 및 폴리머좀에 의한 전달을 포함할 수 있다.

[0200] 핵산 서열, 전이유전자 및/또는 게놈 편집 구조체를 생체외에서, 생체내에서, 시험관내에서 또는 원위치에서 세포 내로 도입하는 것은 비바이러스 벡터를 포함할 수 있다. 비바이러스 벡터는 핵산을 포함할 수 있다. 비바이러스 벡터는 플라스미드 DNA, 선형 이중 가닥의 DNA(dsDNA), 선형 단일 가닥 DNA(ssDNA), DoggyBone™ DNA, 나노플라스미드, 미니서클 DNA, 단일 가닥 올리고데옥시뉴클레오타이드(ssODN), DDNA 올리고뉴클레오타이드, 단일 가닥 mRNA(ssRNA), 및 이중 가닥 mRNA(dsRNA)를 포함할 수 있다. 비바이러스 벡터는 본원에 기재된 바와 같은 트랜스포존을 포함할 수 있다.

[0201] 핵산 서열, 전이유전자 및/또는 게놈 편집 구조체를 생체외에서, 생체내에서, 시험관내에서 또는 원위치에서 세포 내로 도입하는 것은 바이러스 벡터를 포함할 수 있다. 바이러스 벡터는 비통합 비염색체 벡터일 수 있다. 비통합 비염색체 벡터의 비제한적인 예는 아데노 관련 바이러스(AAV), 아데노바이러스, 및 헤르페스 바이러스를 포함한다. 바이러스 벡터는 통합 염색체 벡터일 수 있다. 통합 염색체 벡터의 비제한적인 예는 아데노 관련 벡터(AAV), 렌티바이러스, 및 감마-레트로바이러스를 포함한다.

[0202] 핵산 서열, 전이유전자 및/또는 게놈 편집 구조체를 생체외에서, 생체내에서, 시험관내에서 또는 원위치에서 세포 내로 도입하는 것은 벡터의 조합을 포함할 수 있다. 벡터 조합의 비제한적인 예는 바이러스 및 비바이러스 벡터, 복수의 비바이러스 벡터, 또는 복수의 바이러스 벡터를 포함한다. 벡터 조합의 비제한적인 예는 DNA 유래된 벡터 및 RNA 유래된 벡터의 조합, RNA 및 역전사효소의 조합, 트랜스포존 및 트랜스포사제의 조합, 비바이러스 벡터 및 엔도뉴클레아제의 조합, 및 바이러스 벡터 및 엔도뉴클레아제의 조합을 포함한다.

[0203] 게놈 변형은 핵산 서열을 안정적으로 통합하거나, 핵산 서열을 일시적으로 통합하거나, 핵산 서열의 부위 특이적 통합을 생산하거나, 또는 핵산 서열의 편향된 통합을 생산하기 위해 핵산 서열, 전이유전자 및/또는 게놈 편집 구조체를 생체외에서, 생체내에서, 시험관내에서 또는 원위치에서 세포 내로 도입하는 것을 포함할 수 있다. 핵산 서열은 전이유전자일 수 있다.

[0204] 게놈 변형은 핵산 서열을 안정적으로 통합하기 위해 핵산 서열, 전이유전자 및/또는 게놈 편집 구조체를 생체외에서, 생체내에서, 시험관내에서 또는 원위치에서 세포 내로 도입하는 것을 포함할 수 있다. 안정적인 염색체 통합은 무작위 통합, 부위 특이적 통합, 또는 편향된 통합일 수 있다. 부위 특이적 통합은 지원받지 않거나 지원받을 수 있다. 지원된 부위 특이적 통합은 부위-지정 뉴클레아제와 함께 전달된다. 부위-지정 뉴클레아제는 게놈 통합의 부위의 상류 및 하류 영역에 대해 백분율 상동성을 함유하는 5' 및 3' 뉴클레오타이드 서열 연장을 갖는 전이유전자를 포함한다. 상동성 뉴클레오타이드 연장을 갖는 전이유전자는 상동성 재조합, 미세상동성 매개 말단 결합, 또는 비상동성 말단 결합에 의한 게놈 통합을 가능하게 한다. 부위 특이적 통합은 안전한 하버(harbor) 부위에서 발생할 수 있다. 게놈 안전한 하버 부위는 새로 삽입된 유전 요소가 확실히 기능하고(예를 들어, 발현의 치료적으로 효과적인 수준에서 발현됨) 숙주 유기체에 위험을 유발하는 숙주 게놈에 해로운 변경을 유발하지 않도록 하는 방식으로 새로운 유전 물질의 통합을 수용할 수 있다. 잠재적인 게놈 안전한 하버의 비제한적인 예는 인간 알부민 유전자의 인트론 서열, 아데노 관련 바이러스 부위 1(AAVS1), 염색체 19 상의 AAV 바이러스의 통합의 자연발생 부위, 케모카인(C-C 모티프) 수용체 5(CCR5) 유전자의 부위 및 마우스 Rosa26 유전자좌의 인간 오르토로그의 부위를 포함한다.

[0205] 부위 특이적 전이유전자 통합은 표적 유전자의 발현을 방해하는 부위에서 발생할 수 있다. 표적 유전자 발현의 방해는 인트론, 엑손, 프로모터, 유전 요소, 인핸서, 억제자, 시작 코돈, 정지 코돈, 및 반응 요소에서의 부위 특이적 통합에 의해 발생할 수 있다. 부위 특이적 통합에 의해 표적화되는 표적 유전자의 비제한적인 예는 TRAC, TRAB, PDI, 임의의 면역억제 유전자, 및 동종거부에 관여하는 유전자를 포함한다.

[0206] 부위 특이적 전이유전자 통합은 표적 유전자의 발현을 향상시키는 부위에서 발생할 수 있다. 표적 유전자 발현의 향상은 인트론, 엑손, 프로모터, 유전 요소, 인핸서, 억제자, 시작 코돈, 정지 코돈, 및 반응 요소에서의 부위 특이적 통합에 의해 발생할 수 있다.

[0207] 효소는 전이유전자의 전달 또는 통합을 촉진하기 위해 숙주 게놈에 가닥 파손(strand break)을 생성하는 데 사용될 수 있다. 효소는 단일 가닥 파손 또는 이중 가닥 파손를 생성할 수 있다. 파손 유도 효소의 비제한적인 예는 트랜스포사제, 인테그라아제, 엔도뉴클레아제, CRISPR-Cas9, 전사 활성자-유사 효과기 뉴클레아제(TALEN), 아연 핑거 뉴클레아제(ZFN), Cas-CLOVER™, 및 CPF1을 포함한다. 파손-유도 효소는 DNA로 코딩되어, mRNA로 코딩되어, 단백질로서, 또는 가이드 RNA(gRNA)와의 핵단백질 복합체로서 세포에 전달될 수 있다.

[0208] 부위 특이적 전이유전자 통합은 벡터 매개 통합 부위 편향에 의해 제어될 수 있다. 벡터 매개 통합 부위 편향은 선택된 렌티바이러스 벡터에 의해 또는 선택된 감마-레트로바이러스 벡터에 의해 제어될 수 있다.

[0209] 부위 특이적 전이유전자 통합 부위는 비안정적인 염색체 삽입일 수 있다. 통합된 전이유전자는 침묵, 제거, 절제, 또는 추가로 변형될 수 있다. 게놈 변형은 전이유전자의 비안정적인 통합일 수 있다. 비안정적인 통합은 일시적인 비염색체 통합, 반안정적인 비염색체 통합, 반지속적인 비염색체 삽입, 또는 비안정적인 염색체 삽입일 수 있다. 일시적인 비염색체 삽입은 에피-염색체 또는 세포질일 수 있다. 일 양태에서, 전이유전자의 일시적인 비염색체 삽입은 염색체 내로 통합되지 않으며, 변형된 유전 물질은 세포 분열 동안 복제되지 않는다.

[0210] 게놈 변형은 전이유전자의 반안정적인 또는 지속적인 비염색체 통합일 수 있다. DNA 벡터는 분열하는 세포의 핵에서 자율 복제를 허용하는 비바이러스 벡터의 에피솜 보유를 위해 핵 매트릭스 단백질에 결합하는 스캐폴드/매트릭스 부착 영역(S-MAR) 모듈을 코딩한다.

[0211] 게놈 변형은 전이유전자의 비안정적인 염색체 통합일 수 있다. 통합된 전이유전자는 침묵, 제거, 절제, 또는 추가로 변형될 수 있다.

[0212] 전이유전자 삽입에 의한 게놈 변형은 상동성 재조합(HR), 미세상동성 매개 말단 결합(MMEJ), 비상동성 말단 결합(NHEJ), 트랜스포사제 효소 매개 변형, 인테그라아제 효소 매개 변형, 엔도뉴클레아제 효소 매개 변형, 또는 재조합 효소 매개 변형에 의한 숙주 세포-지시 이중 가닥 파손 복구(상동 지시 복구)를 통해 발생할 수 있다. 전이유전자 삽입에 의한 게놈 변형은 CRISPR-Cas9, TALEN, ZFNs, Cas-CLOVER™, 및 cpf1을 통해 발생할 수 있다.

[0213] 새로운 또는 기존 뉴클레오타이드/핵산을 삽입하는 것을 포함하는 유전자 편집 시스템에서, 삽입 도구(예컨대, DNA 주형 벡터, 전위가능한 요소(트랜스포존 또는 레트로트랜스포존)는 절단 효소(예컨대, 뉴클레아제, 재조합효소, 인테그라아제 또는 트랜스포사제)와 함께 세포에 전달되어야 한다. 재조합효소를 위한 이러한 삽입 도구의 예는 DNA 벡터를 포함할 수 있다. 다른 유전자 편집 시스템은 삽입 벡터와 함께 인테그라아제의 전달, 트랜스포존/레트로트랜스포존과 함께 트랜스포사제의 전달 등을 필요로 한다. 절단 효소로서 사용될 수 있는 예시적인 재조합효소는 CRE 재조합효소이다. 삽입 도구에서 사용될 수 있는 인테그라아제의 비제한적인 예는 AAV, 감마 레트로바이러스, 및 렌티바이러스를 포함하는 많은 바이러스로부터 취한 바이러스 기반 효소를 포함한다. 삽입 도구에서 사용될 수 있는 예시적인 트랜스포존/레트로트랜스포존은 본원에 더 상세히 기재되어 있다.

[0214] 생체외, 생체내, 시험관내 또는 원위치 게놈 변형을 갖는 세포는 생식계열 세포 또는 체세포일 수 있다. 변형된 세포는 인간, 비인간, 포유동물, 랫트, 마우스, 또는 개 세포일 수 있다. 변형된 세포는 분화되거나, 미분화되거나, 또는 불멸화될 수 있다. 변형된 미분화된 세포는 줄기 세포일 수 있다. 변형된 미분화된 세포는 유도된 다능성 줄기 세포일 수 있다. 변형된 세포는 면역 세포일 수 있다. 변형된 세포는 T 세포, 조혈 줄기 세포, 자연 살해 세포, 대식세포, 수지상 세포, 단핵구, 거핵구, 또는 파골세포일 수 있다. 변형된 세포는 세포가 정지 상태이거나, 활성화된 상태이거나, 휴지상태이거나, 간기(interphase)이거나, 전기(prophase)이거나, 중기(metaphase)이거나, 후기(anaphase)이거나, 또는 말기(telophase)인 동안 변형될 수 있다. 변형된 세포는 신선하거나, 동결건조되거나, 부피가 크거나, 하위집단으로 분류되거나, 전혈로부터 유래되거나, 백혈구채집술로부터 유래되거나, 또는 불멸화된 세포주로부터 유래될 수 있다. 백혈구채집술 생성물 또는 혈액으로부터 세포를 단리하기 위한 상세한 설명은 PCT 공개 WO 제2019/173636호 및 제PCT/US2019/049816호에 개시되어 있다.

[0215] 본 개시내용은 유전자 편집 조성물 및/또는 유전자 편집 조성물을 포함하는 세포를 제공한다. 유전자 편집 조성물은 DNA 결합 도메인을 코딩하는 서열 및 뉴클레아제 단백질 또는 이의 뉴클레아제 도메인을 코딩하는 서열을 포함할 수 있다. 뉴클레아제 단백질을 코딩하는 서열 또는 이의 뉴클레아제 도메인을 코딩하는 서열은 DNA 서열, RNA 서열, 또는 이의 조합을 포함할 수 있다. 뉴클레아제 또는 이의 뉴클레아제 도메인은 CRISPR/Cas 단백질, 전사 활성자-유사 효과기 뉴클레아제(TALEN), 아연 핑거 뉴클레아제(ZFN), 및 엔도뉴클레아제 중 하나 이상을 포함할 수 있다.

[0216] 뉴클레아제 또는 이의 뉴클레아제 도메인은 뉴클레아제 불활성화된 Cas(dCas) 단백질 및 엔도뉴클레아제를 포함할 수 있다. 엔도뉴클레아제는 Clo051 뉴클레아제 또는 이의 뉴클레아제 도메인을 포함할 수 있다. 유전자 편집 조성물은 융합 단백질을 포함할 수 있다. 융합 단백질은 뉴클레아제 불활성화된 Cas9(dCas9) 단백질 및 Clo051 뉴클레아제 또는 Clo051 뉴클레아제 도메인을 포함할 수 있다. 유전자 편집 조성물은 가이드 서열을 추가로 포함할 수 있다. 가이드 서열은 RNA 서열을 포함한다.

[0217] 본 개시내용은 효과기에 작동가능하게 연결된 작은 Cas9(Cas9)를 포함하는 조성물을 제공한다. 본 개시내용은 DNA 국소화 구성요소 및 효과기 분자를 포함하거나, 본질적으로 구성되거나 또는 구성되는 융합 단백질을 제공하며, 효과기는 작은 Cas9(Cas9)를 포함한다. 본 개시내용의 작은 Cas9 구조체는 유형 IIS 엔도뉴클레아제를 포함하는 효과기를 포함할 수 있다. 활성 촉매 부위를 갖는 스타필로코커스 아우레우스 Cas9는 서열번호: 43의 아미노산 서열을 포함한다.

[0218] 본 개시내용은 효과기에 작동가능하게 연결된 불활성화된 작은 Cas9(dSaCas9)를 포함하는 조성물을 제공한다. 본 개시내용은 DNA 국소화 구성요소 및 효과기 분자를 포함하거나, 본질적으로 구성되거나 또는 구성되는 융합 단백질을 제공하며, 효과기는 작은 불활성화된 Cas9(dSaCas9)를 포함한다. 본 개시내용의 작은 불활성화된 Cas9(dSaCas9) 구조체는 유형 IIS 엔도뉴클레아제를 포함하는 효과기를 포함할 수 있다. dSaCas9는 서열번호: 44의 아미노산 서열을 포함하고, 이는 촉매 부위를 불활성화시키는 D10A 및 N580A 돌연변이를 포함한다

[0219] 본 개시내용은 효과기에 작동가능하게 연결된 불활성화된 Cas9(dCas9)를 포함하는 조성물을 제공한다. 본 개시내용은 DNA 국소화 구성요소 및 효과기 분자를 포함하거나, 본질적으로 구성되거나 또는 구성되는 융합 단백질을 제공하며, 효과기는 불활성화된 Cas9(dCas9)를 포함한다. 본 개시내용의 불활성화된 Cas9(dCas9) 구조체는 유형 IIS 엔도뉴클레아제를 포함하는 효과기를 포함할 수 있다.

[0220] dCas9는 스트렙토코커스 피오제네스(*Streptoccocus pyogenes*)로부터 단리되거나 유래될 수 있다. dCas9는 아미노산 위치 10 및 840에 치환을 갖는 dCas9를 포함할 수 있고, 이는 촉매 부위를 불활성화시킨다. 일부 양태에서, 이러한 치환은 D10A 및 H840A이다. dCas9는 서열번호: 45 또는 서열번호: 46의 아미노산 서열을 포함할 수 있다.

[0221] 예시적인 Clo051 뉴클레아제 도메인은 서열번호: 47의 아미노산 서열을 포함하거나, 본질적으로 구성되거나 또는 구성된다.

[0222] 예시적인 dCas9-Clo051(Cas-CLOVER) 융합 단백질은 서열번호: 48의 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있다. 예시적인 dCas9-Clo051 융합 단백질은 서열번호: 49의 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩될 수 있다. dCas9-Clo051 융합 단백질을 코딩하는 핵산은 DNA 또는 RNA일 수 있다.

[0223] 예시적인 dCas9-Clo051(Cas-CLOVER) 융합 단백질은 서열번호: 50의 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있다. 예시적인 dCas9-Clo051 융합 단백질은 서열번호: 51의 핵산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성되는 폴리뉴클레오타이드에 의해 코딩될 수 있다. dCas9-Clo051 융합 단백질을 코딩하는 핵산은 DNA 또는 RNA일 수 있다.

[0224] 유전자 편집 조성물을 포함하는 세포는 유전자 편집 조성물을 안정하게 또는 일시적으로 발현할 수 있다. 바람직하게는, 유전자 편집 조성물은 일시적으로 발현된다. 가이드 RNA는 게놈 DNA 서열 내의 표적 서열에 상보적인 서열을 포함할 수 있다. 게놈 DNA 서열 내의 표적 서열은 게놈 DNA 서열의 안전 하버 부위 내의 표적 서열일 수 있다.

[0225] Cas-CLOVER를 포함하는 유전자 편집 조성물, 및 유전자 편집을 위해 이들 조성물을 사용하는 방법은 미국 특허 공개 제2017/0107541호, 제2017/0114149호, 제2018/0187185호 및 미국 특허 제10,415,024호에 상세히 기재되어 있다.

[0226] 유전자 편집 도구는 또한 하나 이상의 폴리(히스티딘) 기반의 마이셀을 사용하여 세포에 전달될 수 있다. 폴리(히스티딘)(예컨대, 폴리(L-히스티딘))은 불포화된 질소 상에 고립 전자쌍을 제공하는 이미다졸 고리로 인해 pH에 민감한 중합체이다. 즉, 폴리(히스티딘)은 양성자화-탈양성화를 통해 양쪽성 특성을 갖는다. 특히, 특정 pH에서, 폴리(히스티딘) 함유 삼블록 공중합체는 표면 상에 양으로 하전된 폴리(히스티딘) 단위를 갖는 마이셀로 조립되어, 음으로 하전된 유전자 편집 분자(들)와의 복합체화를 가능하게 할 수 있다. pH 의존적 방식으로 단백질 및/또는 핵산을 결합 및 방출하는 이러한 나노입자는 원하는 유전자 변형을 수행하기 위해 효율적이고 선택적인 메커니즘을 제공할 수 있다. 특히, 이 마이셀 기반 전달 시스템은 충전된 물질에 대해 실질적인 유연성뿐만 아니라 큰 탑재 용량, 및 나노입자 탑재물의 표적화된 방출을 제공한다. 일 예에서, 이중 가닥 DNA의 부위 특이적 절단은 폴리(히스티딘) 기반 마이셀을 사용한 뉴클레아제의 전달에 의해 가능하게 된다. 특정 이론에 구속되기를 바라지 않지만, 다양한 삼블록 공중합체에 의해 형성된 마이셀에서, 소수성 블록은 응집하여 코어를 형성하고, 이는 말단에 친수성 블록 및 폴리(히스티딘) 블록을 남겨 하나 이상의 주변 층을 형성한다고 여겨진다.

[0227] 일 양태에서, 본 개시내용은 친수성 블록, 소수성 블록, 및 하전된 블록으로 만들어진 삼블록 공중합체를 제공한다. 일부 양태에서, 친수성 블록은 폴리(에틸렌 옥사이드)(PEO)일 수 있고, 하전된 블록은 폴리(L-히스티딘)일 수 있다. 사용될 수 있는 예시적인 삼블록 공중합체는 PEO-b-PLA-b-PHIS이며, 각 블록 내의 반복 단위의 가변적인 수는 설계에 따라 다양하다.

[0228] 삼블록 공중합체를 제조하기 위한 중간체로 사용될 수 있는 이블록 공중합체는 비제한적으로 폴리(락티드)(PLA), 폴리(글리콜리드)(PLGA), 폴리(락틱-co-글리콜산)(PLGA), 폴리(ε-카프로락톤)(PCL), 및 폴리(트리메틸렌 카르보네이트)(PTMC)를 포함하는 다양한 소수성 지방족 폴리(무수물), 폴리(핵산), 폴리(에스테르), 폴리(오르토 에스테르), 폴리(펩타이드), 폴리(포스파젠) 및 폴리(당류)에 결합된 PEG와 화학적으로 동의어인 친수성 생체적합성 폴리(에틸렌 옥사이드)(PEO)를 가질 수 있다. 100% 페길화된 표면으로 구성된 중합체성 마이셀은 개선된 시험관내 화학적 안정성, 증가된 생체내 생체이용률, 및 연장된 혈액 순환 반감기를 보유한다.

[0229] 삼블록 공중합체를 포함하는 것을 포함하는 중합체성 비히클, 폴리머좀 및 폴리(히스티딘) 기반 마이셀, 및 이를 제조하는 방법은 미국 특허 제7,217,427호; 제7,868,512호; 제6,835,394호; 제8,808,748호; 제10,456,452호; 미국 공개 제2014/0363496호; 제2017/0000743호; 및 제2019/0255191호; 및 PCT 공개 WO 제2019/126589호에 더 상세히 기재되어 있다.

[0230] **유도성 아폽토시스 촉진 폴리펩타이드**

[0231] 본원에 개시된 유도성 아폽토시스 촉진 폴리펩타이드는 기존 유도성 폴리펩타이드보다 우수한데, 본 개시내용의 유도성 아폽토시스 촉진 폴리펩타이드가 면역원성이 훨씬 낮기 때문이다. 유도성 아폽토시스 촉진 폴리펩타이드는 재조합 폴리펩타이드이므로 비자연발생이다. 또한, 숙주 인간 면역 시스템이 "비자기"로 인식하여 결과적으로 유도성 아폽토시스 촉진 폴리펩타이드, 유도성 아폽토시스 촉진 폴리펩타이드를 포함하는 세포 또는 유도성 아폽토시스 촉진 폴리펩타이드를 포함하는 조성물 또는 유도성 아폽토시스 촉진 폴리펩타이드를 포함하는 세포를 받는 대상체에서 면역 반응을 유도하는 비인간 서열을 포함하지 않는 유도성 아폽토시스 촉진 폴리펩타이드를 생성하기 위해 조합된 서열.

[0232] 본 개시내용은 리간드 결합 영역, 링커, 및 아폽토시스 촉진 펩타이드를 포함하는 유도성 아폽토시스 촉진 폴리펩타이드를 제공하며, 유도성 아폽토시스 촉진 폴리펩타이드는 비인간 서열을 포함하지 않는다. 특정 양태에서, 비인간 서열은 제한 부위를 포함한다. 특정 양태에서, 리간드 결합 영역은 다량체성 리간드 결합 영역일 수 있다. 특정 양태에서, 아폽토시스 촉진 펩타이드는 카스파아제 폴리펩타이드이다. 카스파아제 폴리펩타이드의 비제한적인 예는 카스파아제 1, 카스파아제 2, 카스파아제 3, 카스파아제 4, 카스파아제 5, 카스파아제 6, 카스파아제 7, 카스파아제 8, 카스파아제 9, 카스파아제 10, 카스파아제 11, 카스파아제 12, 및 카스파아제 14를 포함한다. 바람직하게는, 카스파아제 폴리펩타이드는 카스파아제 9 폴리펩타이드이다. 카스파아제 9 폴리펩타이드는 절단된 카스파아제 9 폴리펩타이드일 수 있다. 유도성 아폽토시스 촉진 폴리펩타이드는 비자연발생일 수 있다. 카스파아제가 카스파아제 9 또는 절단된 카스파아제 9인 경우, 유도성 아폽토시스 촉진 폴리펩타이드는 또한 "iC9 안전 스위치"로서 지칭될 수 있다.

[0233] 유도성 카스파아제 폴리펩타이드는 (a) 리간드 결합 영역, (b) 링커, 및 (c) 카스파아제 폴리펩타이드를 포함할 수 있고, 유도성 아폽토시스 촉진 폴리펩타이드는 비인간 서열을 포함하지 않는다. 특정 양태에서, 유도성 카스파아제 폴리펩타이드는 (a) 리간드 결합 영역, (b) 링커, 및 (c) 절단된 카스파아제 9 폴리펩타이드를

포함하고, 유도성 아폽토시스 촉진 폴리펩타이드는 비인간 서열을 포함하지 않는다.

[0234] 리간드 결합 영역은 FK506 결합 단백질 12(FKBP12) 폴리펩타이드를 포함할 수 있다. FK506 결합 단백질 12(FKBP12) 폴리펩타이드를 포함하는 리간드 결합 영역의 아미노산 서열은 서열의 위치 36에 변형을 포함할 수 있다. 변형은 위치 36에서 페닐알라닌(F)에 대한 발린(V)의 치환(F36V)일 수 있다. FKBP12 폴리펩타이드는 서열번호: 73과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있다. FKBP12 폴리펩타이드는 서열번호: 74와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나 구성되는 폴리뉴클레오타이드에 의해 코딩될 수 있다.

[0235] 링커 영역은 서열번호: 75와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있거나 또는 링커 영역은 서열번호: 76과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나 구성되는 폴리뉴클레오타이드에 의해 코딩될 수 있다. 일부 양태에서, 링커를 코딩하는 핵산 서열은 제한 부위를 포함하지 않는다.

[0236] 절단된 카스파아제 9 폴리펩타이드는 서열의 위치 87에 아르기닌(R)을 포함하지 않는 아미노산 서열을 포함할 수 있다. 대안적으로, 또는 추가로, 절단된 카스파아제 9 폴리펩타이드는 서열의 위치 282에 알라닌(A)을 포함하지 않는 아미노산 서열을 포함할 수 있다. 절단된 카스파아제 9 폴리펩타이드는 서열번호: 77과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있거나 또는 절단된 카스파아제 9 폴리펩타이드는 서열번호: 78과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나 구성되는 폴리뉴클레오타이드에 의해 코딩될 수 있다.

[0237] 특정 양태에서, 폴리펩타이드가 절단된 카스파아제 9 폴리펩타이드를 포함하는 경우, 유도성 아폽토시스 촉진 폴리펩타이드는 서열번호: 79와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함하거나, 본질적으로 구성되거나, 또는 구성될 수 있거나 또는 유도성 아폽토시스 촉진 폴리펩타이드는 서열번호: 5, 서열번호: 6, 또는 서열번호: 80과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 핵산 서열을 포함하거나 구성되는 폴리뉴클레오타이드에 의해 코딩된다.

[0238] 유도성 아폽토시스 촉진 폴리펩타이드는 세포에서 유도성 아폽토시스 촉진 폴리펩타이드의 발현을 개시 및/또는 조절할 수 있는 당업계에 공지된 임의의 프로모터의 전사 조절하에 세포에서 발현될 수 있다.

[0239] 유도성 아폽토시스 촉진 폴리펩타이드의 활성화는, 예를 들어, 조건부로 제어된 단백질 또는 폴리펩타이드를 생산하기 위해 유도제에 의해 매개되는 화학적으로 유도된 이량체화(CID)를 통해 달성될 수 있다. 아폽토시스 촉진 폴리펩타이드는 유도성일뿐만 아니라 이들 폴리펩타이드의 유도는 불안정한 이량체화제의 분해 또는 단량체성 경쟁적 억제제의 투여로 인해 가역적이다.

[0240] 특정 양태에서, 리간드 결합 영역이 위치 36에서 페닐알라닌(F)에 대한 발린(V)의 치환(F36V)을 갖는 FKBP12 폴리펩타이드를 포함하는 경우, 유도제는 AP1903, 합성 약물(CAS 인덱스 명칭: 2-피페리딘카르복실산, 1-[(2S)-1-옥소-2-(3,4,5-트리메톡시페닐)부틸]-, 1,2-에탄디일비스[이미노(2-옥소-2,1-에탄디일)옥시-3,1-페닐렌[(1R)-3-(3,4-디메톡시페닐)프로필리덴]]에스테르, [2S-[1(R*),2R*[S*[S*[1(R*),2R*]]]]-(9Cl) CAS 등록 번호: 195514-63-7; 분자식: C78H98N4O20; 분자량: 1411.65)); AP20187(CAS 등록 번호: 195514-80-8 및 분자식: C82H107N5O20) 또는 AP20187 유사체, 예를 들어, AP1510을 포함할 수 있다. 본원에 사용된 바와 같이, 유도제 AP20187, AP1903 및 AP1510은 상호교환적으로 사용될 수 있다.

[0241] 유도성 아폽토시스 촉진 펩타이드 및 이들 펩타이드를 유도하는 방법은 미국 특허 공개 WO 제2019/0225667호 및 PCT 공개 WO 제2018/068022호에 상세히 기재되어 있다.

[0242] **키메라 자극제 수용체 및 재조합 HLA-E 폴리펩타이드**

[0243] 임의의 환자에 투여하기에 "보편적으로" 안전한 입양 세포 조성물은 동종이식편반응성(alloreactivity)의 유의한 감소 또는 제거를 필요로 한다. 이를 위해, 본 개시내용의 세포(예컨대, 동종이계 세포)는 T-세포 수용체(TCR) 및/또는 주 조직적합성 복합체(MHC)의 부류의 발현 또는 기능을 방해하도록 변형될 수 있다. TCR은 이식편 대 숙주(GvH) 반응을 매개하는 반면, MHC는 숙주 대 이식편(HvG) 반응을 매개한다. 바람직한 양태에서, TCR의 임의의 발현 및/또는 기능은 대상체에게 사망을 유발할 수 있는 T-세포 매개 GvH를 방지하기 위해 제거된다.

따라서, 바람직한 양태에서, 본 개시내용은 순수한 TCR-음성 동종이계 T-세포 조성물을 제공한다(예컨대, 조성물의 각 세포는 검출불가능하거나 존재하지 않을 정도의 낮은 수준으로 발현된다).

[0244] MHC 부류 I(MHC-I, 특이적으로, HLA-A, HLA-B, 및 HLA-C)의 발현 및/또는 기능은 HvG를 방지하고, 결과적으로 대상체에서 세포의 생착을 개선하기 위해 감소되거나 제거되다. 개선된 생착은 세포의 더 긴 지속성을 초래하고, 따라서, 대상체에게 더 큰 치료적 창을 제공한다. 구체적으로, MHC-I의 구조적 요소인 베타-2-마이크로글로불린(B2M)의 발현 및/또는 기능이 감소되거나 제거된다.

[0245] 상기 전략은 추가 도전을 유도한다. T 세포에서 T 세포 수용체(TCR) 넉아웃(KO)은 TCR 복합체의 일부인 CD3-제타(CD3z 또는 CD3ζ)의 발현 손실을 초래한다. TCR-KO T-세포에서 CD3ζ의 손실은 비제한적으로, 작용제 항-CD3 mAb를 포함하는 표준 자극/활성화 시약을 사용하여 이들 세포를 최적으로 활성화시키고 확장시키는 능력을 극적으로 감소시킨다. TCR 복합체의 임의의 한 구성요소의 발현 또는 기능이 중단되면, TCR-알파(TCRα), TCR-베타(TCRα), CD3-감마(CD3γ), CD3-엡실론(CD3ε), CD3-델타(CD3δ), 및 CD3-제타(CD3ζ)를 포함하는 복합체의 모든 구성요소가 손실된다. CD3ε 및 CD3ζ 둘 모두는 T 세포 활성화 및 확장에 필요하다. 작용제 항-CD3 mAb는 전형적으로 CD3ε 및 아마도 복합체 내의 또 다른 단백질을 인식하며, 이는 차례로 CD3ζ에 신호를 전달한다. CD3ζ는 최적의 활성화 및 확장을 위해 T 세포 활성화를 위한 1차 자극(2차 공-자극 신호와 함께)을 제공한다. 정상적인 조건 하에, 완전한 T-세포 활성화는 면역 반응을 증가시키는 하나 이상의 공-자극 수용체(예컨대, CD28, CD2, 4-1BBL)에 의해 매개되는 제2 신호와 함께 TCR의 결합에 좌우된다. 그러나, TCR이 존재하지 않는 경우, T 세포 확장은 작용제 항-CD3 mAb를 포함한 표준 활성화/자극 시약을 사용하여 자극될 때 심각하게 감소된다. 사실, T 세포 확장은 작용제 항-CD3 mAb를 포함한 표준 활성화/자극 시약을 사용하여 자극될 때 정상 확장 수준의 단지 20-40%로 감소된다.

[0246] 따라서, 본 개시내용은 (a) 활성화 구성요소를 포함하는 엑토도메인으로서, 활성화 구성요소는 제1 단백질로부터 단리되거나 유래되는 것인 엑토도메인; (b) 막관통 도메인; 및 (c) 적어도 하나의 신호 전달 도메인을 포함하는 엔도도메인으로서, 적어도 하나의 신호 전달 도메인은 제2 단백질로부터 단리되거나 유래되고, 제1 단백질 및 제2 단백질은 동일하지 않은 것인 엔도도메인을 포함하는 비자연발생 키메라 자극 수용체(CSR)을 제공한다.

[0247] 활성화 구성요소는 하나 이상의 T-세포 수용체(TCR)의 구성요소, TCR 복합체의 구성요소, TCR 공-수용체의 구성요소, TCR 공-자극 단백질의 구성요소, TCR 억제 단백질의 구성요소, 사이토카인 수용체, 및 활성화 구성요소의 작용제가 결합하는 케모카인 수용체의 일부를 포함할 수 있다. 활성화 구성요소는 CD2 세포외 도메인 또는 작용제가 결합하는 이의 일부를 포함할 수 있다.

[0248] 신호 전달 도메인은 하나 이상의 인간 신호 전달 도메인의 구성요소, T-세포 수용체(TCR), TCR 복합체의 구성요소, TCR 공-수용체의 구성요소, TCR 공-자극 단백질의 구성요소, TCR 억제 단백질의 구성요소, 사이토카인 수용체, 및 케모카인 수용체를 포함할 수 있다. 신호 전달 도메인은 CD3 단백질 또는 이의 일부를 포함할 수 있다. CD3 단백질은 CD3ζ 단백질 또는 이의 일부를 포함할 수 있다.

[0249] 엔도도메인은 세포질 도메인을 추가로 포함할 수 있다. 세포질 도메인은 제3 단백질로부터 단리되거나 유래될 수 있다. 제1 단백질 및 제3 단백질은 동일할 수 있다. 엑토도메인은 신호 펩타이드를 추가로 포함할 수 있다. 신호 펩타이드는 제4 단백질로부터 유래될 수 있다. 제1 단백질 및 제4 단백질은 동일할 수 있다. 막관통 도메인은 제5 단백질로부터 단리되거나 유래될 수 있다. 제1 단백질 및 제5 단백질은 동일할 수 있다.

[0250] 일부 양태에서, 활성화 구성요소는 자연발생 분자에 결합하지 않는다. 일부 양태에서, 활성화 구성요소는 자연발생 분자에 결합하지만 CSR은 활성화 구성요소가 자연발생 분자에 결합할 때 신호를 전달하지 않는다. 일부 양태에서, 활성화 구성요소는 비자연발생 분자에 결합하지 않는다. 일부 양태에서, 활성화 구성요소는 자연발생 분자에 결합하지 않지만 비자연발생 분자에 결합한다. CSR은 활성화 구성요소가 비자연발생 분자에 결합할 때 신호를 선택적으로 전달할 수 있다.

[0251] 바람직한 양태에서, 본 개시내용은 (a) 신호 펩타이드 및 활성화 구성요소를 포함하는 엑토도메인으로서, 신호 펩타이드는 CD2 신호 펩타이드 또는 이의 일부를 포함하고 활성화 구성요소는 작용제가 결합하는 CD2 세포외 도메인 또는 이의 일부를 포함하는 것인 엑토도메인; (b) 막관통 도메인으로서, 막관통 도메인은 CD2 막관통 도메인 또는 이의 일부를 포함하는 것인 막관통 도메인; 및 (c) 세포질 도메인 및 적어도 하나의 신호 전달 도메인을 포함하는 엔도도메인으로서, 세포질 도메인은 CD2 세포질 도메인 또는 이의 일부를 포함하고 적어도 하나의 신호 전달 도메인은 CD3ζ 단백질 또는 이의 일부를 포함하는 것인 엔도도메인을 포함하는 비자연발생 키메라 자극 수용체(CSR)를 제공한다. 일부 양태에서, 비자연 CSR은 서열번호: 81과 적어도 75%, 80%, 85%, 90%, 95%,

96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함한다. 바람직한 양태에서, 비자연발생 CSR은 서열번호: 81의 아미노산 서열을 포함한다.

[0252] 본 개시내용은 또한 엑토도메인이 변형을 포함하는 비자연발생 키메라 자극 수용체(CSR)를 제공한다. 변형은 활성화 구성요소 또는 제1 단백질의 야생형 서열과 비교하여 활성화 구성요소 또는 제1 단백질의 아미노산 서열의 돌연변이 또는 절단을 포함할 수 있다. 활성화 구성요소의 아미노산 서열의 돌연변이 또는 절단은작용제가 결합하는 CD2 세포외 도메인 또는 이의 일부의 돌연변이 또는 절단을 포함할 수 있다. CD2 세포외 도메인의 돌연변이 또는 절단은 자연발생 CD58과의 결합을 감소 또는 제거할 수 있다. 일부 양태에서, 돌연변이 또는 절단을 포함하는 CD2 세포외 도메인은 서열번호: 82와 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함한다. 바람직한 양태에서, 돌연변이 또는 절단을 포함하는 CD2 세포외 도메인은 서열번호: 82의 아미노산 서열을 포함한다.

[0253] 바람직한 양태에서, 본 개시내용은 (a) 신호 펩타이드 및 활성화 구성요소를 포함하는 엑토도메인으로서, 신호 펩타이드는 CD2 신호 펩타이드 또는 이의 일부를 포함하고 활성화 구성요소는 작용제가 결합하는 CD2 세포외 도메인 또는 이의 일부를 포함하며 작용제가 결합하는 CD2 세포외 도메인 또는 이의 일부는 돌연변이 또는 절단을 포함하는 것인 엑토도메인; (b) 막관통 도메인으로서, 막관통 도메인은 CD2 막관통 도메인 또는 이의 일부를 포함하는 것인 막관통 도메인; 및 (c) 세포질 도메인 및 적어도 하나의 신호 전달 도메인을 포함하는 엔도도메인으로서, 세포질 도메인은 CD2 세포질 도메인 또는 이의 일부를 포함하고 적어도 하나의 신호 전달 도메인은 CD3 ζ 단백질 또는 이의 일부를 포함하는 것인 엔도도메인을 포함하는 비자연발생 키메라 자극 수용체(CSR)를 제공한다. 일부 양태에서, 비자연 CSR은 서열번호: 83과 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%(또는 이들 사이의 임의의 백분율) 동일한 아미노산 서열을 포함한다. 바람직한 양태에서, 비자연발생 CSR은 서열번호: 83의 아미노산 서열을 포함한다.

[0254] 본 개시내용은 본원에 개시된 임의의 CSR을 코딩하는 핵산 서열을 제공한다. 본 개시내용은 본원에 개시된 임의의 CSR을 코딩하는 핵산 서열을 포함하는 트랜스포존 또는 벡터를 제공한다.

[0255] 본 개시내용은 본원에 개시된 임의의 CSR을 포함하는 세포를 제공한다. 본 개시내용은 본원에 개시된 임의의 CSR을 코딩하는 핵산 서열을 포함하는 세포를 제공한다. 본 개시내용은 본원에 개시된 임의의 CSR을 코딩하는 핵산 서열을 포함하는 벡터를 포함하는 세포를 제공한다. 본 개시내용은 본원에 개시된 임의의 CSR을 코딩하는 핵산 서열을 포함하는 트랜스포존을 포함하는 세포를 제공한다.

[0256] 본원에 개시된 변형된 세포는 동종이계 세포 또는 자가조직 세포일 수 있다. 일부 바람직한 양태에서, 변형된 세포는 동종이계 세포이다. 일부 양태에서, 변형된 세포는 자가조직 T-세포 또는 변형된 자가조직 CAR T-세포이다. 일부 바람직한 양태에서, 변형된 세포는 동종이계 T-세포 또는 변형된 동종이계 CAR T-세포이다.

[0257] 본 개시내용은 본원에 개시된 임의의 CSR을 포함하는 조성물을 제공한다. 본 개시내용은 본원에 개시된 임의의 CSR을 코딩하는 핵산 서열을 포함하는 조성물을 제공한다. 본 개시내용은 본원에 개시된 임의의 CSR을 코딩하는 핵산 서열을 포함하는 벡터를 포함하는 조성물을 제공한다. 본 개시내용은 본원에 개시된 임의의 CSR을 코딩하는 핵산 서열을 포함하는 트랜스포존을 포함하는 조성물을 제공한다. 본 개시내용은 본원에 개시된 변형된 세포를 포함하는 조성물 또는 본원에 개시된 복수의 변형된 세포를 포함하는 조성물을 제공한다.

[0258] 본 개시내용은 (a) T-세포 수용체(TCR)를 코딩하는 내인성 서열의 변형으로서, 변형은 TCR의 발현 또는 활성의 수준을 감소시키거나 제거하는 것인 변형; 및 (b) (i) 활성화 구성요소를 포함하는 엑토도메인으로서, 활성화 구성요소는 제1 단백질로부터 단리되거나 유래되는 것인 엑토도메인; (ii) 막관통 도메인; 및 (iii) 적어도 하나의 신호 전달 도메인을 포함하는 엔도도메인으로서, 적어도 하나의 신호 전달 도메인은 제2 단백질로부터 단리되거나 유래되는 것인 엔도도메인을 포함하는 키메라 자극 수용체(CSR)로서; 제1 단백질 및 제2 단백질은 동일하지 않은 것인 키메라 자극 수용체(CSR)를 포함하는 변형된 T 림프구(T-세포)를 제공한다.

[0259] 변형된 T-세포는 유도성 아폽토시스 촉진 폴리펩타이드를 추가로 포함할 수 있다. 변형된 T-세포는 베타-2-마이크로글로불린(B2M)을 코딩하는 내인성 서열의 변형을 추가로 포함할 수 있고, 변형은 주 조직 적합성 복합체(MHC) 부류 I(MHC-I)의 발현 또는 활성의 수준을 감소시키거나 제거한다.

[0260] 변형된 T-세포는 HLA 부류 I 조직적합성 항원, 알파 사슬 E(HLA-E) 폴리펩타이드를 포함하는 비자연발생 폴리펩타이드를 추가로 포함할 수 있다. HLA-E 폴리펩타이드를 포함하는 비자연발생 폴리펩타이드는 B2M 신호 펩타이드를 추가로 포함할 수 있다. HLA-E 폴리펩타이드를 포함하는 비자연발생 폴리펩타이드는 B2M 폴리펩타이드를 추가로 포함할 수 있다. HLA-E 폴리펩타이드를 포함하는 비자연발생 폴리펩타이드는 링커를 추가로 포함할 수

있고, 링커는 B2M 폴리펩타이드 및 HLA-E 폴리펩타이드 사이에 위치한다. HLA-E 폴리펩타이드를 포함하는 비자연발생 폴리펩타이드는 펩타이드 및 B2M 폴리펩타이드를 추가로 포함할 수 있다. HLA-E 폴리펩타이드를 포함하는 비자연발생 폴리펩타이드는 B2M 신호 펩타이드 및 펩타이드 사이에 위치한 제1 링커, 및 B2M 폴리펩타이드 및 HLA-E를 코딩하는 펩타이드 사이에 위치한 제2 링커를 추가로 포함할 수 있다.

[0261] 변형된 T-세포는 비자연발생 항원 수용체, 치료적 폴리펩타이드를 코딩하는 서열, 또는 이의 조합을 추가로 포함할 수 있다. 비자연발생 항원 수용체는 키메라 항원 수용체(CAR)를 포함할 수 있다.

[0262] CSR은 변형된 T-세포에서 일시적으로 발현될 수 있다. CSR은 변형된 T-세포에서 안정적으로 발현될 수 있다. HLA-E 폴리펩타이드를 포함하는 폴리펩타이드는 변형된 T-세포에서 일시적으로 발현될 수 있다. HLA-E 폴리펩타이드를 포함하는 폴리펩타이드는 변형된 T-세포에서 안정적으로 발현될 수 있다. 유도성 아폽토시스 촉진 폴리펩타이드는 변형된 T-세포에서 일시적으로 발현될 수 있다. 유도성 아폽토시스 촉진 폴리펩타이드는 변형된 T-세포에서 안정적으로 발현될 수 있다. 비자연발생 항원 수용체 또는 치료적 단백질을 코딩하는 서열은 변형된 T-세포에서 일시적으로 발현될 수 있다. 비자연발생 항원 수용체 또는 치료적 단백질을 코딩하는 서열은 변형된 T-세포에서 안정적으로 발현될 수 있다.

[0263] 비제한적으로 본원에 상세히 기재된 바와 같은 dCas9-Clo051을 포함하는 RNA 가이드된 융합 단백질을 포함하는 유전자 편집 조성물은 내인성 T-세포 수용체의 발현을 표적화하고 감소시키거나 제거하는 데 사용될 수 있다. 바람직한 양태에서, 유전자 편집 조성물은 내인성 T-세포 수용체를 코딩하는 유전자, 유전자의 일부, 또는 유전자의 조절 요소(예컨대, 프로모터)를 표적화하고 결실시킨다. TCR-알파(TCR-α)를 표적화하고 결실시키고, TCR-베타(TCR-β)를 표적화하고 결실시키며, 베타-2-마이크로글로불린(β2M)을 표적화하고 결실시키기 위한 가이드 RNA(gRNA) 주형의 생성을 위한 프라이머(T7 프로모터, 게놈 표적 서열, 및 gRNA 스캐폴드 포함)의 비제한적인 예가 PCT 출원 제PCT/US2019/049816호에 개시되어 있다.

[0264] 비제한적으로, dCas9-Clo051을 포함하는 RNA 가이드된 융합 단백질을 포함하는 유전자 편집 조성물은 내인성 MHCI, MHCII, 또는 MHC 활성화제의 발현을 표적화하고 감소시키거나 제거하는 데 사용될 수 있다. 바람직한 양태에서, 유전자 편집 조성물은 내인성 MHCI, MHCII, 또는 MHC 활성화제의 하나 이상의 구성요소를 코딩하는 유전자, 유전자의 일부, 또는 유전자의 조절 요소(예컨대, 프로모터)를 표적화하고 결실시킨다. MHC 활성화제를 표적화하고 결실시키기 위한 가이드 RNA(gRNA)의 비제한적인 예가 PCT 출원 제PCT/US2019/049816호에 개시되어 있다.

[0265] 비자연발생 키메라 자극 수용체, TCR-알파(TCR-α), TCR-베타(TCR-β), 및/또는 베타-2-마이크로글로불린(β2M)을 코딩하는 내인성 서열의 유전적 변형, 및 HLA 부류 I 조직적합성 항원, 알파 사슬 E(HLA-E) 폴리펩타이드를 포함하는 비자연발생 폴리펩타이드의 상세한 설명이 PCT 출원 제PCT/US2019/049816호에 개시되어 있다.

[0266] **제제, 투여량 및 투여**

[0267] 본 개시내용은 본원에 기재된 조성물의 제제, 투여량 및 투여 방법을 제공한다.

[0268] 개시된 조성물 및 약학 조성물은, 비제한적으로, 희석제, 결합제, 안정화제, 완충제, 염, 친유성 용매, 보존제, 보조제 등과 같은 임의의 적합한 보조제 중 적어도 하나를 추가로 포함할 수 있다. 약학적으로 허용가능한 보조제가 바람직하다. 이러한 멸균 용액의 비제한적인 예 및 이의 제조 방법은 비제한적으로 문헌[Gennaro, Ed., Remington's Pharmaceutical Sciences, 18th Edition, Mack Publishing Co.(Easton, Pa.) 1990 및 "Physician's Desk Reference", 52nd ed., Medical Economics(Montvale, N.J.) 1998]에서와 같이 당업계에 널리 알려져 있다. 당업계에 널리 알려져 있거나 본원에 기재된 바와 같은 단백질 스캐폴드, 단편 또는 변이체 조성물의 투여 방식, 용해도 및/또는 안정성에 적합한 약학적으로 허용가능한 담체가 일상적으로 선택될 수 있다.

[0269] 사용하기에 적합한 약제학적 부형제 및 첨가제의 비제한적인 예는 단백질, 펩타이드, 아미노산, 지질, 및 탄수화물(예컨대, 단당류, 이당류, 삼당류, 사당류, 및 올리고당을 포함하는 당류; 유도체화된 당류, 예컨대 알디톨(alditol), 알돈산(aldonic acid), 에스테르화 당류 등; 및 다당류 또는 당 중합체)을 포함하고, 이는 단독으로 또는 조합하여 존재할 수 있고, 이는 단독으로 또는 조합하여 1-99.99 중량% 또는 부피%를 포함한다. 단백질 부형제의 비제한적인 예는 혈청 알부민, 예컨대 인간 혈청 알부민(HSA), 재조합 인간 알부민(rHA), 젤라틴, 카세인 등을 포함한다. 완충 용량에서도 기능할 수 있는 대표적인 아미노산/단백질 구성요소는 알라닌, 글리신, 아르기닌, 베타인, 히스티딘, 글루탐산, 아스파르트산, 시스테인, 리신, 류신, 이소류신, 발린, 메티오닌, 페닐알라닌, 아스파르탐 등을 포함한다. 하나의 바람직한 아미노산은 글리신이다.

[0270] 사용하기에 적합한 탄수화물 부형제의 비제한적인 예는 단당류, 예컨대 프럭토스, 말토스, 갈락토스, 글루코스,

D-만노스, 소르보스 등; 이당류, 예컨대 락토스, 수크로스, 트레할로스, 셀로바이오스 등; 다당류, 예컨대 라피노스, 멜레지토스, 말토덱스트린, 덱스트란, 전분 등; 및 알디톨, 예컨대, 만니톨, 자일리톨, 말티톨, 락티톨, 자일리톨 소르비톨(glucitol), 미오이노시톨 등을 포함한다. 바람직하게는, 탄수화물 부형제는 만니톨, 트레할로스, 및/또는 라피노스이다.

[0271] 조성물은 또한 완충제 또는 pH-조절제를 포함하고; 전형적으로, 완충제는 유기 산 또는 염기로부터 제조된 염이다. 대표적인 완충제는 유기 산 염, 예컨대 시트르산, 아스코르브산, 글루콘산, 카르본산, 타르타르산, 석신산, 아세트산, 또는 프탈산의 염; 트리스, 트로메타민 하이드로클로라이드, 또는 포스페이트 완충제를 포함한다. 바람직한 완충제는 유기 산 염, 예컨대 시트레이트이다.

[0272] 추가로, 개시된 조성물은 중합체성 부형제/첨가제, 예컨대 폴리비닐피롤리돈, 피콜(중합체성 당), 덱스트레이트(예컨대, 사이클로덱스트린, 예컨대 2-하이드록시프로필-β-사이클로덱스트린), 폴리에틸렌 글리콜, 향미제, 항균제, 감미제, 항산화제, 정균제, 계면활성제(예컨대, 폴리소르베이트, 예컨대 "TWEEN 20" 및 "TWEEN 80"), 지질(예컨대, 인지질, 지방산), 스테로이드(예컨대, 콜레스테롤), 및 킬레이트제(예컨대, EDTA)를 포함할 수 있다.

[0273] 본원에 개시된 조성물 또는 약학 조성물의 치료적 유효량을 투여하기 위해 많은 공지되고 개발된 방식이 사용될 수 있다. 투여 방식의 비제한적인 예는 볼루스, 협측, 주입, 관절내, 기관지내, 복내(intraabdominal), 피막내, 연골내, 강내, 복내, 소뇌내, 뇌실내, 결장내, 자궁경부내, 위내, 간내, 병변내, 근육내, 심근내, 비강내, 안내, 골내(intraosseous), 골내(intraosteal), 골반내, 심낭내, 복강내(intraperitoneal), 흉막내, 전립선내, 폐내, 직장내, 신장내, 망막내, 척수내, 활막내, 흉부내, 자궁내, 종양내, 정맥내, 방광내, 경구, 비경구, 직장, 설하, 피하, 경피 또는 질 수단을 포함한다.

[0274] 본 개시내용의 조성물은 비경구(피하, 근육내 또는 정맥내) 또는 임의의 다른 투여의 사용을 위해, 특히 액체 용액 또는 현탁액의 형태로; 질 또는 직장 투여에 사용하기 위해, 특히 반고체 형태, 예컨대, 비제한적으로, 크림 및 좌제로; 협측, 또는 설하 투여를 위해, 예컨대, 비제한적으로, 정제 또는 캡슐 형태로; 또는 비강내로, 예컨대, 비제한적으로, 분말, 점비제 또는 에어로졸 또는 특정 제제의 형태로; 또는 경피로, 예컨대 비제한적으로 피부 구조를 변형시키거나 경피 패치에서 약물 농도를 증가시키기 위해 디메틸 설폭사이드와 같은 화학적 강화제를 갖거나(Junginger, et al. In "Drug Permeation Enhancement;" Hsieh, D. S., Eds., pp. 59-90(Marcel Dekker, Inc. New York 1994,), 또는 피부에 단백질 및 펩타이드를 함유하는 제제의 적용(WO 제98/53847호), 또는 전기천공과 같은 일시적인 수송 경로를 생성하거나, 또는 이온영동(iontophoresis)과 같이 피부를 통해 하전된 약물의 이동성을 증가시키는 전하의 적용, 또는 초음파영동(sonophoresis)과 같은 초음파의 적용(미국 특허 제4,309,989호 및 제4,767,402gh)(상기 간행물 및 특허는 전체가 참조로 본원에 포함됨)을 가능하게 하는 산화제를 갖는 겔, 연고, 로션, 현탁액 또는 패치 전달 시스템으로 제조될 수 있다.

[0275] 비경구 투여를 위해, 본원에 개시된 임의의 조성물은 약학적으로 허용가능한 비경구 비히클과 함께 회합되거나 또는 별도로 제공되는 용액, 현탁액, 유화액, 입자, 분말, 또는 동결건조 분말로서 제제화될 수 있다. 비경구 투여를 위한 제제는 일반적인 부형제로서 멸균수 또는 식염수, 폴리알킬렌 글리콜, 예컨대 폴리에틸렌 글리콜, 식물성 오일, 수소화된 나프탈렌 등을 함유할 수 있다. 주사를 위한 수성 또는 유성 현탁액은 공지된 방법에 따라 적절한 유화제 또는 가습제 및 현탁제를 사용하여 제조될 수 있다. 주사를 위한 제제는 용매 중의 비독성, 비경구 투여가능한 희석제, 예컨대 수성 용액, 멸균 주사가능한 용액 또는 현탁액일 수 있다. 사용가능한 비히클 또는 용매로서, 물, 링거액, 등장 식염수 등이 허용되고; 일반 용매 또는 현탁 용매로서, 멸균 비휘발성 오일이 사용될 수 있다. 이러한 목적을 위해, 천연 또는 합성 또는 반합성 지방유 또는 지방산; 천연 또는 합성 또는 반합성 모노- 또는 디- 또는 트리-글리세라이드를 포함하는 임의의 종류의 비휘발성 오일 및 지방산이 사용될 수 있다. 비경구 투여는 당업계에 공지되어 있으며, 비제한적으로, 일반적인 주사 수단, 미국 특허 제5,851,198호에 기재된 바와 같은 가스 압력 무바늘 주사 장치, 및 미국 특허 제5,839,446호에 기재된 바와 같은 레이저 천공기 장치를 포함한다.

[0276] 경구 투여를 위한 제제는 장벽의 투과성을 인공적으로 증가시키기 위한 보조제(예컨대, 레소르시놀 및 비이온성 계면활성제, 예컨대 폴리옥시에틸렌 올레일 에테르 및 n-헥사데실폴리에틸렌 에테르)의 공동 투여뿐만 아니라 효소 분해를 억제하기 위한 효소 억제제(예컨대, 췌장 트립신 억제제, 디이소프로필플루오로포스페이트(DFF) 및 트라실롤)의 공동 투여에 의존한다. 단백질 및 단백질 스캐폴드를 포함하는 친수성 제제 및 경구, 협측, 점막, 비강, 폐, 질 막관통, 또는 직장 투여용으로 의도된 적어도 2개의 계면활성제의 조합의 전달을 위한 제제가 미국 특허 제6,309,663호에 기재되어 있다. 경구 투여를 위한 고체 유형 제형의 활성 성분 화합물은 수크로스, 락

토스, 셀룰로오스, 만니톨, 트레할로스, 라피노스, 말티톨, 덱스트란, 전분, 아가, 아르기네이트, 키틴, 키토산, 펙틴, 트라가칸트 검, 아라비아 검, 젤라틴, 콜라겐, 카세인, 알부민, 합성 또는 빈합성 중합체, 및 글리세라이드를 포함하는 적어도 하나의 첨가제와 혼합될 수 있다. 이러한 투여 형태는 또한 다른 유형(들)의 첨가제, 예컨대, 불활성 희석제, 윤활제, 예컨대 마그네슘 스테아레이트, 파라벤, 보존제, 예컨대 소르브산, 아스코르브산, 알파-토코페롤, 항산화제, 예컨대 시스테인, 붕해제, 결합제, 증점제, 완충제, 감미제, 향미제, 방향제 등을 포함할 수 있다.

[0277] 정제 및 환제는 장용 코팅 제제로 추가로 가공될 수 있다. 경구 투여를 위한 액체 제제는 의약용으로 허용가능한 유화액, 시럽, 엘릭시르, 현탁액 및 용액 제제를 포함한다. 이러한 제제는 상기 분야에서 통상적으로 사용되는 불활성 희석제, 예컨대, 물을 함유할 수 있다. 리포좀은 또한 인슐린 및 헤파린에 대한 약물 전달 시스템으로서 기술되었다(미국 특허 제4,239,754호). 보다 최근에는, 혼합된 아미노산의 인공 중합체의 미소구체(프로티노이드(proteinoid))가 의약품을 전달하는 데 사용되었다(미국 특허 제4,925,673호). 또한, 미국 특허 제5,879,681호 및 미국 특허 제5,871,753호에 기재되고 생물학적 활성제를 경구로 전달하는 데 사용된 담체 화합물은 당업자에 공지되어 있다.

[0278] 폐 투여를 위해, 바람직하게는, 본원에 기재된 조성물 또는 약학 조성물은 폐 또는 부비동의 하기도에 도달하기에 효과적인 입자 크기로 전달된다. 조성물 또는 약학 조성물은 흡입에 의한 치료제의 투여를 위해 당업계에 공지된 다양한 흡입 또는 비강 장치에 의해 전달될 수 있다. 환자의 부비동 또는 폐포에 에어로졸화된 제제를 침착시킬 수 있는 이러한 장치는 계량 흡입기, 분무기(예컨대, 제트 분무기, 초음파 분무기), 건조 분말 생성기, 분무기 등을 포함한다. 이러한 모든 장치는 본원에 기재된 조성물 또는 약학 조성물을 에어로졸에 분배하기 위해 투여에 적합한 제제를 사용할 수 있다. 이러한 에어로졸은 용액(수성 및 비수성 모두) 또는 고체 입자로 구성될 수 있다. 추가로, 본원에 기재된 조성물 또는 약학 조성물을 포함하는 스프레이는 적어도 하나의 단백질 스캐폴드의 현탁액 또는 용액을 압력 하에 노즐을 통과시킴으로써 생산될 수 있다. 계량 흡입기(MDI)에서, 추진제, 본원에 기재된 조성물 또는 약학 조성물, 및 임의의 부형제 또는 다른 첨가제가 액화 압축 가스를 포함하는 혼합물로서 캐니스터(canister)에 포함된다. 계량 밸브의 구동은, 바람직하게는 약 10 μm 미만, 바람직하게는, 약 1 μm 내지 약 5 μm, 및, 가장 바람직하게는, 약 2 μm 내지 약 3 μm 크기 범위의 입자를 함유하는 에어로졸로서 혼합물을 방출한다. 폐 투여, 제제 및 관련 장치의 보다 상세한 설명은 PCT 공개 WO 제2019/049816호에 개시되어 있다.

[0279] 점막 표면을 통한 흡수를 위해, 조성물은 복수의 서브마이크론 입자, 점막접착성 거대분자, 생활성 펩타이드, 및 유화액 입자의 점막접착을 달성함으로써 점막 표면을 통한 흡수를 촉진하는 수성 연속상을 포함하는 유화액을 포함한다(미국 특허 제5,514,670호). 본 개시내용의 유화액의 적용에 적합한 점막 표면은 각막, 결막, 협측, 설하, 비강, 질, 폐, 위, 장, 및 직장 투여 경로를 포함할 수 있다. 질 또는 직장 투여를 위한 제제, 예컨대, 좌제는 부형제, 예를 들어, 폴리알킬렌글리콜, 바셀린, 코코넛 버터 등을 함유할 수 있다. 비강내 투여를 위한 제제는 고체일 수 있고, 부형제로서, 예를 들어, 락토스를 함유할 수 있거나 점비제의 수성 또는 유성 용액일 수 있다. 협측 투여를 위해, 부형제는 당류, 칼슘 스테아레이트, 마그네슘 스테아레이트, 전호화된 전분 등을 포함한다(미국 특허 제5,849,695호). 점막 투여 및 제제의 보다 상세한 설명이 PCT 공개 WO 제2019/049816호에 개시되어 있다.

[0280] 경피 투여를 위해, 본원에 개시된 조성물 또는 약학 조성물은 리포좀 또는 중합체성 나노입자, 마이크로입자, 마이크로캡슐, 또는 미소구체(달리 언급되지 않는 한 집합적으로 마이크로입자로서 지칭됨)와 같은 전달 장치에 캡슐화된다. 합성 중합체, 예컨대 폴리하이드록시산, 예컨대 폴리락트산, 폴리글리콜산 및 이의 공중합체, 폴리오르토에스테르, 폴리무수물, 및 폴리포스파젠, 및 천연 중합체, 예컨대 콜라겐, 폴리아미노산, 알부민 및 다른 단백질, 알지네이트 및 다른 다당류, 및 이의 조합으로 제조된 마이크로입자를 포함하는 많은 적합한 장치가 알려져 있다(미국 특허 제5,814,599호). 경피 투여, 제제 및 적합한 장치의 보다 상세한 설명이 PCT 공개 WO 제2019/049816호에 개시되어 있다.

[0281] 개시된 화합물을 장기간에 걸쳐, 예를 들어, 단일 투여로부터 1주 내지 1년의 기간 동안 대상체에게 전달하는 것이 바람직할 수 있다. 다양한 서방성, 데포 또는 임플란트 투여 형태가 이용될 수 있다. 예를 들어, 투여 형태는 체액에서 낮은 정도의 용해도를 갖는 화합물의 약학적으로 허용가능한 무독성 염, 예를 들어, (a) 다가산(polybasic acid), 예컨대 인산, 황산, 시트르산, 타르타르산, 탄닌산, 파모산, 알긴산, 폴리글루탐산, 나프탈렌 모노- 또는 디-설폰산, 폴리갈락투론산 등을 갖는 산 부가 염; (b) 다가 금속 양이온, 예컨대 아연, 칼슘, 비스무스, 바륨, 마그네슘, 알루미늄, 구리, 코발트, 니켈, 카드뮴 등을 갖는 염 또는 예컨대, N,N'-디벤질-에틸렌디아민 또는 에틸렌디아민으로부터 형성된 유기 양이온을 갖는 염; 또는 (c) (a) 및 (b)의 조합, 예컨대,

아연 탄네이트 염을 함유할 수 있다. 추가로, 개시된 화합물 또는, 바람직하게는, 방금 기재된 것과 같은 비교적 불용성인 염은, 예컨대, 주사에 적합한 참기름을 이용하여 겔, 예를 들어, 알루미늄 모노스테아레이트 겔로 제제화될 수 있다. 특히 바람직한 염은 아연 염, 아연 탄네이트 염, 파모에이트 염 등이다. 주사를 위한 서방성 데포 제제의 또 다른 유형은 느리게 분해하는 무독성, 비항원성 중합체, 예컨대 미국 특허 제3,773,919호에 기재된 바와 같은 폴리락트산/폴리글리콜산 중합체에서 캡슐화를 위해 분산된 화합물 또는 염을 함유할 수 있다. 화합물 또는, 바람직하게는, 상기 기재된 것과 같은 비교적 불용성인 염은 또한 특히 동물에서 사용하기 위해 콜레스테롤 매트릭스 실라스틱(silastic) 펠렛으로 제제화될 수 있다. 추가적인 서방성, 데포 또는 임플란트 제제, 예컨대, 기체 또는 액체 리포좀은 문헌에 공지되어 있다(미국 특허 제5,770,222호 및 "Sustained and Controlled Release Drug Delivery Systems", J. R. Robinson ed., Marcel Dekker, Inc., N.Y., 1978).

[0282] 적합한 투여량은 당업계에 잘 알려져 있다. 예컨대, 문헌[Wells et al., eds., Pharmacotherapy Handbook, 2nd Edition, Appleton and Lange, Stamford, Conn. (2000); PDR Pharmacopoeia, Tarascon Pocket Pharmacopoeia 2000, Deluxe Edition, Tarascon Publishing, Loma Linda, Calif. (2000); Nursing 2001 Handbook of Drugs, 21st edition, Springhouse Corp., Springhouse, Pa., 2001; Health Professional's Drug Guide 2001, ed., Shannon, Wilson, Stang, Prentice-Hall, Inc, Upper Saddle River, N.J]을 참고한다. 바람직한 용량은 선택적으로 약 0.1-99 및/또는 100-500 mg/kg/투여, 또는 이의 임의의 범위, 값 또는 분수를 포함할 수 있거나, 또는 단일 또는 다중 투여당 약 0.1-5000 μg/ml 혈청 농도, 또는 이의 임의의 범위, 값 또는 분수를 달성할 수 있다. 본원에 개시된 조성물 또는 약학 조성물을 위한 바람직한 투여량 범위는 대상체의 체중의 약 1 mg/kg, 최대 약 3, 약 6 또는 약 12 mg/kg이다.

[0283] 대안적으로, 투여되는 투여량은 공지된 인자, 예컨대 특정 제제의 약력학적 특징, 및 그의 투여 방식 및 경로; 수여자의 연령, 건강, 및 체중; 증상의 성질 및 정도, 동시 치료의 종류, 치료 빈도, 및 원하는 효과에 따라 달라질 수 있다. 일반적으로 활성 성분의 투여량은 체중의 킬로그램당 약 0.1 내지 100 밀리그램일 수 있다. 일반적으로, 투여당 또는 지속 방출 형태에서 킬로그램당 0.1 내지 50, 및 바람직하게는, 0.1 내지 10 밀리그램이 원하는 결과를 얻는 데 효과적이다.

[0284] 비제한적인 예로서, 인간 또는 동물의 치료는 단일, 주입 또는 반복 용량을 사용하여, 본원에 개시된 조성물 또는 약학 조성물의 1회 또는 주기적 투여량으로서 매일 약 0.1 내지 100 mg/kg 또는 이의 임의의 범위, 값 또는 분수, 1-40일 중 적어도 하나, 또는 대안적으로 또는 추가적으로, 1-52주 중 적어도 하나, 또는, 대안적으로 또는 추가적으로, 1-20년 중 적어도 하나, 또는 이의 임의의 조합으로 제공될 수 있다.

[0285] 내부 투여에 적합한 투여 형태는 일반적으로 단위 또는 용기당 약 0.001 밀리그램 내지 약 500 밀리그램의 활성 성분을 함유한다. 이러한 약학 조성물에서, 활성 성분은 일반적으로 조성물의 총 중량을 기준으로 약 0.5-99.999 중량%의 양으로 존재할 것이다.

[0286] 유효량은 단일(예컨대, 볼루스), 다중 또는 연속 투여당 약 0.001 내지 약 500 mg/kg의 양을 포함할 수 있거나, 본원에 기재되거나 관련 분야에서 알려진 바와 같이, 공지된 방법을 사용하여 수행되고 결정된 바와 같이, 단일, 다중, 또는 연속 투여당 0.01-5000 μg/ml 혈청 농도, 또는 이 안의 임의의 효과적인 범위 또는 값의 혈청 농도를 달성하는 양을 포함할 수 있다.

[0287] 이를 필요로 하는 대상체에게 투여될 조성물이 본원에 개시된 바와 같은 변형된 세포인 양태에서, 세포는 약 $1x10^3$ 내지 $1x10^{15}$ 세포; 약 $1x10^4$ 내지 $1x10^{12}$ 세포; 약 $1x10^5$ 내지 $1x10^{10}$ 세포; 약 $1x10^6$ 내지 $1x10^9$ 세포; 약 $1x10^6$ 내지 $1x10^8$ 세포; 약 $1x10^6$ 내지 $1x10^7$ 세포; 또는 약 $1x10^6$ 내지 $25x10^6$ 세포로 투여될 수 있다. 일 양태에서, 세포는 약 $5x10^6$ 내지 $25x10^6$ 세포로 투여된다.

[0288] 약학적으로 허용가능한 부형제, 제제, 투여량 및 개시된 조성물 및 약학 조성물의 투여 방법의 보다 상세한 설명이 PCT 공개 WO 제2019/049816호에 개시되어 있다.

[0289] **본 개시내용의 조성물을 사용하는 방법**

[0290] 본 개시내용은 전위 효율을 개선하기 위한 개시된 조성물의 용도를 제공한다. 구체적으로, 방법은 세포 또는 복수의 세포를 (a) 제1 역위 말단 반복(ITR) 또는 제1 ITR을 코딩하는 서열, (b) 제2 ITR 또는 제2 ITR을 코딩하는 서열, 및 (c) 인트라-ITR 서열 또는 인트라-ITR을 코딩하는 서열을 포함하는 제1 핵산 서열로서, 인트라-ITR 서열은 트랜스포존 서열 또는 트랜스포존을 코딩하는 서열을 포함하는 것인 제1 핵산 서열; 및 인터-ITR 서열 또는 인터-ITR을 코딩하는 서열을 포함하는 제2 핵산 서열로서, 인터-ITR 서열의 길이는 700개 이하의 뉴클레오

타이드이고, 세포 또는 복수의 세포 내의 전위 효율은 700개 뉴클레오타이드보다 큰 제2 핵산 서열 또는 인터-ITR 서열을 포함하는 동일한 조성물과 비교하여 개선되는 것인 제2 핵산 서열을 포함하는 조성물과 접촉시키는 단계를 포함한다. 일 양태에서, 전위 효율은 적어도 5%, 적어도 10%, 적어도 15%, 적어도 20%, 적어도 25%, 적어도 30%, 적어도 35%, 적어도 40%, 적어도 45%, 적어도 50%, 적어도 55%, 적어도 60%, 적어도 65%, 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 적어도 95%, 적어도 90% 개선된다.

[0291] 본 개시내용은 개시된 조성물 및 약학 조성물을 사용하여, 예컨대 세포, 조직, 장기, 동물, 또는 대상체에게 조성물 또는 약학 조성물의 치료적 유효량을 투여하거나 이와 접촉시켜, 당업계에 공지되거나 본원에 기재된 바와 같이 세포, 조직, 장기, 동물, 또는 대상체에서 질환 또는 장애를 치료하기 위한 개시된 조성물 또는 약학 조성물의 용도를 제공한다. 일 양태에서, 대상체는 포유동물이다. 바람직하게는, 대상체는 인간이다. 용어 "대상체" 및 "환자"는 본원에서 상호교환적으로 사용된다.

[0292] 본 개시내용은 세포, 조직, 장기, 동물 또는 대상체에서 적어도 하나의 악성 질환 또는 장애를 조절하거나 치료하기 위한 방법을 제공한다. 바람직하게는, 악성 질환은 암이다. 악성 질환 또는 장애의 비제한적인 예는 백혈병, 급성 백혈병, 급성 림프구성 백혈병(acute lymphoblastic leukemia, ALL), 급성 림프성 백혈병(acute lymphocytic leukemia), B-세포, T-세포 또는 FAB ALL, 급성 골수성 백혈병(acute myeloid leukemia, AML), 급성 골수성 백혈병(acute myelogenous leukemia), 만성 골수성 백혈병(chronic myelocytic leukemia, CML), 만성 림프성 백혈병(chronic lymphocytic leukemia, CLL), 털세포 백혈병(hairy cell leukemia), 골수이형성 증후군(myelodyplastic syndrome, MDS), 호지킨병, 악성 림프종, 비호지킨 림프종, 버킷 림프종, 다발성 골수종, 카포시 육종, 결장직장 암종, 췌장 암종, 비인두 암종, 악성 조직구증, 부신생물 증후군/악성 고칼슘혈증, 고형 종양, 방광암, 유방암, 결장직장암, 자궁내막암, 두부암, 경부암, 유전성 비용종증 암, 호지킨 림프종, 간암, 폐암, 비소 세포 폐암, 난소암, 췌장암, 전립선암, 신장 세포 암종, 고환암, 선암종, 육종, 악성 흑색종, 혈관종, 전이성 질환, 암 관련 골 흡수, 암 관련 골 통증 등을 포함한다.

[0293] 바람직한 양태에서, 악성 질환 또는 장애의 치료는 입양 세포 요법을 포함한다. 예를 들어, 일 양태에서, 본 개시내용은 이를 필요로 하는 대상체에게 투여하기 위해 선택되고/거나 확장된 단백질 스캐폴드(예컨대, 본 개시내용의 조성물을 이용하여 세포에 전달된 scFv, 단일 도메인 항체, 센티린)를 포함하는 적어도 하나의 개시된 단백질 스캐폴드 및/또는 CAR을 발현하는 변형된 세포를 제공한다. 변형된 세포는 실온 및 체온을 포함하는 임의의 온도에서 보관하기 위해 제제화될 수 있다. 변형된 세포는 동결보존 및 후속 해동을 위해 제제화될 수 있다. 변형된 세포는 멸균 포장으로부터 대상체에게 직접 투여하기 위해 약학적으로 허용가능한 담체에서 제제화될 수 있다. 변형된 세포는 최소 수준의 세포 기능 및 CAR 발현을 보장하기 위해 세포 생존력 및/또는 CAR 발현 수준의 지표와 함께 약학적으로 허용가능한 담체에서 제제화될 수 있다. 변형된 세포는 추가 확장을 억제하고/거나 세포 사멸을 방지하기 위해 하나 이상의 시약과 함께 규정된 밀도로 약학적으로 허용가능한 담체에서 제제화될 수 있다.

[0294] 임의의 방법은 본원에 개시된 임의의 조성물 또는 약학 조성물의 유효량을 이러한 조절, 치료 또는 요법을 필요로 하는 세포, 조직, 장기, 동물 또는 대상체에게 투여하는 단계를 포함할 수 있다. 이러한 방법은 선택적으로 이러한 질환 또는 장애를 치료하기 위한 공-투여 또는 조합 요법을 추가로 포함할 수 있고, 본원에 개시된 임의의 조성물 또는 약학 조성물의 투여는 적어도 하나의 화학치료제(예컨대, 알킬화제, 유사분열 억제제, 방사성의 약품)의 전, 동시에 및/또는 후에 투여하는 것을 추가로 포함한다.

[0295] 일부 양태에서, 대상체는 투여 후에 이식편 대 숙주(GvH) 및/또는 숙주 대 이식편(HvG)을 발생시키지 않는다. 일 양태에서, 투여는 전신이다. 전신 투여는 당업계에 공지되고 본원에 상세히 기재된 임의의 수단일 수 있다. 바람직하게는, 전신 투여는 정맥내 주사 또는 정맥내 주입에 의한 것이다. 일 양태에서, 투여는 국소이다. 국소 투여는 당업계에 공지되고 본원에 상세히 기재된 임의의 수단일 수 있다. 바람직하게는, 국소 투여는 종양내 주사 또는 주입, 척수내 주사 또는 주입, 뇌실내 주사 또는 주입, 안내 주사 또는 주입, 또는 골내 주사 또는 주입에 의한다.

[0296] 일부 양태에서, 치료적 유효 용량은 단일 용량이다. 일부 양태에서, 단일 용량은 동시에 제조된 적어도 2, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100개 또는 이들 사이의 임의의 수의 용량 중 하나이다. 일부 양태에서, 조성물이 자가조직 세포 또는 동종이계 세포인 경우, 용량은 세포가 질환 또는 장애를 치료하기에 충분한 시간 동안 생착 및/또는 지속하기에 충분한 양이다.

[0297] 일 예에서, 본 개시내용은 단백질 스캐폴드를 포함하는 조성물 또는 단백질 스캐폴드(예컨대, scFv, 단일 도메인 항체, 센티린)를 포함하는 CAR을 대상체에게 투여하는 단계를 포함하는 이를 필요로 하는 대상체에서 암을

치료하는 방법을 제공하며, 항체 또는 CAR은 종양 세포 상의 항원에 특이적으로 결합한다. 조성물이 변형된 세포 또는 세포 집단을 포함하는 양태에서, 세포 또는 세포 집단은 자가조직 또는 동종이계일 수 있다.

[0298] 본원에 기재된 치료 방법의 일부 양태에서, 치료는 변형되거나 종료될 수 있다. 구체적으로, 치료에 사용되는 조성물이 유도성 아폽토시스 촉진 폴리펩타이드를 포함하는 양태에서, 아폽토시스는 세포를 유도제와 접촉시킴으로써 세포에서 선택적으로 유도될 수 있다. 치료는, 예를 들어, 회복의 징후 또는 질환 중증도/진행의 감소 징후, 질환 완화/중단의 징후, 및/또는 부정적인 사건의 발생에 대한 응답으로 변형되거나 종료될 수 있다. 일부 양태에서, 방법은 유도제의 억제제를 투여하여 세포 요법의 변형을 억제함으로써 세포 요법의 기능 및/또는 효능을 회복시키는 단계를 포함한다(예를 들어, 질병의 징후 또는 증상이 다시 나타나거나 중증도의 증가 및/또는 부정적인 사건이 해결되는 경우).

[0299] **단백질 스캐폴드 생산, 스크리닝 및 정제**

[0300] 본 개시내용의 적어도 하나의 단백질 스캐폴드(예컨대, 단클론 항체, 키메라 항체, 단일 도메인 항체, VHH, VH, 단일 사슬 가변 단편(scFv), 센티린, 항원-결합 단편(Fab) 또는 Fab 단편)는 선택적으로 당업계에 널리 알려진 바와 같은 세포주, 혼합된 세포주, 불멸화된 세포 또는 불멸화된 세포의 클론 집단에 의해 생산될 수 있다. 예컨대, 문헌[Ausubel, et al., ed., Current Protocols in Molecular Biology, John Wiley & Sons, Inc., NY, N.Y. (1987-2001); Sambrook, et al., Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor, N.Y. (1989); Harlow and Lane, Antibodies, a Laboratory Manual, Cold Spring Harbor, N.Y. (1989); Colligan, et al., eds., Current Protocols in Immunology, John Wiley & Sons, Inc., NY (1994-2001); Colligan et al., Current Protocols in Protein Science, John Wiley & Sons, NY, N.Y., (1997-2001)]을 참고한다.

[0301] 단백질 스캐폴드로부터의 아미노산은 면역원성을 감소시키거나 결합, 친화성, 온-속도, 오프-속도, 결합력, 특이성, 반감기, 안정성, 용해도 또는 당업계에 공지된 바와 같은 임의의 다른 적합한 특징을 감쇠, 향상 또는 변형시키기 위해 변경, 부가 및/또는 결실될 수 있다.

[0302] 선택적으로, 단백질 스캐폴드는 항원에 대한 높은 친화성 및 다른 유리한 생물학적 특성을 유지하도록 조작될 수 있다. 이러한 목표를 달성하기 위해, 스캐폴드 단백질은 선택적으로 부모 및 조작된 서열의 3차원 모델을 사용하여 부모 서열 및 다양한 개념적 조작된 생성물의 분석 과정에 의해 제조될 수 있다. 3차원 모델이 일반적으로 이용가능하며, 당업자에게 익숙하다. 선택된 후보 서열의 가능한 3차원 형태 구조를 예시하고 표시하며 가능한 면역원성을 측정할 수 있는 컴퓨터 프로그램(예컨대, 캘리포니아 몬로비아의 Xencor사의 Immunofilter 프로그램)이 이용가능하다. 이러한 디스플레이의 검사는 후보 서열의 기능에서 잔기의 가능한 역할의 분석, 즉 그의 항원에 결합하는 후보 단백질 스캐폴드의 능력에 영향을 미치는 잔기의 분석을 허용한다. 이러한 방식으로, 표적 항원(들)에 대한 친화성과 같은 원하는 특징이 달성되도록 부모 및 참조 서열로부터 잔기가 선택되고 조합될 수 있다. 대안적으로, 또는 상기 절차에 부가적으로, 다른 적합한 조작 방법이 사용될 수 있다.

[0303] 단백질 스캐폴드를 유사한 단백질 또는 단편에 대한 특이적 결합에 대해 스크리닝하는 것은 뉴클레오타이드(DNA 또는 RNA 디스플레이) 또는 펩타이드 디스플레이 라이브러리, 예를 들어, 시험관내 디스플레이를 사용하여 편리하게 달성될 수 있다. 이러한 방법은 원하는 기능 또는 구조를 갖는 개별 구성원에 대한 많은 펩타이드 집합의 스크리닝을 포함한다. 디스플레이된 뉴클레오타이드 또는 펩타이드 서열은 3 내지 5000개 이상의 뉴클레오타이드 또는 아미노산 길이, 빈번하게는 5-100개의 아미노산 길이, 및 종종 약 8 내지 25개의 아미노산 길이일 수 있다. 펩타이드 라이브러리를 생성하기 위한 직접적인 화학적 합성 방법 외에도, 몇 가지 재조합 DNA 방법이 기재되었다. 한 가지 유형은 박테리오파아지 또는 세포의 표면 상에 펩타이드 서열의 디스플레이를 포함한다. 각 박테리오파아지 또는 세포는 특정 디스플레이된 펩타이드 서열을 코딩하는 뉴클레오타이드 서열을 함유한다. 이러한 방법은 PCT 특허 공개 WO 제91/17271호, WO 제91/18980호, WO 제91/19818호, 및 WO 제93/08278호에 기재되어 있다.

[0304] 펩타이드의 라이브러리를 생성하기 위한 다른 시스템은 시험관내 화학적 합성 및 재조합 방법 모두의 양태를 갖는다. PCT 특허 공개 WO 제92/05258호, WO 제92/14843호, 및 WO 제96/19256호를 참고한다. 또한, 미국 특허 제5,658,754호; 및 제5,643,768호를 참고한다. 펩타이드 디스플레이 라이브러리, 벡터, 및 스크리닝 키트는 Invitrogen(Carlsbad, Calif.), 및 Cambridge Antibody Technologies(Cambridgeshire, UK)와 같은 공급업체로부터 상업적으로 이용가능하다. 예컨대, Enzon에 양도된 미국 특허 제4,704,692호, 제4,939,666호, 제4,946,778호, 제5,260,203호, 제5,455,030호, 제5,518,889호, 제5,534,621호, 제5,656,730호, 제5,763,733호, 제5,767,260호, 제5856456호; Dyax에 양도된 제5,223,409호, 제5,403,484호, 제5,571,698호, 제5,837,500호,

Affymax에 양도된 제5,427,908호, 제5,580,717호; Cambridge Antibody Technologies에 양도된 제5,885,793호; Genentech에 양도된 제5,750,373호, 상기 Xoma, Colligan에 양도된 제5,618,920호, 제5,595,898호, 제5,576,195호, 제5,698,435호, 제5,693,493호, 제5,698,417호; 상기 Ausubel; 또는 상기 Sambrook을 참고한다.

[0305] 본 개시내용의 단백질 스캐폴드는 광범위한 친화성(KD)으로 인간 또는 다른 포유동물 단백질에 결합할 수 있다. 바람직한 양태에서, 본 개시내용의 적어도 하나의 단백질 스캐폴드는 높은 친화성으로, 예를 들어, 당업자에 의해 실시되는 바와 같이 표면 플라즈몬 공명 또는 Kinexa 방법에 의해 결정된 바와 같이, 약 $10^{-7}$ M 이하의 KD, 예컨대 비제한적으로, 0.1-9.9(또는 이 안의 임의의 범위 또는 값) X $10^{-8}$, $10^{-9}$, $10^{-10}$, $10^{-11}$, $10^{-12}$, $10^{-13}$, $10^{-14}$, $10^{-15}$ 또는 이 안의 임의의 범위 또는 값으로 표적 단백질에 선택적으로 결합할 수 있다.

[0306] 항원에 대한 단백질 스캐폴드의 친화성 또는 결합력은 임의의 적합한 방법을 사용하여 실험적으로 결정될 수 있다(예를 들어, Berzofsky, et al., "Antibody-Antigen Interactions," In Fundamental Immunology, Paul, W. E., Ed., Raven Press: New York, N.Y. (1984); Kuby, Janis Immunology, W.H. Freeman and Company: New York, N.Y. (1992); 및 본원에 기재된 방법 참조). 특정 단백질 스캐폴드-항원 상호작용의 측정된 친화성은 상이한 조건(예컨대, 염 농도, pH) 하에 측정되는 경우 달라질 수 있다. 따라서, 친화성 및 다른 항원-결합 매개변수(예컨대, KD, Kon, Koff)의 측정은 바람직하게는 단백질 스캐폴드 및 항원의 표준화된 용액, 및 본원에 기재된 완충제와 같은 표준화된 완충제로 이루어진다.

[0307] 경쟁 분석은 어떤 단백질, 항체, 및 다른 길항제가 표적 단백질에의 결합에 대해 단백질 스캐폴드와 경쟁하는지 및/또는 에피토프 영역을 공유하는지 결정하기 위해 단백질 스캐폴드를 이용하여 수행될 수 있다. 당업자에게 쉽게 공지된 이러한 분석은 단백질 상의 제한된 수의 결합 부위에 대한 길항제 또는 리간드 사이의 경쟁을 평가한다. 단백질 및/또는 항체는 경쟁 전 또는 후에 고정되거나 불용화되고, 표적 단백질에 결합된 샘플은, 예를 들어, 경사분리(decanting)에 의해(단백질/항체가 사전에 불용화된 경우) 또는 원심분리에 의해(단백질/항체가 경쟁적 반응 후에 침전된 경우) 결합되지 않은 샘플로부터 분리된다. 또한, 경쟁적 결합은 표적 단백질에 대한 단백질 스캐폴드의 결합 또는 결합 부족에 의해 기능이 변경되는지 여부, 예컨대, 단백질 스캐폴드가, 예를 들어, 표지의 효소 활성을 억제하거나 강화하는지 여부에 의해 결정될 수 있다. ELISA 및 다른 기능 분석이 당업계에 잘 알려진 바와 같이 사용될 수 있다.

[0308] 핵산 분자

[0309] 단백질 스캐폴드를 코딩하는 본 개시내용의 핵산 분자는 mRNA, hnRNA, tRNA 또는 임의의 다른 형태와 같은 RNA의 형태, 또는 비제한적으로, 클로닝에 의해 수득되거나 합성적으로 생산된 cDNA 및 게놈 DNA를 포함하는 DNA의 형태, 또는 임의의 조합일 수 있다. DNA는 삼중 가닥, 이중 가닥 또는 단일 가닥, 또는 이의 임의의 조합일 수 있다. DNA 또는 RNA의 적어도 하나의 가닥의 임의의 부분은 센스 가닥으로도 알려진 코딩 가닥일 수 있거나, 안티센스 가닥으로도 지칭되는 비코딩 가닥일 수 있다.

[0310] 본 개시내용의 단리된 핵산 분자는 선택적으로, 하나 이상의 인트론, 예컨대, 비제한적으로, 적어도 하나의 단백질 스캐폴드의 적어도 하나의 특정 부분과 함께, 오픈 리딩 프레임(ORF)을 포함하는 핵산 분자; 표적 단백질에 결합하는 단백질 스캐폴드 또는 루프 영역에 대한 코딩 서열을 포함하는 핵산 분자; 및 유전적 코드의 축퇴성으로 인해, 본원에 기재된 바와 같은 및/또는 당업계에 공지된 바와 같은 단백질 스캐폴드를 여전히 코딩하는 상기 기재된 것과 실질적으로 상이한 뉴클레오타이드 서열을 포함하는 핵산 분자를 포함할 수 있다. 물론, 유전적 코드는 당업계에 잘 알려져 있다. 따라서, 본 개시내용의 특정 단백질 스캐폴드를 코딩하는 이러한 축퇴성 핵산 변이체를 생성하는 것은 당업자에게 일상적일 것이다. 예컨대, 상기 문헌[Ausubel, et al.]을 참고하며, 이러한 핵산 변이체는 본 개시내용에 포함된다.

[0311] 본원에 표시된 바와 같이, 단백질 스캐폴드를 코딩하는 핵산을 포함하는 본 개시내용의 핵산 분자는, 비제한적으로, 그 자체로 단백질 스캐폴드 단편의 아미노산 서열을 코딩하는 것; 전체 단백질 스캐폴드 또는 이의 일부에 대한 코딩 서열; 단백질 스캐폴드, 단편 또는 일부에 대한 코딩 서열뿐만 아니라 전사, 스플라이싱을 포함하는 mRNA 가공 및 폴리아데닐화 신호(예를 들어, mRNA의 리보솜 결합 및 안정성)에서 역할을 하는 전사된 비번역된 서열과 같은 비코딩 5' 및 3' 서열을 비제한적으로 포함하는 추가의 비코딩 서열과 함께, 적어도 하나의 인트론과 같은 전술한 추가의 코딩 서열을 갖거나 갖지 않는 적어도 하나의 신호 리더 또는 융합 펩타이드에 대한 코딩 서열과 같은 추가의 서열; 추가의 기능을 제공하는 것과 같은 추가의 아미노산을 코딩하는 추가의 코딩 서열을 포함할 수 있다. 따라서, 단백질 스캐폴드를 코딩하는 서열은 마커 서열, 예컨대 단백질 스캐폴드 단편 또

는 일부를 포함하는 융합된 단백질 스캐폴드의 정제를 촉진하는 펩타이드를 코딩하는 서열에 융합될 수 있다.

[0312] 본원에 기재된 바와 같은 폴리뉴클레오타이드에 선택적으로 혼성화하는 폴리뉴클레오타이드

[0313] 본 개시내용은 선택적 혼성화 조건 하에서 본원에 개시된 폴리뉴클레오타이드에 혼성화하는 단리된 핵산을 제공한다. 따라서, 폴리뉴클레오타이드는 이러한 폴리뉴클레오타이드를 포함하는 핵산을 단리, 검출, 및/또는 정량화하는 데 사용될 수 있다. 예를 들어, 본 개시내용의 폴리뉴클레오타이드는 기탁된 라이브러리에서 부분적인 또는 전장 클론을 확인, 단리 또는 증폭하는 데 사용될 수 있다. 폴리뉴클레오타이드는 인간 또는 포유동물 핵산 라이브러리로부터 단리되거나 이에 상보적인 게놈 또는 cDNA 서열일 수 있다.

[0314] 바람직하게는, cDNA 라이브러리는 적어도 80% 전장 서열, 바람직하게는, 적어도 85% 또는 90% 전장 서열, 및, 더욱 바람직하게는, 적어도 95% 전장 서열을 포함한다. cDNA 라이브러리는 희귀 서열의 표현을 증가시키기 위해 정규화될 수 있다. 낮거나 중간 정도의 엄격성 혼성화 조건은 전형적으로 상보적인 서열에 비해 감소된 서열 동일성을 갖는 서열과 함께 사용되지만 이에 국한되지 않는다. 중간 정도 및 높은 엄격성 조건은 더 큰 동일성의 서열에 대해 선택적으로 사용될 수 있다. 낮은 엄격성 조건은 약 70% 서열 동일성을 갖는 서열의 선택적 혼성화를 허용하며 이종상동성(orthologous) 또는 유사(paralogous) 서열을 확인하는 데 사용될 수 있다.

[0315] 선택적으로, 폴리뉴클레오타이드는 본원에 기재된 폴리뉴클레오타이드에 의해 코딩된 단백질 스캐폴드의 적어도 일부를 코딩할 것이다. 폴리뉴클레오타이드는 본 개시내용의 단백질 스캐폴드를 코딩하는 폴리뉴클레오타이드에 대한 선택적 혼성화에 사용될 수 있는 핵산 서열을 포함한다. 예컨대, 상기 Ausubel; 상기 Colligan을 참고하며, 각각은 전체가 참조로 본원에 포함되어 있다.

[0316] 핵산의 제작

[0317] 본 개시내용의 단리된 핵산은 당업계에 널리 공지된 바와 같이 (a) 재조합 방법, (b) 합성 기술, (c) 정제 기술, 및/또는 (d) 이의 조합을 사용하여 제조될 수 있다.

[0318] 핵산은 본 개시내용의 폴리뉴클레오타이드 외의 서열을 편리하게 포함할 수 있다. 예를 들어, 하나 이상의 엔도뉴클레아제 제한 부위를 포함하는 다중 클로닝 부위는 폴리뉴클레오타이드의 단리를 돕기 위해 핵산 내로 삽입될 수 있다. 또한, 번역가능한 서열은 본 개시내용의 번역된 폴리뉴클레오타이드의 단리를 돕기 위해 삽입될 수 있다. 예를 들어, 헥사-히스티딘 마커 서열은 본 개시내용의 단백질을 정제하기 위한 편리한 수단을 제공한다. 코딩 서열을 제외한 본 개시내용의 핵산은 선택적으로 본 개시내용의 폴리뉴클레오타이드의 클로닝 및/또는 발현을 위한 벡터, 어댑터, 또는 링커이다.

[0319] 클로닝 및/또는 발현에서 기능을 최적화하기 위해, 폴리뉴클레오타이드의 단리를 돕기 위해, 또는 폴리뉴클레오타이드의 세포 내로의 도입을 개선하기 위해 추가의 서열이 이러한 클로닝 및/또는 발현 서열에 첨가될 수 있다. 클로닝 벡터, 발현 벡터, 어댑터, 및 링커의 사용은 당업계에 널리 알려져 있다(예컨대, 상기 Ausubel; 또는 상기 Sambrook 참고).

[0320] 핵산을 제작하기 위한 재조합 방법

[0321] 본 개시내용의 단리된 핵산 조성물, 예컨대 RNA, cDNA, 게놈 DNA, 또는 이의 임의의 조합은 당업자에게 공지된 많은 클로닝 방법을 사용하여 생물학적 공급원으로부터 수득될 수 있다. 일부 양태에서, 엄격한 조건 하에 본 개시내용의 폴리뉴클레오타이드에 선택적으로 혼성화하는 올리고뉴클레오타이드 프로브가 cDNA 또는 게놈 DNA 라이브러리에서 원하는 서열을 확인하는 데 사용된다. RNA의 단리, 및 cDNA 및 게놈 라이브러리의 제작은 당업자에게 널리 알려져 있다(예컨대, 상기 Ausubel; 또는 상기 Sambrook 참고).

[0322] 핵산 스크리닝 및 단리 방법

[0323] cDNA 또는 게놈 라이브러리는 본 개시내용의 폴리뉴클레오타이드의 서열에 기초한 프로브를 사용하여 스크리닝될 수 있다. 프로브는 게놈 DNA 또는 cDNA 서열과 혼성화하여 동일하거나 상이한 유기체에서 상동 유전자를 단리하는 데 사용될 수 있다. 당업자는 다양한 정도의 혼성화 엄격성이 분석에서 사용될 수 있음을 이해할 것이며; 혼성화 또는 세척 매질은 엄격할 수 있다. 혼성화 조건이 더욱 엄격해짐에 따라, 이합체 형성이 일어나기 위해서는 프로브 및 표적 사이에 더 큰 상보성 정도가 있어야 한다. 엄격성 정도는 온도, 이온 강도, pH 및 포름아미드와 같은 부분적으로 변성하는 용매의 존재 중 하나 이상에 의해 제어될 수 있다. 예를 들어, 혼성화 엄격성은, 예를 들어, 포름아미드의 농도를 0% 내지 50%의 범위 내로 조작하는 것을 통해 반응 용액의 극성을 변경함으로서 편리하게 달라진다. 검출가능한 결합에 필요한 상보성(서열 동일성)의 정도는 혼성화 매질 및/또는 세척 매질의 엄격성에 따라 달라질 것이다. 상보성 정도는 최적으로 100%, 또는 70-100%, 또는 이 안의 임의

의 범위 또는 값일 것이다. 그러나, 프로브 및 프라이머 내의 사소한 서열 변이는 혼성화 및/또는 세척 매질의 엄격성을 감소시킴으로써 보상될 수 있음을 이해해야 한다.

[0324] RNA 또는 DNA의 증폭 방법은 당업계에 널리 알려져 있으며, 본원에 제시된 교시 및 지침에 기초하여 과도한 실험 없이 본 개시내용에 따라 사용될 수 있다.

[0325] DNA 또는 RNA 증폭의 공지된 방법은, 비제한적으로, 중합효소 연쇄 반응(PCR) 및 관련 증폭 과정(예컨대, Mullis, et al의 미국 특허 제4,683,195호, 제4,683,202호, 제4,800,159호, 제4,965,188호; Tabor, et al의 제4,795,699호 및 제4,921,794호; Innis의 제5,142,033호; Wilson, et al의 제5,122,464호; Innis의 제5,091,310호; Gyllensten, et al의 제5,066,584호; Gelfand, et al의 제4,889,818호; Silver, et al의 제4,994,370호; Biswas의 제4,766,067호; Ringold의 제4,656,134호 참고) 및 이중 가닥의 DNA 합성을 위한 주형으로서 표적 서열에 대한 안티센스 RNA를 사용하는 RNA 매개 증폭(Malek, et al의 미국 특허 제5,130,238호, 상표명 NASBA를 가짐)을 포함하며, 이들 문헌의 전체 내용이 참조로 본원에 포함된다(예컨대, 상기 Ausubel; 또는 상기 Sambrook 참고)

[0326] 예를 들어, 중합효소 연쇄 반응(PCR) 기술은 게놈 DNA 또는 cDNA 라이브러리로부터 직접 본 개시내용의 폴리뉴클레오타이드 및 관련 유전자의 서열을 증폭하는 데 사용될 수 있다. PCR 및 다른 시험관내 증폭 방법은 또한, 예를 들어, 발현될 단백질을 코딩하는 핵산 서열을 클로닝하는데, 샘플에서 원하는 mRNA의 존재를 검출하기 위한 프로브로서 사용하기 위한 핵산을 제조하는데, 핵산 시퀀싱을 위해, 또는 다른 목적을 위해 유용할 수 있다. 시험관내 증폭 방법을 통해 숙련자를 지시하기에 충분한 기술의 예는 문헌[상기 Berger, 상기 Sambrook, 및 상기 Ausubel 뿐만 아니라 Mullis, et al., 미국 특허 제4,683,202 (1987); 및 Innis, et al., PCR Protocols A Guide to Methods and Applications, Eds., Academic Press Inc., San Diego, Calif.(1990)]에서 발견된다. 게놈 PCR 증폭을 위한 상업적으로 이용가능한 키트는 당업계에 알려져 있다. 예컨대, Advantage-GC 게놈 PCR 키트(Clontech)를 참고한다. 추가로, 예컨대, T4 유전자 32 단백질(Boehringer Mannheim)이 긴 PCR 생성물의 수율을 개선하는 데 사용될 수 있다.

[0327] 핵산을 제작하기 위한 합성 방법

[0328] 본 개시내용의 단리된 핵산은 또한 공지된 방법에 의해 직접 화학적 합성에 의해 제조될 수 있다(예컨대, 상기 Ausubel, et al 참고). 화학적 합성은 일반적으로 단일 가닥의 올리고뉴클레오타이드를 생산하며, 이는 상보적 서열과의 혼성화에 의해, 또는 주형으로서 단일 가닥을 사용하는 DNA 중합효소를 이용한 중합에 의해 이중 가닥의 DNA로 전환될 수 있다. 당업자는 DNA의 화학적 합성이 약 100개 이상의 염기의 서열로 제한될 수 있지만, 더 긴 서열이 더 짧은 서열의 결찰에 의해 수득될 수 있음을 인식할 것이다.

[0329] 재조합 발현 카세트

[0330] 본 개시내용은 본 개시내용의 핵산을 포함하는 재조합 발현 카세트를 추가로 제공한다. 본 개시내용의 핵산 서열, 예를 들어, 본 개시내용의 단백질 스캐폴드를 코딩하는 cDNA 또는 게놈 서열은 적어도 하나의 원하는 숙주 세포 내로 도입될 수 있는 재조합 발현 카세트를 제작하는 데 사용될 수 있다. 재조합 발현 카세트는 전형적으로 의도된 숙주 세포에서 폴리뉴클레오타이드의 전사를 지시할 전사 개시 조절 서열에 작동가능하게 연결된 본 개시내용의 폴리뉴클레오타이드를 포함할 것이다. 이종 및 비이종(즉, 내인성) 프로모터 모두가 본 개시내용의 핵산의 발현을 지시하는 데 사용될 수 있다.

[0331] 일부 양태에서, 프로모터, 인핸서, 또는 다른 요소로서 작용하는 단리된 핵산이 본 개시내용의 폴리뉴클레오타이드의 발현을 상향 또는 하향 조절하기 위해 본 개시내용의 폴리뉴클레오타이드의 비이종 형태의 적절한 위치에(상류, 하류 또는 인트론 내에) 도입될 수 있다. 예를 들어, 내인성 프로모터는 돌연변이, 결실 및/또는 치환에 의해 생체내 또는 시험관내에서 변경될 수 있다.

[0332] 발현 벡터 및 숙주 세포

[0333] 본 개시내용은 또한 본 개시내용의 단리된 핵산 분자를 포함하는 벡터, 재조합 벡터로 유전적으로 조작된 숙주 세포, 및 당업계에 널리 공지된 바와 같은 재조합 기술에 의한 적어도 하나의 단백질 스캐폴드의 생산에 관한 것이다. 예컨대, 상기 Sambrook, et al; 상기 Ausubel, et al을 참고하며, 이들 각각은 전체가 참고로 본원에 포함되어 있다.

[0334] 폴리뉴클레오타이드는 숙주에서 증식을 위한 선택 마커를 함유하는 벡터에 선택적으로 연결될 수 있다. 일반적으로, 플라스미드 벡터는 인산 칼슘 침전물과 같은 침전물에, 또는 하전된 지질과의 복합체에 도입된다. 벡터가

바이러스인 경우, 그것은 적절한 패키징 세포주를 사용하여 시험관내에서 패키징된 다음 숙주 세포 내로 형질도입될 수 있다.

[0335] DNA 삽입물은 적절한 프로모터에 작동가능하게 연결되어야 한다. 발현 구조체는 전사 개시, 종결을 위한 부위 및, 전사된 영역에서, 번역을 위한 리보솜 결합 부위를 추가로 함유할 것이다. 구조체에 의해 발현되는 성숙한 전사체의 코딩 부분은 바람직하게는 시작에서 번역 개시 및 번역될 mRNA의 말단에 적절하게 위치한 종결 코돈(예컨대, UAA, UGA 또는 UAG)을 포함할 것이며, UAA 및 UAG가 포유동물 또는 진핵 세포 발현에 바람직하다.

[0336] 발현 벡터는 바람직하게는 그러나 선택적으로 적어도 하나의 선택 마커를 포함할 것이다. 이러한 마커는, 예컨대, 비제한적으로, 암피실린, 제오신(*Sh ble* 유전자), 푸로마이신(*pac* 유전자), 하이그로마이신 B(*hygB* 유전자), G418/제네티신(*neo* 유전자), DHFR(디하이드로폴레이트 환원효소를 코딩하고 메토트렉세이트에 대한 내성을 부여함), 미코페놀산, 또는 글루타민 합성효소(GS, 미국 특허 제5,122,464호; 제5,770,359호; 제5,827,739호), 블라스티시딘(*bsd* 유전자), 진핵 세포 배양에 대한 내성 유전자뿐만 아니라 암피실린, 제오신(*Sh ble* 유전자), 푸로마이신(*pac* 유전자), 하이그로마이신 B(*hygB* 유전자), G418/제네티신(*neo* 유전자), 카나마이신, 스펙티노마이신, 스트렙토마이신, 카르베니실린, 블레오마이신, 에리트로마이신, 폴리믹신 B, 또는 E. coli 및 다른 박테리아 또는 원핵생물에서 배양하기 위한 테트라사이클린 내성 유전자를 포함한다(상기 특허는 전체적으로 참조로 본원에 포함됨). 전술한 숙주 세포를 위한 적절한 배양 배지 및 조건은 당업계에 공지되어 있다. 적합한 벡터는 당업자에게 쉽게 명백할 것이다. 숙주 세포 내로 벡터 구조체의 도입은 인산 칼슘 형질감염, DEAE-덱스트란 매개 형질감염, 양이온성 지질 매개 형질감염, 전기천공, 형질도입, 감염 또는 다른 공지된 방법에 의해 수행될 수 있다. 이러한 방법은 상기 Sambrook, 1-4 및 16-18장; 상기 Ausubel, 1, 9, 13, 15, 16장과 같이 당업계에 기재되어 있다.

[0337] 발현 벡터는 바람직하게는 그러나 선택적으로 본 개시내용의 조성물 및 방법에 의해 변형된 세포의 단리를 위한 적어도 하나의 선택가능한 세포 표면 마커를 포함할 것이다. 본 개시내용의 선택가능한 세포 표면 마커는 세포 또는 세포의 서브세트를 세포의 또 다른 정의된 서브세트와 구별하는 표면 단백질, 당단백질, 또는 단백질의 그룹을 포함한다. 바람직하게는 선택가능한 세포 표면 마커는 본 개시내용의 조성물 또는 방법에 의해 변형된 세포를 본 개시내용의 조성물 또는 방법에 의해 변형되지 않은 세포와 구별한다. 이러한 세포 표면 마커는, 예컨대, 비제한적으로, CD19, CD271, CD34, CD22, CD20, CD33, CD52, 또는 임의의 이의 조합의 절단된 또는 전장 형태와 같은 "지정 클러스터(cluster of designation)" 또는 "분류 결정자(classification determinant)" 단백질(종종 "CD"로 약칭됨)을 포함한다. 세포 표면 마커는 자살 유전자 마커 RQR8을 추가로 포함한다(Philip B et al. Blood. 2014 Aug 21; 124(8):1277-87).

[0338] 발현 벡터는 바람직하게는 그러나 선택적으로 본 개시내용의 조성물 및 방법에 의해 변형된 세포의 단리를 위한 적어도 하나의 선택가능한 약물 내성 마커를 포함할 것이다. 본 개시내용의 선택가능한 약물 내성 마커는 야생형 또는 돌연변이체 Neo, DHFR, TYMS, FRANCF, RAD51C, GCS, MDR1, ALDH1, NKX2.2, 또는 이의 임의의 조합을 포함할 수 있다.

[0339] 본 개시내용의 적어도 하나의 단백질 스캐폴드는 융합 단백질과 같은 변형된 형태로 발현될 수 있고, 분비 신호뿐만 아니라 추가의 이종 기능 영역을 포함할 수 있다. 예를 들어, 추가의 아미노산, 특히 하전된 아미노산의 영역은 정제 동안, 또는 후속 취급 및 저장 동안 숙주 세포에서 안정성 및 지속성을 개선하기 위해 단백질 스캐폴드의 N-말단에 추가될 수 있다. 또한, 펩타이드 모이어티는 정제를 촉진하기 위해 본 개시내용의 단백질 스캐폴드에 첨가될 수 있다. 이러한 영역은 단백질 스캐폴드 또는 이의 적어도 하나의 단편의 최종 제조 전에 제거될 수 있다. 이러한 방법은 상기 Sambrook, 17.29-17.42 및 18.1-18.74장; 상기 Ausubel, 16, 17 및 18장과 같이 많은 표준 실험실 매뉴얼에 기재되어 있다.

[0340] 당업자는 본 개시내용의 단백질을 코딩하는 핵산의 발현에 이용가능한 수많은 발현 시스템을 잘 안다. 대안적으로, 본 개시내용의 핵산은 본 개시내용의 단백질 스캐폴드를 코딩하는 내인성 DNA를 함유하는 숙주 세포에서 켜짐으로써(조작에 의해) 숙주 세포에서 발현될 수 있다. 이러한 방법은, 예컨대, 그 전체가 참조로 본원에 포함되어 있는 미국 특허 제5,580,734호, 제5,641,670호, 제5,733,746호, 및 제5,733,761호에 기재된 바와 같이 당업계에 널리 알려져 있다.

[0341] 단백질 스캐폴드, 이의 특정 부분 또는 변이체의 생산에 유용한 세포 배양의 예시는 당업계에 알려진 바와 같은 박테리아, 효모, 및 포유동물 세포이다. 포유동물 세포 시스템은 종종 세포의 단층 형태일 것이지만 포유동물 세포 현탁액 또는 생물반응기가 또한 사용될 수 있다. 온전한 글리코실화된 단백질을 발현할 수 있는 많은 적합한 숙주 세포주가 당업계에서 개발되었고, 이는 COS-1(예컨대, ATCC CRL 1650), COS-7(예컨대, ATCC CRL-

1651), HEK293, BHK21(예컨대, ATCC CRL-10), CHO(예컨대, ATCC CRL 1610) 및 BSC-1(예컨대, ATCC CRL-26) 세포주, Cos-7 세포, CHO 세포, hep G2 세포, P3X63Ag8.653, SP2/0-Ag14, 293 세포, HeLa 세포 등을 포함하며, 이는 예를 들어, American Type Culture Collection, Manassas, Va.(www.atcc.org)로부터 쉽게 이용가능하다. 바람직한 숙주 세포는 림프계 기원의 세포, 예컨대 골수종 및 림프종 세포를 포함한다. 특히 바람직한 숙주 세포는 P3X63Ag8.653 세포(ATCC 등록번호 CRL-1580) 및 SP2/0-Ag14 세포(ATCC 등록번호 CRL-1851)이다. 바람직한 양태에서, 재조합 세포는 P3X63Ab8.653 또는 SP2/0-Ag14 세포이다.

[0342] 이들 세포를 위한 발현 벡터는 하기 발현 제어 서열, 예컨대, 비제한적으로, 복제 원점; 프로모터(예컨대, 후기 또는 초기 SV40 프로모터, CMV 프로모터(미국 특허 제5,168,062호; 제5,385,839호), HSV tk 프로모터, pgk(phosphoglycerate kinase) 프로모터, EF-1 알파 프로모터(미국 특허 제5,266,491호), 적어도 하나의 인간 프로모터; 인핸서, 및/또는 가공 정보 부위, 예컨대 리보솜 결합 부위, RNA 스플라이스 부위, 폴리아데닐화 부위(예컨대, SV40 대형 T Ag 폴리 A 부가 부위), 및 전사 종결자 서열 중 하나 이상을 포함할 수 있다. 예컨대, 상기 Ausubel et al; 상기 Sambrook, et al을 참고한다. 본 개시내용의 핵산 또는 단백질의 생산에 유용한 다른 세포는 공지되어 있고/거나, 예를 들어, American Type Culture Collection Catalogue of Cell Lines and Hybridomas(www.atcc.org) 또는 다른 공지된 또는 상업적인 출처로부터 이용가능하다.

[0343] 진핵 숙주 세포가 사용되는 경우, 폴리아데닐화 또는 전사 종결자 서열이 전형적으로 벡터 내로 혼입된다. 종결자 서열의 예는 소 성장 호르몬 유전자로부터의 폴리아데닐화 서열이다. 전사체의 정확한 스플라이싱을 위한 서열도 포함될 수 있다. 스플라이싱 서열의 예는 SV40으로부터의 VP1 인트론이다(Sprague, et al., J. Virol. 45:773-781(1983)). 추가로, 숙주 세포에서 복제를 제어하기 위한 유전자 서열이 당업계에 공지된 바와 같이 벡터 내로 혼입될 수 있다.

[0344] 단백질 스캐폴드 정제

[0345] 단백질 스캐폴드는, 비제한적으로, 단백질 A 정제, 황산 암모늄 또는 에탄올 침전, 산 추출, 음이온 또는 양이온 교환 크로마토그래피, 포스포셀룰로오스 크로마토그래피, 소수성 상호작용 크로마토그래피, 친화성 크로마토그래피, 하이드록실아파타이트 크로마토그래피 및 렉틴 크로마토그래피를 포함하는 잘 알려진 방법에 의해 재조합 세포 배양물로부터 회수 및 정제될 수 있다. 고성능 액체 크로마토그래피("HPLC")가 또한 정제에 사용될 수 있다. 예컨대, 각각 전체가 참조로 본원에 포함된 문헌[Colligan, Current Protocols in Immunology, or Current Protocols in Protein Science, John Wiley & Sons, NY, N.Y., (1997-2001), 예컨대, 1, 4, 6, 8, 9, 10장]을 참고한다.

[0346] 본 개시내용의 단백질 스캐폴드는 정제된 생성물, 화학적 합성 절차의 생성물, 및, 예를 들어, E. coli, 효모, 고등 식물, 곤충 및 포유동물 세포를 포함하는 원핵 또는 진핵 숙주로부터 재조합 기술에 의해 생산된 생성물을 포함한다. 재조합 생산 절차에 사용된 숙주 세포에 따라, 본 개시내용의 단백질 스캐폴드는 글리코실화될 수 있거나 글리코실화되지 않을 수 있다. 이러한 방법은 많은 표준 실험실 매뉴얼, 예컨대 상기 Sambrook, 섹션 17.37-17.42; 상기 Ausubel, 10, 12, 13, 16, 18 및 20장, 상기 Colligan, Protein Science, 12-14장에 기재되어 있으며, 이들 모두는 그 전체가 참조로 본원에 포함되어 있다.

[0347] 아미노산 코드

[0348] 본 개시내용의 단백질 스캐폴드를 구성하는 아미노산은 종종 약칭된다. 아미노산 지정은 당업계에서 널리 이해되는 바와 같이 아미노산을 그의 단일 글자 코드, 그의 세 글자 코드, 이름, 또는 3개의 뉴클레오타이드 코돈(들)에 의해 지정함으로써 표시될 수 있다(Alberts, B., et al., Molecular Biology of The Cell, Third Ed., Garland Publishing, Inc., New York, 1994 참고). 본 개시내용의 단백질 스캐폴드는 본원에 명시된 바와 같이 자발적 또는 돌연변이 및/또는 인간 조작으로부터의 하나 이상의 아미노산 치환, 결실 또는 부가를 포함할 수 있다. 기능에 필수적인 본 개시내용의 단백질 스캐폴드 내의 아미노산은 부위 지정 돌연변이유발 또는 알라닌 스캐닝 돌연변이유발과 같은 당업계에 공지된 방법에 의해 확인될 수 있다(예컨대, 상기 Ausubel, 8, 15장; Cunningham and Wells, Science 244:1081-1085(1989)). 후자의 절차는 분자 내의 모든 잔기에 단일 알라닌 돌연변이를 도입한다. 이후, 생성된 돌연변이체 분자는, 비제한적으로, 적어도 하나의 중화 활성과 같은 생물학적 활성에 대해 시험된다. 단백질 스캐폴드 결합에 중요한 부위는 또한 결정화, 핵 자기 공명 또는 광친화성 표지화와 같은 구조적 분석에 의해 확인될 수 있다(Smith, et al., J. Mol. Biol. 224:899-904(1992) 및 de Vos, et al., Science 255:306-312(1992)).

[0349] 당업자가 인식하는 바와 같이, 본 개시내용은 본 개시내용의 적어도 하나의 생물학적 활성 단백질 스캐폴드를

포함한다. 생물학적 활성 단백질 스캐폴드는 천연(비합성), 내인성 또는 관련되고 알려진 단백질 스캐폴드의 비활성의 적어도 20%, 30%, 또는 40%, 및, 바람직하게는, 적어도 50%, 60%, 또는 70%, 및, 가장 바람직하게는, 적어도 80%, 90%, 또는 95%-99% 이상의 비활성을 갖는다. 효소 활성 및 기질 특이성의 측정값을 분석하고 정량화하는 방법은 당업자에게 잘 알려져 있다.

[0350] 또 다른 양태에서, 본 개시내용은 유기 모이어티의 공유 부착에 의해 변형된 본원에 기재된 바와 같은 단백질 스캐폴드 및 단편에 관한 것이다. 이러한 변형은 개선된 약동학적 특성(예컨대, 증가된 생체내 혈청 반감기)를 갖는 단백질 스캐폴드 단편을 생산할 수 있다. 유기 모이어티는 선형 또는 분지형 친수성 중합체성 그룹, 지방산 그룹, 또는 지방산 에스테르 그룹일 수 있다. 특정 양태에서, 친수성 중합체성 그룹은 약 800 내지 약 120,000 달톤의 분자량을 가질 수 있고, 폴리알칸 글리콜(예컨대, 폴리에틸렌 글리콜(PEG), 폴리프로필렌 글리콜(PPG)), 탄수화물 중합체, 아미노산 중합체 또는 폴리비닐 피롤리돈일 수 있으며, 지방산 또는 지방산 에스테르 그룹은 약 8 내지 약 40개의 탄소 원자를 포함할 수 있다.

[0351] 본 개시내용의 변형된 단백질 스캐폴드 및 단편은 항체에 직접 또는 간접적으로 공유 결합된 하나 이상의 유기 모이어티를 포함할 수 있다. 본 개시내용의 단백질 스캐폴드 또는 단편에 결합된 각각의 유기 모이어티는 독립적으로 친수성 중합체성 그룹, 지방산 그룹 또는 지방산 에스테르 그룹일 수 있다. 본원에 사용된 바와 같이, 용어 "지방산"은 모노-카르복실산 및 디-카르복실산을 포함한다. 본원에 사용된 바와 같이, "친수성 중합체성 그룹"은 옥탄보다 물에서 더 잘 용해되는 유기 중합체를 지칭한다. 예를 들어, 폴리리신은 옥탄보다 물에서 더 잘 용해된다. 따라서, 폴리리신의 공유 부착에 의해 변형된 단백질 스캐폴드는 본 개시내용에 의해 포함된다. 본 개시내용의 단백질 스캐폴드를 변형시키는 데 적합한 친수성 중합체는 선형 또는 분지형일 수 있고, 예를 들어, 폴리알칸 글리콜(예컨대, PEG, 모노메톡시-폴리에틸렌 글리콜(mPEG), PPG 등), 탄수화물(예컨대, 덱스트란, 셀룰로오스, 올리고당, 다당류 등), 친수성 아미노산의 중합체(예컨대, 폴리리신, 폴리아르기닌, 폴리아스파르테이트 등), 폴리알칸 옥사이드(예컨대, 폴리에틸렌 옥사이드, 폴리프로필렌 옥사이드 등) 및 폴리비닐 피롤리돈을 포함한다. 바람직하게는, 본 개시내용의 단백질 스캐폴드를 변형하는 친수성 중합체는 별도의 분자 엔티티로서 약 800 내지 약 150,000 달톤의 분자량을 갖는다. 예를 들어, PEG5000 및 PEG20,000이 사용될 수 있으며, 여기서 아래 첨자는 중합체의 평균 분자량 달톤이다. 친수성 중합체성 그룹은 1 내지 약 6개의 알킬, 지방산 또는 지방산 에스테르 그룹으로 치환될 수 있다. 지방산 또는 지방산 에스테르 그룹으로 치환되는 친수성 중합체는 적합한 방법을 사용하여 제조될 수 있다. 예를 들어, 아민 그룹을 포함하는 중합체는 지방산 또는 지방산 에스테르의 카르복실레이트에 결합될 수 있고, 지방산 또는 지방산 에스테르 상의 활성화된 카르복실레이트(예컨대, N,N-카르보닐 디이미다졸로 활성화됨)는 중합체 상의 하이드록실 그룹에 결합될 수 있다.

[0352] 본 개시내용의 단백질 스캐폴드를 변형하는 데 적합한 지방산 및 지방산 에스테르는 포화될 수 있거나 하나 이상의 불포화 단위를 함유할 수 있다. 본 개시내용의 단백질 스캐폴드를 변형하는 데 적합한 지방산은, 예를 들어, n-도데카노에이트(C12, 라우레이트), n-테트라데카노에이트(C14, 미리스테이트), n-옥타데카노에이트(C18, 스테아레이트), n-에이코사노에이트(C20, 아라키데이트), n-도코사노에이트(C22, 베헤네이트), n-트리아콘타노에이트(C30), n-테트라콘타노에이트(C40), cis-△9-옥타데카노에이트(C18, 올레에이트), 모든 시스-△5,8,11,14-에이코사테트라에노에이트(C20, 아라키도네이트), 옥탄디오산, 테트라데칸디오산, 옥타데칸디오산, 도코산디오산 등을 포함한다. 적합한 지방산 에스테르는 선형 또는 분지형 저급 알킬 그룹을 포함하는 디카르복실산의 모노-에스테르를 포함한다. 저급 알킬 그룹은 1 내지 약 12, 바람직하게는, 1 내지 약 6개의 탄소 원자를 포함할 수 있다.

[0353] 변형된 단백질 스캐폴드 및 단편은 적합한 방법을 사용하여, 예컨대 하나 이상의 변형제와의 반응에 의해 제조될 수 있다. 본원에서 사용된 바와 같이 "변형제"는 활성화 그룹을 포함하는 적합한 유기 그룹(예컨대, 친수성 중합체, 지방산, 지방산 에스테르)을 지칭한다. "활성화 그룹"은 적절한 조건 하에 제2 화학 그룹과 반응하여 변형제 및 제2 화학 그룹 사이에 공유 결합을 형성할 수 있는 화학적 모이어티 또는 작용기이다. 예를 들어, 아민 반응성 활성화 그룹은 친전자성 그룹, 예컨대 토실레이트, 메실레이트, 할로(클로로, 브로모, 플루오로, 아이오도), N-하이드록시석신이미딜 에스테르(NHS) 등을 포함한다. 티올과 반응할 수 있는 활성화 그룹은, 예를 들어, 말레이미드, 아이오도아세틸, 아크릴로일, 피리딜 디설파이드, 5-티올-2-니트로벤조산 티올(TNB-티올) 등을 포함한다. 알데하이드 작용기는 아민 함유 또는 히드라지드 함유 분자에 결합될 수 있고, 아지드 그룹은 3가 포스포러스 그룹과 반응하여 포스포르아미데이트 또는 포스포르이미데이트 연결을 형성할 수 있다. 분자에 활성화 그룹을 도입하는 적합한 방법은 당업계에 알려져 있다(예를 들어, Hermanson, G. T., Bioconjugate Techniques, Academic Press: San Diego, Calif.(1996) 참고). 활성화 그룹은 유기 그룹(예컨대, 친수성 중합체, 지방산, 지방산 에스테르)에 직접 결합될 수 있거나, 또는 링커 모이어티, 예를 들어, 하나 이상의 탄소 원

자가 산소, 질소 또는 황과 같은 헤테로원자에 의해 대체될 수 있는 2가 C1-C12 그룹을 통해 결합될 수 있다. 적합한 링커 모이어티는, 예를 들어, 테트라에틸렌 글리콜, -(CH2)3-, -NH-(CH2)6-NH-, -(CH2)2-NH- 및 -CH2-O-CH2-CH2-O-CH2-CH2-O-CH-NH-를 포함한다. 링커 모이어티를 포함하는 변형제는, 예를 들어, 모노-Boc-알킬디아민(예컨대, 모노-Boc-에틸렌디아민, 모노-Boc-디아미노헥산)을 1-에틸-3-(3-디메틸아미노프로필) 카르보디이미드(EDC)의 존재 하에 지방산과 반응시켜 유리 아민 및 지방산 카르복실레이트 사이에 아미드 결합을 형성함으로써 생산될 수 있다. Boc 보호 그룹은 트리플루오로아세트산(TFA)으로 처리하여 기재된 바와 같이 또 다른 카르복실레이트에 결합될 수 있는 1차 아민을 노출시킴으로써 생성물로부터 제거될 수 있거나, 또는 말레익 무수물과 반응하여 생성된 생성물을 고리화하여 지방산의 활성화된 말레이미도 유도체를 생산할 수 있다(예를 들어, 이의 전체 교시가 본원에 참고로 포함된 Thompson, et al., WO 제92/16221호 참고)

[0354] 본 개시내용의 변형된 단백질 스캐폴드는 단백질 스캐폴드 또는 단편을 변형제와 반응시킴으로써 생산될 수 있다. 예를 들어, 유기 모이어티는 아민 반응성 변형제, 예를 들어, PEG의 NHS 에스테르를 사용함으로써 비부위특이적 방식으로 단백질 스캐폴드에 결합될 수 있다. 본 개시내용의 단백질 스캐폴드의 특정 부위에 결합된 유기 모이어티를 포함하는 변형된 단백질 스캐폴드 및 단편은 적합한 방법, 예컨대 역 단백질분해(Fisch et al., Bioconjugate Chem., 3:147-153(1992); Werlen et al., Bioconjugate Chem., 5:411-417(1994); Kumaran et al., Protein Sci. 6(10):2233-2241(1997); Itoh et al., Bioorg. Chem., 24(1): 59-68(1996); Capellas et al., Biotechnol. Bioeng., 56(4):456-463(1997)), 및 문헌[Hermanson, G. T., Bioconjugate Techniques, Academic Press: San Diego, Calif.(1996)]에 기재된 방법을 사용하여 제조될 수 있다.

[0355] **정의**

[0356] 본 개시내용 전반에 걸쳐 사용된 바와 같이, 단수 형태 "a," "and," 및 "the"는 문맥이 명백하게 달리 지시하지 않는 한 복수 지시 대상을 포함한다. 따라서, 예를 들어, "방법"에 대한 언급은 복수의 이러한 방법을 포함하고 "용량"에 대한 언급은 하나 이상의 용량 및 당업자에게 알려진 이의 등가물 등에 대한 언급을 포함한다.

[0357] 용어 "약" 또는 "대략"은 당업자에 의해 결정된 바와 같은 특정 값에 대한 허용가능한 오차 범위 내를 의미하며, 이는 값이 측정 또는 결정되는 방법, 예컨대, 측정 시스템의 한계에 부분적으로 의존한다. 예를 들어, "약"은 1 이상의 표준 편차 이내를 의미할 수 있다. 대안적으로, "약"은 주어진 값의 최대 20%, 또는 최대 10%, 또는 최대 5%, 또는 최대 1%의 범위를 의미할 수 있다. 대안적으로, 특히 생물학적 시스템 또는 과정과 관련하여, 용어는 값의 10배 이내, 바람직하게는 5배 이내, 및 더욱 바람직하게는 2배 이내를 의미할 수 있다. 특정 값이 출원 및 청구범위에 기재된 경우, 달리 언급되지 않는 한 특정 값에 대한 허용가능한 오차 범위 내를 의미하는 용어 "약"이 추정되어야 한다.

[0358] 본 개시내용은 단리된 또는 실질적으로 정제된 폴리뉴클레오타이드 또는 단백질 조성물을 제공한다. "단리된" 또는 "정제된" 폴리뉴클레오타이드 또는 단백질, 또는 이의 생물학적 활성 부분은 그의 자연발생 환경에서 발견되는 폴리뉴클레오타이드 또는 단백질을 정상적으로 동반하거나 이와 상호작용하는 구성요소가 실질적으로 또는 본질적으로 없다. 따라서, 단리된 또는 정제된 폴리뉴클레오타이드 또는 단백질은 재조합 기술에 의해 생산될 때 다른 세포 물질 또는 배양 배지가 실질적으로 없거나, 또는 화학적으로 합성될 때 화학적 전구체 또는 다른 화학물질이 실질적으로 없다. 최적으로, "단리된" 폴리뉴클레오타이드는 폴리뉴클레오타이드가 유래된 유기체의 게놈 DNA에서 폴리뉴클레오타이드에 자연적으로 측접하는 서열(최적으로는 단백질 코딩 서열)(즉, 폴리뉴클레오타이드의 5' 및 3' 말단에 위치한 서열)이 없다. 예를 들어, 다양한 양태에서, 단리된 폴리뉴클레오타이드는 폴리뉴클레오타이드가 유래된 세포의 게놈 DNA에서 폴리뉴클레오타이드에 자연적으로 측접하는 뉴클레오타이드 서열의 약 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, 또는 0.1 kb 미만을 함유할 수 있다. 세포 물질이 실질적으로 없는 단백질은 약 30%, 20%, 10%, 5%, 또는 1%(건조 중량) 미만의 오염 단백질을 갖는 단백질 제제를 포함한다. 본 개시내용의 단백질 또는 이의 생물학적 활성 부분이 재조합적으로 생산되는 경우, 최적으로 배양 배지는 화학적 전구체 또는 비단백질 관심 화학물질의 약 30%, 20%, 10%, 5%, 또는 1%(건조 중량) 미만을 나타낸다.

[0359] 본 개시내용은 개시된 DNA 서열 및 이들 DNA 서열에 의해 코딩된 단백질의 단편 및 변이체를 제공한다. 본 개시내용 전반에 걸쳐 사용된 바와 같이, 용어 "단편"은 DNA 서열의 일부 또는 아미노산 서열의 일부 및 그에 따라 코딩되는 단백질을 지칭한다. 코딩 서열을 포함하는 DNA 서열의 단편은 천연 단백질의 생물학적 활성 및 이에 따라 본원에 기재된 바와 같은 표적 DNA 서열에 대한 DNA 인식 또는 결합 활성을 보유하는 단백질 단편을 코딩할 수 있다. 대안적으로, 혼성화 프로브로서 유용한 DNA 서열의 단편은 일반적으로 생물학적 활성을 유지하거나 프로모터 활성을 유지하지 않는 단백질을 코딩하지 않는다. 따라서, DNA 서열의 단편은 적어도 약 20개의 뉴클레오타이드, 약 50개의 뉴클레오타이드, 약 100개의 뉴클레오타이드, 및 본 개시내용의 최대 전장 폴리뉴클레오

타이드 범위일 수 있다.

[0360] 본 개시내용의 핵산 또는 단백질은 후속적으로 최종 목적지 벡터로 조립될 수 있는 표적 벡터에서 단량체 단위 및/또는 반복 단위를 사전 조립하는 것을 포함하는 모듈식 접근법에 의해 제작될 수 있다. 본 개시내용의 폴리펩타이드는 본 개시내용의 반복 단량체를 포함할 수 있고, 후속적으로 최종 목적지 벡터로 조립될 수 있는 표적 벡터에서 반복 단위를 사전 조립함으로써 모듈식 접근법에 의해 제작될 수 있다. 본 개시내용은 이 방법에 의해 생산된 폴리펩타이드뿐만 아니라 이들 폴리펩타이드를 코딩하는 핵산 서열을 제공한다. 본 개시내용은 이러한 모듈식 접근법에 의해 생산된 폴리펩타이드를 코딩하는 핵산 서열을 포함하는 숙주 유기체 및 세포를 제공한다.

[0361] 용어 "항체"는 가장 넓은 의미로 사용되며, 구체적으로 단일 단클론 항체(작용제 및 길항제 항체 포함) 및 폴리에피토프 특이성을 갖는 항체 조성물을 포함한다. 또한 본원에 정의된 바와 같은 본원의 항체의 천연 또는 합성 유사체, 돌연변이체, 변이체, 대립유전자, 상동체 및 오르소로그(ortholog)(본원에서 집합적으로 "유사체"로 지칭됨)를 사용하는 것은 본원의 범위에 속한다. 따라서, 이의 양태에 따르면, 용어 "본원의 항체"는 가장 넓은 의미에서 또한 이러한 유사체를 포함한다. 일반적으로, 이러한 유사체에서, 하나 이상의 아미노산 잔기는 본원에 정의된 바와 같은 본원의 항체와 비교하여 대체, 결실 및/또는 부가될 수 있다.

[0362] 본원에 사용된 바와 같이 "항체 단편", 및 모든 문법적 변형은 온전한 항체의 항원 결합 부위 또는 가변 영역을 포함하는 온전한 항체의 일부로서 정의되며, 상기 일부는 온전한 항체의 Fc 영역의 불변 중쇄 도메인(즉, 항체 아이소타입에 따라 CH2, CH3, 및 CH4)이 없다. 항체 단편의 예는 Fab, Fab', Fab'- SH, $F(ab')_2$, 및 Fv 단편; 디아바디; 비제한적으로 (1) 단일-사슬 Fv(scFv) 분자, (2) 결합된 중쇄 모이어티 없이, 오직 하나의 경쇄 가변 도메인 또는 경쇄 가변 도메인의 3개의 CDR을 함유하는 이의 단편을 함유하는 단일 사슬 폴리펩타이드 및 (3) 결합된 경쇄 모이어티 없이, 오직 하나의 중쇄 가변 영역 또는 중쇄 가변 영역의 3개의 CDR을 함유하는 이의 단편을 함유하는 단일 사슬 폴리펩타이드를 포함하는, 연속적인 아미노산 잔기의 하나의 중단되지 않은 서열로 구성되는 1차 구조를 갖는 폴리펩타이드인 임의의 항체 단편(본원에서 "단일-사슬 항체 단편" 또는 "단일 사슬 폴리펩타이드"로 지칭됨); 및 항체 단편으로부터 형성된 다중특이적 또는 다가 구조를 포함한다. 하나 이상의 중쇄를 포함하는 항체 단편에서, 중쇄(들)는 온전한 항체의 비Fc 영역에서 발견되는 임의의 불변 도메인 서열(예컨대, IgG 아이소타입에서 CHI)을 함유할 수 있고/거나, 온전한 항체에서 발견되는 임의의 힌지 영역 서열을 함유할 수 있고/거나, 힌지 영역 서열 또는 중쇄(들)의 불변 도메인 서열에 융합되거나 위치하는 류신 지퍼 서열을 함유할 수 있다. 용어는 일반적으로 단일 단량체성 가변 항체 도메인(예를 들어, 낙타과로부터 유래)을 갖는 항체 단편을 지칭하는 단일 도메인 항체("sdAB")를 추가로 포함한다. 이러한 항체 단편 유형은 당업자에 의해 쉽게 이해될 것이다.

[0363] "결합"은 거대분자 사이의(예컨대, 단백질 및 핵산 사이의) 서열 특이적 비공유 상호작용을 지칭한다. 상호작용이 그 전체로서 서열 특이적인 한, 결합 상호작용의 모든 구성요소가 서열 특이적(예컨대, DNA 백본에서 포스페이트 잔기와의 접촉)일 필요는 없다.

[0364] 용어 "포함하는"은 조성물 및 방법이 인용된 요소를 포함하지만, 다른 요소를 배제하지 않는다는 것을 의미하는 것으로 의도된다. 조성물 및 방법을 정의하는데 사용될 때 "본질적으로 구성되는"은 의도된 목적을 위해 사용될 때 조합에 임의의 본질적인 중요한 다른 요소를 배제하는 것을 의미한다. 따라서, 본원에 정의된 바와 같이 요소로 본질적으로 구성되는 조성물은 미량의 오염물질 또는 불활성 담체를 배제하지 않을 것이다. "구성되는"은 다른 성분 및 실질적인 방법 단계의 미량을 초과하는 원소를 배제하는 것을 의미한다. 이러한 각각의 전환 용어에 의해 정의되는 양태는 본 개시내용의 범위 내에 속한다.

[0365] 용어 "에피토프"는 폴리펩타이드의 항원 결정기를 지칭한다. 에피토프는 공간적 형태에 3개의 아미노산을 포함할 수 있고, 이는 에피토프에 고유하다. 일반적으로, 에피토프는 적어도 4, 5, 6, 또는 7개의 이러한 아미노산으로 구성되고, 보다 일반적으로, 적어도 8, 9, 또는 10개의 이러한 아미노산으로 구성된다. 아미노산의 공간적 형태를 결정하는 방법은 당업계에 알려져 있으며, 예를 들어, x-선 결정학 및 2차원 핵 자기 공명을 포함한다.

[0366] 본원에 사용된 바와 같이, "발현"은 폴리뉴클레오타이드가 mRNA로 전사되는 과정 및/또는 전사된 mRNA가 후속적으로 펩타이드, 폴리펩타이드, 또는 단백질로 번역되는 과정을 지칭한다. 폴리뉴클레오타이드가 게놈 DNA로부터 유래된 경우, 발현은 진핵 세포에서 mRNA의 스플라이싱을 포함할 수 있다.

[0367] "유전자 발현"은 유전자에 함유된 정보를 유전자 생성물로 전환시키는 것을 지칭한다. 유전자 생성물은 유전자의 직접적인 전사 생성물(예컨대, mRNA, tRNA, rRNA, 안티센스 RNA, 리보자임, shRNA, 마이크로 RNA, 구조적 RNA 또는 임의의 다른 유형의 RNA) 또는 mRNA의 번역에 의해 생산된 단백질일 수 있다. 유전자 생성물은 또한

캡핑, 폴리아데닐화, 메틸화, 및 편집과 같은 과정에 의해 변형된 RNA, 및, 예를 들어, 메틸화, 아세틸화, 인산화, 유비퀴틴화, ADP-리보실화, 미리스틸화, 및 글리코실화에 의해 변형된 단백질을 포함한다.

[0368] 유전자 발현의 "조절(Modulation 또는 regulation)"은 유전자의 활성의 변화를 지칭한다. 발현의 조절은, 비제한적으로, 유전자 활성화 및 유전자 억제를 포함할 수 있다.

[0369] 용어 "작동가능하게 연결된(operatively linked)" 또는 그의 균등물(예컨대, "작동가능하게 연결된(linked operatively)"은 2개 이상의 분자가 하나 또는 두 분자 또는 이의 조합에 기인하는 기능에 영향을 미치기 위해 상호작용할 수 있도록 서로에 대해 위치하는 것을 의미한다.

[0370] 비공유 연결된 구성요소 및 비공유 연결된 구성요소의 제조 및 사용 방법이 개시되어 있다. 다양한 구성요소는 본원에 기재된 바와 같이 다양한 상이한 형태를 취할 수 있다. 예를 들어, 비공유 연결된(즉, 작동가능하게 연결된) 단백질은 당업계의 하나 이상의 문제를 피하는 일시적인 상호작용을 허용하는 데 사용될 수 있다. 단백질과 같은 비공유 연결된 구성요소가 결합하고 해리하는 능력은 원하는 활성을 위해 이러한 결합이 필요한 상황에서만 또는 주로 이러한 상황에서 기능적 결합을 가능하게 한다. 연결은 원하는 효과를 허용하기에 충분한 기간일 수 있다.

[0371] 단백질을 유기체의 게놈 내의 특정 유전자좌로 향하게 하는 방법이 개시된다. 방법은 DNA 국소화 구성요소를 제공하고 효과기 분자를 제공하는 단계를 포함할 수 있고, DNA 국소화 구성요소 및 효과기 분자는 비공유 연결을 통해 작동가능하게 연결될 수 있다.

[0372] 용어 "scFv"는 단일 사슬 가변 단편을 지칭한다. scFv는 링커 펩타이드로 연결된 면역글로불린의 중쇄(VH) 및 경쇄(VL)의 가변 영역의 융합 단백질이다. 링커 펩타이드는 약 5 내지 40개의 아미노산 또는 약 10 내지 30개의 아미노산 또는 약 5, 10, 15, 20, 25, 30, 35, 또는 40개의 아미노산 길이일 수 있다. 단일 사슬 가변 단편은 완전한 항체 분자에서 발견되는 불변 Fc 영역이 없으므로, 항체를 정제하는 데 사용되는 공통 결합 부위(예컨대, 단백질 G)가 없다. 용어는 세포의 세포질에서 안정하고 세포내 단백질에 결합할 수 있는 항체인 인트라바디인 scFv를 추가로 포함한다.

[0373] 용어 "단일 도메인 항체"는 특정 항원에 선택적으로 결합할 수 있는 단일 단량체성 가변 항체 도메인을 갖는 항체 단편을 의미한다. 단일 도메인 항체는 일반적으로 중쇄 항체, 또는 공통 IgG의 하나의 가변 도메인(VH)을 포함하는 약 110개의 아미노산 길이의 펩타이드 사슬이며, 이는 일반적으로 전체 항체와 유사한 항원에 대한 친화성을 갖지만, 더 내열성이고 세제 및 고농도의 요소에 대해 안정하다. 예는 낙타과 또는 생선 항체로부터 유래된 것이다. 대안적으로, 단일 도메인 항체는 4개의 사슬을 갖는 공통 쥣과 또는 인간 IgG로부터 제조될 수 있다.

[0374] 본원에 사용된 바와 같이 용어 "특이적으로 결합하는" 및 "특이적 결합"은 항체, 항체 단편 또는 나노바디가 상이한 항원의 균질한 혼합물에 존재하는 특정 항원에 우선적으로 결합하는 능력을 지칭한다. 일부 양태에서, 특이적 결합 상호작용은 샘플에서 바람직한 항원 및 바람직하지 않은 항원을 구별할 것이다. 일부 양태에서, 약 10 내지 100배 초과(예컨대, 약 1000배 또는 10,000배 초과). "특이성"은 나노바디와 같은 면역글로불린 또는 면역글로불린 단편이 상이한 항원 표적에 비해 하나의 항원 표적에 우선적으로 결합하는 능력을 지칭하며, 반드시 높은 친화성을 의미하지 않는다.

[0375] "표적 부위" 또는 "표적 서열"은 결합을 위한 충분한 조건이 존재한다면 결합 분자가 결합할 핵산의 부분을 정의하는 핵산 서열이다.

[0376] 용어 "핵산" 또는 "올리고뉴클레오타이드" 또는 "폴리뉴클레오타이드"는 함께 공유 결합된 적어도 2개의 뉴클레오타이드를 지칭한다. 단일 가닥의 묘사는 또한 상보적 가닥의 서열을 정의한다. 따라서, 핵산은 또한 묘사된 단일 가닥의 상보적 가닥을 포함할 수 있다. 본 개시내용의 핵산은 또한 동일한 구조를 보유하거나 동일한 단백질을 코딩하는 실질적으로 동일한 핵산 및 이의 상보체를 포함한다.

[0377] 본 개시내용의 프로브는 엄격한 혼성화 조건 하에 표적 서열에 혼성화할 수 있는 단일 가닥의 핵산을 포함할 수 있다. 따라서, 본 개시내용의 핵산은 엄격한 혼성화 조건 하에 혼성화하는 프로브를 지칭할 수 있다.

[0378] 본 개시내용의 핵산은 단일 또는 이중 가닥일 수 있다. 본 개시내용의 핵산은 분자의 대부분이 단일 가닥인 경우에도 이중 가닥의 서열을 함유할 수 있다. 본 개시내용의 핵산은 분자의 대부분이 이중 가닥인 경우에도 단일 가닥의 서열을 함유할 수 있다. 본 개시내용의 핵산은 게놈 DNA, cDNA, RNA, 또는 이의 하이브리드를 포함할 수 있다. 본 개시내용의 핵산은 데옥시리보- 및 리보-뉴클레오타이드의 조합을 함유할 수 있다. 본 개시내용의 핵

산은 우라실, 아데닌, 티민, 시토신, 구아닌, 이노신, 잔틴, 하이포잔틴, 이소시토신 및 이소구아닌을 포함하는 염기의 조합을 함유할 수 있다. 본 개시내용의 핵산은 비천연 아미노산 변형을 포함하도록 합성될 수 있다. 본 개시내용의 핵산은 화학적 합성 방법에 의해 또는 재조합 방법에 의해 수득될 수 있다.

[0379] 본 개시내용의 핵산, 이들의 전체 서열, 또는 이의 임의의 부분은 비자연발생일 수 있다. 본 개시내용의 핵산은 자연적으로 발생하지 않아 전체 핵산 서열을 비자연발생으로 만드는 하나 이상의 돌연변이, 치환, 결실, 또는 삽입을 함유할 수 있다. 본 개시내용의 핵산은 하나 이상의 복제, 역전 또는 반복 서열을 함유할 수 있고, 생성된 서열은 자연적으로 발생하지 않아 전체 핵산 서열을 비자연발생으로 만든다. 본 개시내용의 핵산은 자연적으로 발생하지 않아 전체 핵산 서열을 비자연발생으로 만드는 변형된, 인공, 또는 합성 뉴클레오타이드를 함유할 수 있다.

[0380] 유전적 코드의 중복성을 감안할 때, 복수의 뉴클레오타이드 서열은 임의의 특정 단백질을 코딩할 수 있다. 이러한 모든 뉴클레오타이드 서열이 본원에서 고려된다.

[0381] 본 개시내용 전반에 걸쳐 사용된 바와 같이, 용어 "작동가능하게 연결된"은 공간적으로 연결된 프로모터의 제어 하에 있는 유전자의 발현을 지칭한다. 프로모터는 제어 하에 있는 유전자의 5'(상류) 또는 3'(하류)에 위치할 수 있다. 프로모터 및 유전자 사이의 거리는 프로모터 및 프로모터가 유래된 유전자에서 그것이 제어하는 유전자 사이의 거리와 거의 동일할 수 있다. 프로모터 및 유전자 사이의 거리 변화는 프로모터 기능의 손실 없이 수용될 수 있다.

[0382] 본 개시내용 전반에 걸쳐 사용된 바와 같이, 용어 "프로모터"는 세포에서 핵산의 발현을 부여, 활성화 또는 향상시킬 수 있는 합성 또는 천연 유래된 분자를 지칭한다. 프로모터는 발현을 추가로 향상시키고/거나 이의 공간적 발현 및/또는 일시적 발현을 변경하기 위해 하나 이상의 특이적 전사 조절 서열을 포함할 수 있다. 프로모터는 또한 전사의 시작 부위로부터 수천 염기쌍만큼 위치할 수 있는 원위 인핸서 또는 억제자 요소를 포함할 수 있다. 프로모터는 바이러스, 박테리아, 진균, 식물, 곤충, 및 동물을 포함하는 공급원으로부터 유래될 수 있다. 프로모터는 발현이 일어나는 세포, 조직 또는 장기에 대해 또는, 발현이 일어나는 발달 단계에 대해, 또는 생리적 스트레스, 병원체, 금속 이온, 또는 유도제와 같은 외부 자극에 대한 반응으로 항시적으로 또는 차등적으로 유전자 구성요소의 발현을 조절할 수 있다. 프로모터의 대표적인 예는 박테리오파아지 T7 프로모터, 박테리오파아지 T3 프로모터, SP6 프로모터, lac 오퍼레이터-프로모터, tac 프로모터, SV40 후기 프로모터, SV40 초기 프로모터, RSV-LTR 프로모터, CMV IE 프로모터, EF-1 알파 프로모터, CAG 프로모터, SV40 초기 프로모터 또는 SV40 후기 프로모터 및 CMV IE 프로모터를 포함한다.

[0383] 본 개시내용 전반에 걸쳐 사용된 바와 같이, 용어 "실질적으로 상보적인"은 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 180, 270, 360, 450, 540개 이상의 뉴클레오타이드 또는 아미노산의 영역에 걸쳐 제2 서열의 상보체와 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% 또는 99% 동일한 제1 서열을 지칭하거나, 또는 두 서열이 엄격한 혼성화 조건 하에서 혼성화한다는 것을 지칭한다.

[0384] 본 개시내용 전반에 걸쳐 사용된 바와 같이, 용어 "실질적으로 동일한"은 제1 및 제2 서열이 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 180, 270, 360, 450, 540개 이상의 뉴클레오타이드 또는 아미노산의 영역에 걸쳐, 또는 제1 서열이 제2 서열의 상보체와 실질적으로 동일한 경우 핵산에 대해, 적어도 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98% 또는 99% 동일한 것을 지칭한다.

[0385] 본 개시내용 전반에 걸쳐 사용된 바와 같이, 핵산을 설명하기 위해 사용될 때 용어 "변이체"는 (i) 참조된 뉴클레오타이드 서열의 일부 또는 단편; (ii) 참조된 뉴클레오타이드 서열의 상보체 또는 이의 부분; (iii) 참조된 핵산 또는 이의 상보체와 실질적으로 동일한 핵산; 또는 (iv) 엄격한 조건 하에 참조된 핵산, 이의 상보체, 또는 이와 실질적으로 동일한 서열에 혼성화하는 핵산을 지칭한다.

[0386] 본 개시내용 전반에 걸쳐 사용된 바와 같이, 용어 "벡터"는 복제 원점을 함유하는 핵산 서열을 지칭한다. 벡터는 바이러스 벡터, 박테리오파아지, 박테리아 인공 염색체 또는 효모 인공 염색체일 수 있다. 벡터는 DNA 또는 RNA 벡터일 수 있다. 벡터는 자가 복제하는 염색체외 벡터일 수 있고, 바람직하게는, DNA 플라스미드이다. 벡터는 아미노산과 DNA 서열, RNA 서열, 또는 DNA 및 RNA 서열 모두와의 조합을 포함할 수 있다.

[0387] 본 개시내용 전반에 걸쳐 사용된 바와 같이, 펩타이드 또는 폴리펩타이드를 설명하기 위해 사용될 때 용어 "변이체"는 아미노산의 삽입, 결실, 또는 보존적 치환에 의해 아미노산 서열이 상이하지만 적어도 하나의 생물학적

활성을 유지하는 펩타이드 또는 폴리펩타이드를 지칭한다. 변이체는 또한 적어도 하나의 생물학적 활성을 유지하는 아미노산 서열을 갖는 참조된 단백질과 실질적으로 동일한 아미노산 서열을 갖는 단백질을 의미할 수 있다.

[0388] 아미노산의 보존적 치환, 즉, 아미노산을 유사한 특성(예컨대, 하전된 영역의 친수성, 정도 및 분포)의 상이한 아미노산으로 대체하는 것은 전형적으로 사소한 변화를 수반하는 것으로 당업계에서 인식된다. 이러한 사소한 변화는 부분적으로 당업계에서 이해되는 바와 같이 아미노산의 소수성 지수(hydropathic index)를 고려함으로써 확인될 수 있다(Kyte et al., J. Mol. Biol. 157: 105-132(1982)). 아미노산의 소수성 지수는 그의 소수성과 전하의 고려에 기초한다. 유사한 소수성 지수의 아미노산은 치환될 수 있고, 여전히 단백질 기능을 유지한다. 일 양태에서, ±2의 소수성 지수를 갖는 아미노산은 치환된다. 아미노산의 친수성은 또한 단백질이 생물학적 기능을 유지하게 할 치환을 밝히는 데 사용될 수 있다. 펩타이드의 문맥에서 아미노산의 친수성을 고려하면 항원성 및 면역원성과 상관관계가 있는 것으로 보고된 유용한 척도인 펩타이드의 가장 큰 국소 평균 친수성을 계산할 수 있다(그 전체가 참조로 본원에 포함된 미국 특허 제4,554,101호).

[0389] 유사한 친수성 값을 갖는 아미노산의 치환은 생물학적 활성, 예를 들어 면역원성을 유지하는 펩타이드를 생성할 수 있다. 치환은 서로 ±2 이내의 친수성 값을 갖는 아미노산을 이용하여 수행될 수 있다. 아미노산의 소수성 지수 및 친수성 값은 아미노산의 특정 측쇄에 의해 영향을 받는다. 이러한 관찰과 일치하게도, 생물학적 기능과 양립할 수 있는 아미노산 치환은 소수성, 친수성, 전하, 크기, 및 다른 특성에 의해 밝혀진 바와 같이 아미노산의 상대적 유사성, 및 특히 상기 아미노산의 측쇄에 의존하는 것으로 이해된다.

[0390] 본원에 사용된 바와 같이, "보존적" 아미노산 치환은 하기 표 A, B, 또는 C에 기재된 바와 같이 정의될 수 있다. 일부 양태에서, 융합 폴리펩타이드 및/또는 이러한 융합 폴리펩타이드를 코딩하는 핵산은 본 개시내용의 폴리펩타이드를 코딩하는 폴리뉴클레오타이드의 변형에 의해 도입된 보존적 치환을 포함한다. 아미노산은 물리적 특성 및 2차 및 3차 단백질 구조에 대한 기여에 따라 분류될 수 있다. 보존적 치환은 하나의 아미노산을 유사한 특성을 갖는 또 다른 아미노산으로 치환하는 것이다. 예시적인 보존적 치환이 표 A에 제시되어 있다.

[0391] 표 A -- 보존적 치환 I

| 측쇄 특징 | | 아미노산 |
|---|---|---|
| 지방족 | 비극성 | G A P I L V F |
| | 극성 - 하전되지 않음 | C S T M N Q |
| | 극성 - 하전됨 | D E K R |
| 방향족 | | H F W Y |
| 기타 | | N Q D E |

[0392]

[0393] 대안적으로, 보존적 아미노산은 표 B에 제시된 바와 같이 문헌[Lehninger, (Biochemistry, 제2 Edition; Worth Publishers, Inc. NY, N.Y.(1975), pp. 71-77)]에 기재된 바와 같이 분류될 수 있다.

[0394] 표 B -- 보존적 치환 II

| 측쇄 특징 | | 아미노산 |
|---|---|---|
| 비극성 (소수성) | 지방족: | A L I V P |
| | 방향족: | F W Y |
| | 황 함유: | M |
| | 경계: | G Y |
| 하전되지 않음-극성 | 하이드록실: | S T Y |
| | 아미드: | N Q |
| | 설프하이드릴: | C |
| | 경계: | G Y |
| 양으로 하전됨 (염기성): | | K R H |
| 음으로 하전됨 (산성): | | D E |

[0395]

[0396] 대안적으로, 예시적인 보존적 치환이 표 C에 제시되어 있다.

[0397] 표 C -- 보존적 치환 III

| 원래의 잔기 | 예시적인 치환 |
|---|---|
| Ala (A) | Val Leu Ile Met |
| Arg (R) | Lys His |
| Asn (N) | Gln |
| Asp (D) | Glu |
| Cys (C) | Ser Thr |
| Gln (Q) | Asn |
| Glu (E) | Asp |
| Gly (G) | Ala Val Leu Pro |
| His (H) | Lys Arg |
| Ile (I) | Leu Val Met Ala Phe |
| Leu (L) | Ile Val Met Ala Phe |
| Lys (K) | Arg His |
| Met (M) | Leu Ile Val Ala |
| Phe (F) | Trp Tyr Ile |
| Pro (P) | Gly Ala Val Leu Ile |
| Ser (S) | Thr |
| Thr (T) | Ser |
| Trp (W) | Tyr Phe Ile |
| Tyr (Y) | Trp Phe Thr Ser |
| Val (V) | Ile Leu Met Ala |

[0398]

[0399] 본 개시내용의 폴리펩타이드는 아미노산 잔기의 하나 이상의 삽입, 결실, 또는 치환, 또는 이의 임의의 조합뿐만 아니라 아미노산 잔기의 삽입, 결실, 또는 치환 이외의 변형을 갖는 폴리펩타이드를 포함하고자 하는 것으로

이해되어야 한다. 본 개시내용의 폴리펩타이드 또는 핵산은 하나 이상의 보존적 치환을 포함할 수 있다.

[0400] 본 개시내용 전반에 걸쳐 사용된 바와 같이, 용어 전술한 아미노산 치환 중 "하나 초과"는 언급된 아미노산 치환 중 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20개 이상을 지칭한다. 용어 "하나 초과"는 언급된 아미노산 치환 중 2, 3, 4, 또는 5개를 지칭할 수 있다.

[0401] 본 개시내용의 폴리펩타이드 및 단백질, 이들의 전체 서열, 또는 이의 임의의 부분은 비자연발생일 수 있다. 본 개시내용의 폴리펩타이드 및 단백질은 자연적으로 발생하지 않아 전체 아미노산 서열을 비자연발생으로 만드는 하나 이상의 돌연변이, 치환, 결실, 또는 삽입을 함유할 수 있다. 본 개시내용의 폴리펩타이드 및 단백질은 하나 이상의 복제, 역전 또는 반복 서열을 함유할 수 있고, 생성된 서열은 자연적으로 발생하지 않아 전체 아미노산 서열 서열을 비자연발생으로 만든다. 본 개시내용의 폴리펩타이드 및 단백질은 자연적으로 발생하지 않아 전체 아미노산 서열 서열을 비자연발생으로 만드는 변형된, 인공, 또는 합성 아미노산을 함유할 수 있다.

[0402] 본 개시내용 전반에 걸쳐 사용된 바와 같이, "서열 동일성"은 기본 매개변수를 사용하여 NCBI(National Center for Biotechnology Information) ftp 사이트로부터 검색될 수 있는 2개의 서열을 블라스팅하기 위한 독립 실행형 BLAST 엔진 프로그램(b12seq)을 사용하여 결정될 수 있다(Tatusova and Madden, FEMS Microbiol Lett., 1999, 174, 247-250; 그 전체가 참조로 본원에 포함됨). 2개 이상의 핵산 또는 폴리펩타이드 서열의 문맥에서 사용될 때 용어 "동일한" 또는 "동일성"은 각각의 서열의 특정 영역에 걸쳐 동일한 잔기의 특정 백분율을 지칭한다. 백분율은 2개의 서열을 최적으로 정렬하고, 특정 영역에 걸쳐 2개의 서열을 비교하고, 두 서열에서 동일한 잔기가 발생하는 위치의 수를 계산하여 일치된 위치의 수를 산출하고, 일치된 위치의 수를 특정 영역에서의 위치의 총 수로 나누고, 결과에 100을 곱하여 서열 동일성의 백분율을 산출함으로써 계산될 수 있다. 2개의 서열이 길이가 상이하거나 정렬이 하나 이상의 엇갈린 말단을 생성하고 특정 비교 영역이 단일 서열만을 포함하는 경우, 단일 서열의 잔기는 계산의 분자가 아닌 분모에 포함된다. DNA 및 RNA를 비교할 때, 티민(T) 및 우라실(U)은 동등한 것으로 간주될 수 있다. 동일성은 수동으로 또는 BLAST 또는 BLAST 2.0과 같은 컴퓨터 서열 알고리즘을 사용하여 수행될 수 있다.

[0403] 본 개시내용 전반에 걸쳐 사용된 바와 같이, 용어 "내인성"은 표적 유전자 또는 그것이 도입되는 숙주 세포와 자연적으로 회합되는 핵산 또는 단백질 서열을 지칭한다.

[0404] 본 개시내용 전반에 걸쳐 사용된 바와 같이, 용어 "외인성"은 표적 유전자 또는 그것이 도입되는 숙주 세포와 자연적으로 회합되지 않는 핵산 또는 단백질 서열을 지칭하며, 이는 자연발생 핵산, 예컨대, DNA 서열의 비자연발생 다중 사본, 또는 비자연발생 게놈 위치에 위치한 자연발생 핵산 서열을 포함한다.

[0405] 본 개시내용은 DNA 서열을 포함하는 폴리뉴클레오타이드 구조체를 숙주 세포 내로 도입하는 방법을 제공한다. "도입"은 구조체가 숙주 세포의 내부에 접근하는 방식으로 폴리뉴클레오타이드 구조체를 세포에 제시하는 것으로 의도된다. 본 개시내용의 방법은 폴리뉴클레오타이드 구조체가 숙주의 하나의 세포의 내부에 접근하는 것을 제외하면 폴리뉴클레오타이드 구조체를 숙주 세포 내로 도입하기 위한 특정 방법에 좌우되지 않는다. 폴리뉴클레오타이드 구조체를 박테리아, 식물, 진균 및 동물에 도입하기 위한 방법은 당업계에 알려져 있으며, 이는 비제한적으로, 안정한 형질전환 방법, 일시적 형질전환 방법, 및 바이러스 매개 방법을 포함한다.

[0406] 실시예 1: PiggyBac 나노트랜스포존을 이용한 개선된 전위

[0407] 인간 범(pan) T 세포의 게놈 내로의 전위 효율에 대한 piggyBac 트랜스포존 플라스미드 백본을 단축하여 ITR 사이의 거리를 감소시키는 효과를 평가하였다. 각각 GFP를 코딩하는 트랜스포존을 포함하는, 전체 크기의 piggyBac 플라스미드(FP)(도 1), piggyBac 나노트랜스포존(NT)(도 1) 및 piggyBac 짧은 나노트랜스포존(NTS)(도 3)을 제작하였다. FP 백본은 박테리아 pUC 복제 원점 뿐만 아니라 Kan/Neo 내성 유전자를 코딩하였다. NT 백본은 RNA-OUT 요소 뿐만 아니라 R6K 미니 복제 원점을 포함하는 항생제가 없는 수크로스 선택가능한 나노플라스미드 백본을 코딩하였다. NTS는 백본 RNA-OUT 요소 및 R6K 미니 복제 원점이 트랜스포존 요소 내부에 배치되어 ITR 사이의 거리를 추가로 감소시킨다는 점에서 NT와 상이하였다. 도 1은 전체 플라스미드 및 NT 나노트랜스포존 사이의 차이를 예시한다. 트랜스포존의 크기는 일정하게 유지되었지만(3,606 bp), NT 백본은 더 짧아서, ITR 측접 거리를 표 1에 상세히 설명된 바와 같이 2,034 bp에서 493 bp로 4배 넘게 효과적으로 감소시켰다. 도 3은 piggyBac NT 및 piggyBac NTS 사이의 차이를 예시한다. 트랜스포존의 크기가 3,614 bp에서 4,069로 증가하였지만(RNA-OUT 및 R6K 서열을 혼입하기 위해), NTS 백본은 더 짧아서, ITR 측접 거리를 표 2에 상세히 설명된 바와 같이 485 bp에서 48 bp로 10배 넘게 효과적으로 감소시켰다.

[0408] **표 1**

| | 플라스미드 크기 | 트랜스포존 크기 | ITR 측접 거리 |
|---|---|---|---|
| PB-FP | 5,640 | 3,606 | 2,034 |
| PB-NT | 4,099 | 3,606 | 493 |
| 배수 변화 | 1.38 | 1.00 | 4.13 |

[0409]

[0410] **표 2**

| | 플라스미드 크기 | 트랜스포존 크기 | ITR 측접 거리 |
|---|---|---|---|
| PB-NT | 4,099 | 3,614 | 485 |
| PB-NTS | 4,117 | 4,069 | 48 |
| 배수 변화 | 1.00 | 0.89 | 10.10 |

[0411]

[0412] piggyBac FP 또는 piggyBac NT를 Super piggyBac 트랜스포사제 효소(SPB)를 코딩하는 mRNA와 함께 또는 없이 전기천공(EP)을 통해 인간 범 T 세포에 전달하였다(도 2). 추가로, FP 및 NT를 등몰 또는 등질량으로 세포에 전달하였다. EP 후, 세포를 표준 TCR 활성화 시약을 사용하여 자극시키고, GFP 발현을 15일 후 FACS에 의해 평가하였다. 이러한 데이터는 ITR 측접 거리를 4배 넘게(2,034 bp에서 493 bp로) 감소시키는 것이 FP GFP 트랜스포존 플라스미드와 비교하여 등몰 및 등질량에서 더 큰 수준의 GFP 전위를 초래하였음을 보여준다. 또한, GFP 발현은 트랜스포존의 안정적인 통합의 결과였는데, GFP 발현이 SPB의 부재 하에 전기천공된 T 세포에서 검출되지 않았기 때문이었다.

[0413] 도 1 및 도 2는 piggyBac 트랜스포존 플라스미드 백본을 단축시키는 것이 인간 범 T 세포 내로의 트랜스포존의 전위 효율을 증가시켰음을 입증한다. 그러나, 총 플라스미드 크기가 FP 및 NT 사이에서 동일하지 않았기 때문에, NT에 의한 향상된 전위 효율이 더 작은 플라스미드(등몰; 더 적은 DNA가 세포에 전달됨), 전달되는 더 많은 총 플라스미드(등몰), 또는 piggyBac ITR에 측접하는 더 짧은 거리의 결과였는지 여부가 불분명하다. 인간 범 T 세포에 전달된 DNA는 면역조절 효과를 유발할 수 있고 독성일 수 있기 때문에, NT에 의한 향상된 전위 효율은 더 적은 총 DNA(등몰)의 전달 또는 더 많은 플라스미드(등질량)의 전달의 결과일 수 있다. 이를 시험하기 위해, 나노플라스미드 백본이 인슐레이터 및 ITR 사이에 위치한 트랜스포존 내에 재위치한 NTS를 제작하였다(도 3). NT 및 NTS의 크기는 일정하게 유지되었지만(4,099 및 4,117 bp), NTS에서의 ITR 측접 거리는 10배 더 짧았다(485 bp에서 48 bp로)(표 2). 등몰/등질량의 NT 또는 NTS를 SPB를 코딩하는 mRNA와 함께 전기천공(EP)을 통해 인간 범 T 세포에 전달하였다. EP 후, 세포를 표준 TCR 활성화 시약을 사용하여 자극시키고, GFP 발현을 15일 후에 FACS에 의해 평가하였다. 이러한 데이터는 총 플라스미드 크기를 일정하게 유지하면서 ITR 측접 거리를 10배 넘게(485 bp에서 48 bp로) 감소시키는 것이 NT GFP 나노트랜스포존과 비교하여 NTS에 의한 더 큰 수준의 GFP 전위를 초래하였음을 보여준다(도 4).

[0414] 실시예 2: BCMA CAR 및 PSMA CAR 나노트랜스포존의 전위

[0415] 전체 크기의 piggyBac 플라스미드(FP) 또는 piggyBac 나노트랜스포존(NT)을 코딩하는 항-BCMA CAR 및 항-PSMA CAR을 Super piggyBac 트랜스포사제 효소를 코딩하는 mRNA와 함께 전기천공(EP)을 통해 인간 범 T 세포에 등질량으로 전달하였다(도 5). EP 후, 세포를 선택 시약의 부재 하에 표준 TCR 활성화 시약을 사용하여 자극시키고, CAR 발현을 5일 후에 FACS에 의해 평가하였다. 이러한 데이터는 두 NT가 FP 트랜스포존 플라스미드와 비교하여 등질량에서 더 큰 수준의 전위를 초래하였음을 보여준다. 이것은 CAR-T 세포가 2명의 상이한 정상 공여체의 인간 범 T 세포로부터 생산되었을 때 사실이었다. 전체 크기의 piggyBac 플라스미드(FP) 또는 piggyBac 나노트랜스포존(NT)을 사용하여 생산된 항-BCMA CAR 및 항-PSMA CAR T 세포를 본원에 기재된 바와 같이 생산하였다. 표시된 효과기 대 표적 비율에서 CAR-T 세포에 의한 BCMA(K562.BCMA) 또는 PSMA(K562.PSMA)를 발현하도록 조작된 K562 세포의 사멸(도 6). 이들 데이터는 모든 CAR-T 세포가 FP 또는 NT를 사용하여 생산되었는지 여부에 관계없이 항원 의존적 방식으로 표적 종양 세포를 사멸시킬 수 있었음을 보여준다. 이것은 2명의 상이한 정상 공여체의 인간 범 T 세포로부터 생산된 CAR-T 세포에 대해 사실이었다. 도 7은 항-BCMA CAR 또는 항-PSMA CAR 나노트

랜스포존(NT)을 사용하여 생산된 인간 CAR-T 세포가 표현형 조성에서 대등하였음을 보여주는 일련의 그래프이다. 전체 크기의 piggyBac 플라스미드(FP) 또는 piggyBac 나노트랜스포존(NT)을 사용하여 생산된 항-BCMA CAR 및 항-PSMA CAR T 세포를 본원에 기재된 바와 같이 제조하였다. 기억 T 세포 마커 및 활성화/소진 마커의 표현형 분석(데이터는 나타내지 않음)을 수행하였다. 이들 데이터는 모든 CAR-T 세포가 FP 또는 NT를 사용하여 생산되었는지 여부에 관계없이 CD45RA+CD62L+(Tscm), CD45RA-CD62L+(Tcm), CD45RA-CD62L-(Tem), 및 CD45RA+CD62L-(Teff) 세포의 유사한 표현형 조성을 나타내었음을 보여준다. 또한, CCR7(CD197), CD127, CD27, LAG3, TIM3, CXCR3, PD-1, 및 CD25의 대등한 수준의 발현이 관찰되었다(데이터는 나타내지 않음). 이것은 2명의 상이한 정상 공여체의 인간 범 T 세포로부터 생산된 CAR-T 세포에 대해 사실이었다. 통합된 트랜스포존의 평균 카피 수를 정량적 PCR에 의해 측정하였다. 이들 데이터는 2명의 상이한 공여체에서, 모든 CAR-T 세포가 FP 또는 NT를 사용하여 생산되었는지 여부에 관계없이 트랜스포존의 유사한 통합된 카피 수를 나타내었음을 보여준다(도 8).

[0416] 고도의 다량체성 로트(7% 단량체성)를 고도의 단량체성 로트(87% 단량체성)와 상이한 비율로 혼합함으로써 상이한 단량체성 순도의 항-BCMA CAR piggyBac 나노트랜스포존을 생산하였다. 2개의 로트는 유전자 시퀀싱에 의해 단량체성 또는 다량체성 구조에서 1차 수준에서 동일하고 3차 수준에서만 상이한 것으로 확인되었다. 혼합된 항-BCMA CAR NT의 각각의 새로운 로트를 제한 절단의 부재 하에 아가로스 겔에서 실행하여 생성된 단량체성 대 다량체성 나노트랜스포존의 비율을 밝혀내었고; 많은 다양한 단량체성 순도를 생성하였다(7%, 32%, 45%, 59%, 65%, 72%, 및 87%). 겔 상에서, 다량체성 NT는 단량체성 NT보다 느리게 이동하였다(도 9). 도 9에 도시된 겔의 밴드는 다량체성(위) 및 단량체성(아래) 나노트랜스포존을 예시하기 위해 직사각형에 의해 박스표시되어 있으며; 넘버링은 위에서 아래로, 왼쪽에서 오른쪽으로 진행된다: 1(다량체성) 및 2(단량체성)[7% 단량체성 순도], 3 및 4[32% 단량체성 순도], 5 및 6[45% 단량체성 순도], 7 및 8[59% 단량체성 순도], 9 및 10[65% 단량체성 순도], 11 및 12[72% 단량체성 순도], 13 및 14[블랭크], 15 및 16[87% 단량체성 순도].

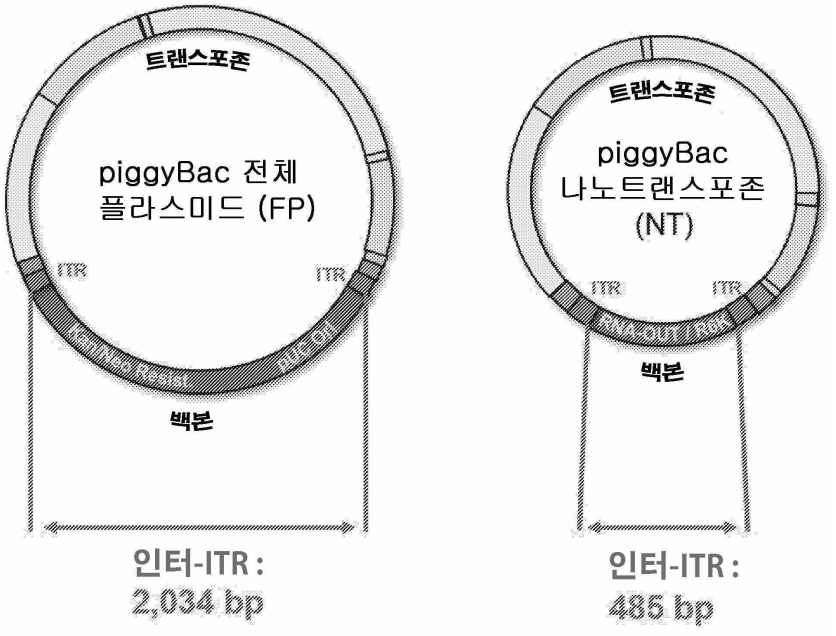
[0417] 상이한 단량체성 순도의 항-BCMA CAR piggyBac NT를 Super piggyBac 트랜스포사제 효소를 코딩하는 mRNA와 함께 전기천공(EP)을 통해 인간 범 T 세포에 전달하였다(도 10). 대조군으로서, 94% 단량체성 순도의 전체 크기의 항-BCMA CAR 플라스미드(FP)를 또한 등몰량으로 전달하였다. EP 후, 세포를 선택 시약의 부재 하에 표준 TCR 활성화 시약을 사용하여 자극시키고 CAR 발현을 5일 후에 FACS에 의해 평가하였다. 이들 데이터는 2명의 개별 공여체(공여체 #3 및 공여체 #2)에서 단량체성 순도가 전위 효율에 긍정적인 영향을 미친다는 것을 보여준다. 또한, 이들 데이터는 NT가 FP 트랜스포존 플라스미드 등몰량과 비교하여 더 큰 수준의 전위를 초래하였음을 보여준다.

[0418] 실시예 3: P-PSMA-101 나노트랜스포존의 전임상 평가

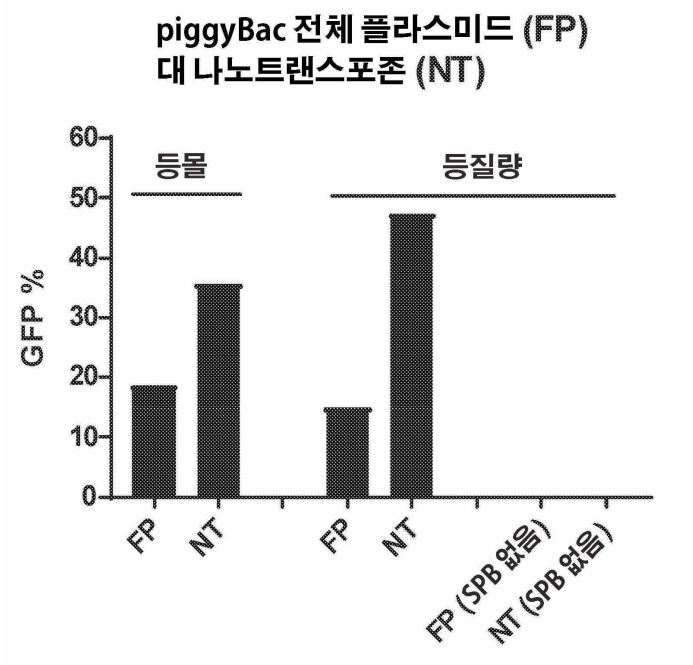
[0419] 쥣과 이종이식편 모델을 사용하여 '스트레스' 용량에서 전장 플라스미드(FLP) 대 나노트랜스포존(NT)에 의해 전달될 때 P-PSMA-101 트랜스포존의 효능을 전임상 환경에서 평가하였다. NSG 마우스에 피하(SC) 주사된 루시퍼라제 발현 LNCaP 세포주(LNCaP.luc)를 사용한 쥣과 이종이식편 모델을 이용하여 2명의 상이한 정상 공여체로부터의 총 CAR-T 세포의 2개의 상이한 '스트레스' 용량(2.5x10^6 또는 4x10^6)에서 전장 플라스미드(FLP) 또는 나노트랜스포존(NT)에 의해 전달된 바와 같은 P-PSMA-101 트랜스포존의 생체내 항종양 효능을 평가하였다(도 11). 모든 CAR-T 세포를 FLP 또는 NT 전달을 사용하여 P-PSMA-101 트랜스포존의 piggyBac(PB) 전달을 사용하여 생산하였다. 마우스에게 LNCaP를 겨드랑이로 주사하고, 종양이 확립되었을 때(캘리퍼 측정에 의해 100-200 $mm^3$) 치료하였다. FLP 및 NT에 의한 트랜스포존 전달 사이의 가능한 효능의 기능적 차이를 감지하는 데 있어서 더 큰 해상도를 위해 마우스에게 IV 주사에 의해 P-PSMA-101 CAR-T의 2개의 상이한 '스트레스' 용량($2.5x10^6$ 또는 $4x10^6$)을 처리하였다. 오차 막대와 함께 그룹 평균(위) 및 개별 마우스(아래)로서 나타낸 바와 같은 대조군 마우스(검정색), 공여체 #1 FLP 마우스(빨간색), 공여체 #1 NT 마우스(파란색), 공여체 #2 FLP 마우스(주황색), 및 공여체 #2 NT 마우스(녹색)에 대한 캘리퍼 측정에 의한 종양 부피 평가(도 12). y-축은 캘리퍼 측정에 의해 평가된 종양 부피($mm^3$)를 보여준다. x-축은 T 세포 치료 후 일수를 보여준다. NT에 의해 전달될 때, '스트레스' 용량에서 P-PSMA-101 트랜스포존은 확립된 SC LNCaP.luc 고형 종양에 대해 FLP 및 대조군 마우스와 비교하여 캘리퍼에 의해 측정된 바와 같이 향상된 항종양 효능을 입증하였다.

도면

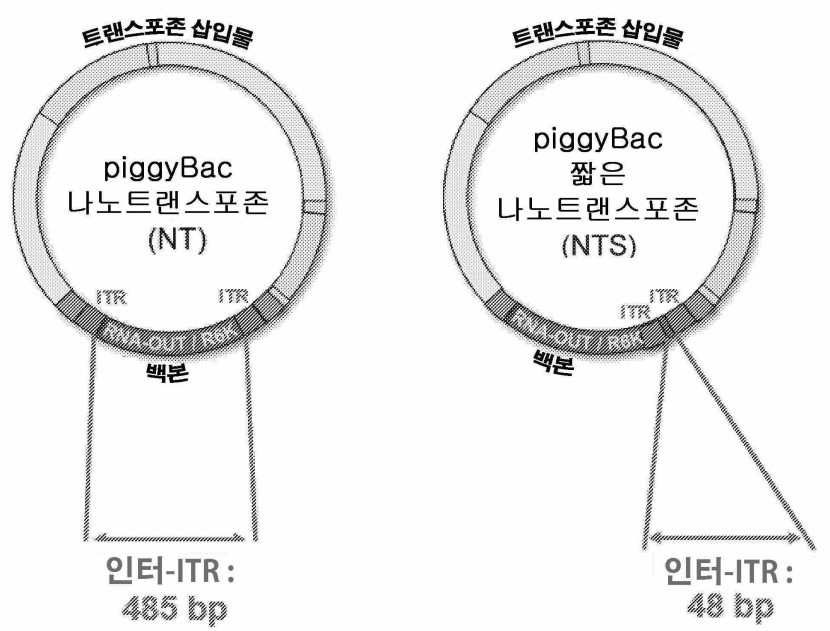
도면1


트랜스포존
piggyBac 전체
플라스미드 (FP)
ITR
ITR
백본
인터-ITR :
2,034 bp
트랜스포존
piggyBac
나노트랜스포존
(NT)
ITR
ITR
백본
인터-ITR :
485 bp

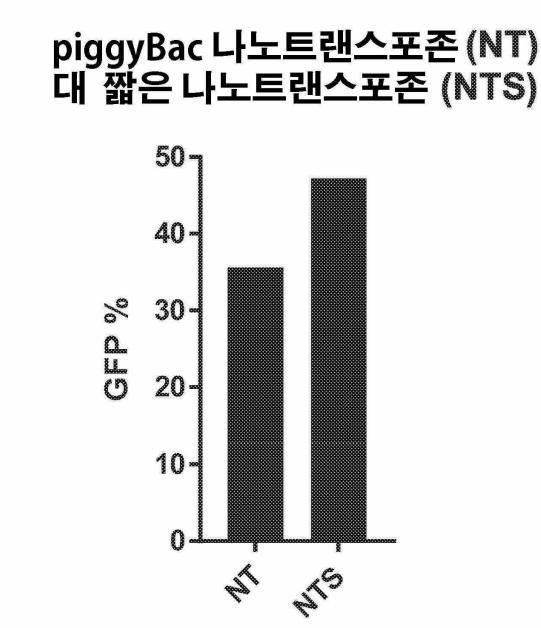
도면2


piggyBac 전체 플라스미드 (FP)
대 나노트랜스포존 (NT)
등몰
등질량
GFP %
60
50
40
30
20
10
0
FP
NT
FP
NT
FP (SPB 없음)
NT (SPB 없음)

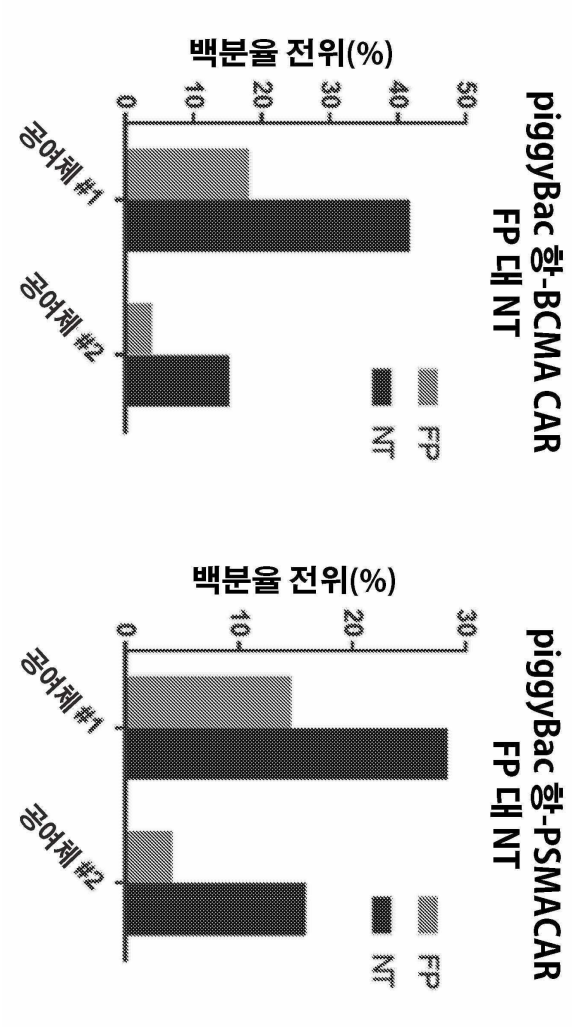
도면3


트랜스포존 삽입물
piggyBac
나노트랜스포존
(NT)
ITR
ITR
백본
인터-ITR :
485 bp
트랜스포존 삽입물
piggyBac
짧은
나노트랜스포존
(NTS)
ITR
ITR
백본
인터-ITR :
48 bp

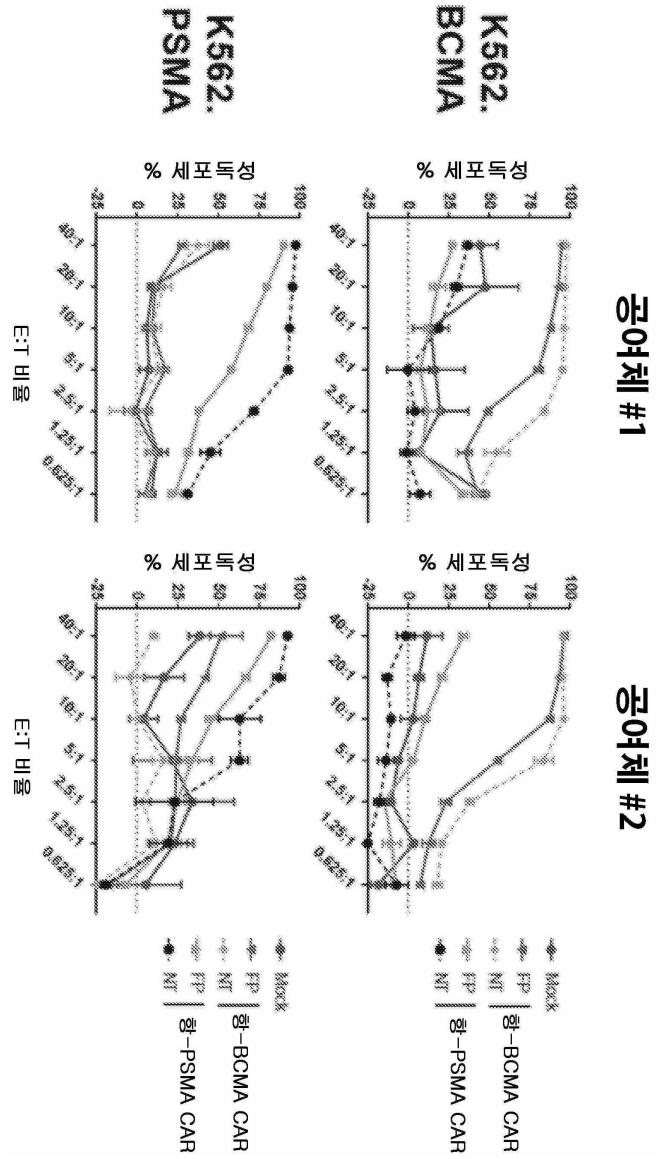
도면4


piggyBac 나노트랜스포존 (NT)
대 짧은 나노트랜스포존 (NTS)
GFP %
50
40
30
20
10
0
NT
NTS

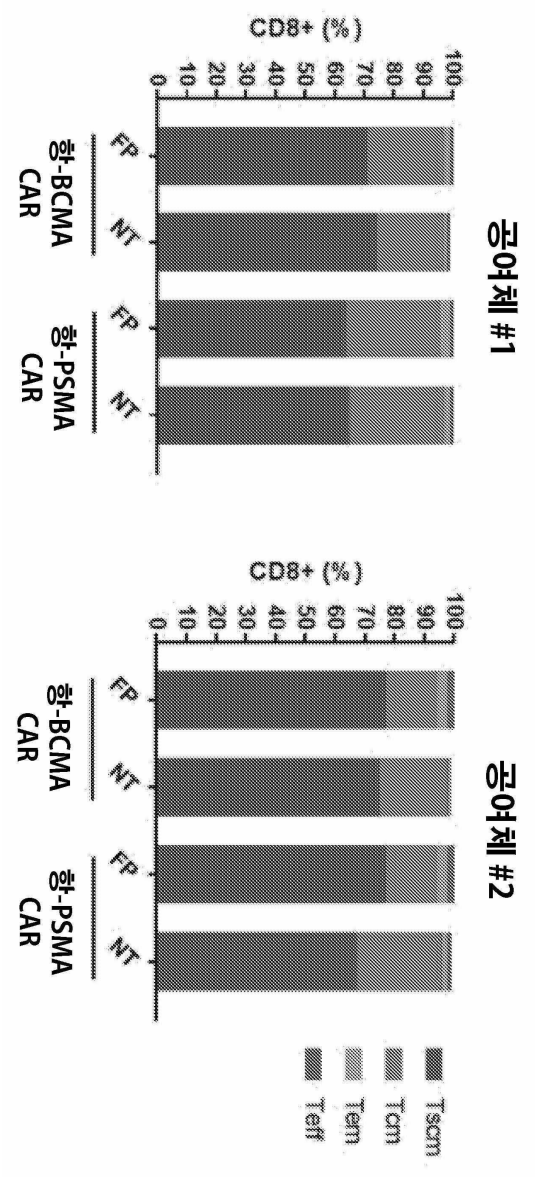
도면5


piggyBac 항-BCMA CAR
FP 대 NT
백분율 전위(%)
0
10
20
30
40
50
공여체 #1
공여체 #2
NT
FP
piggyBac 항-PSMACAR
FP 대 NT
백분율 전위(%)
0
10
20
30
공여체 #1
공여체 #2
NT
FP

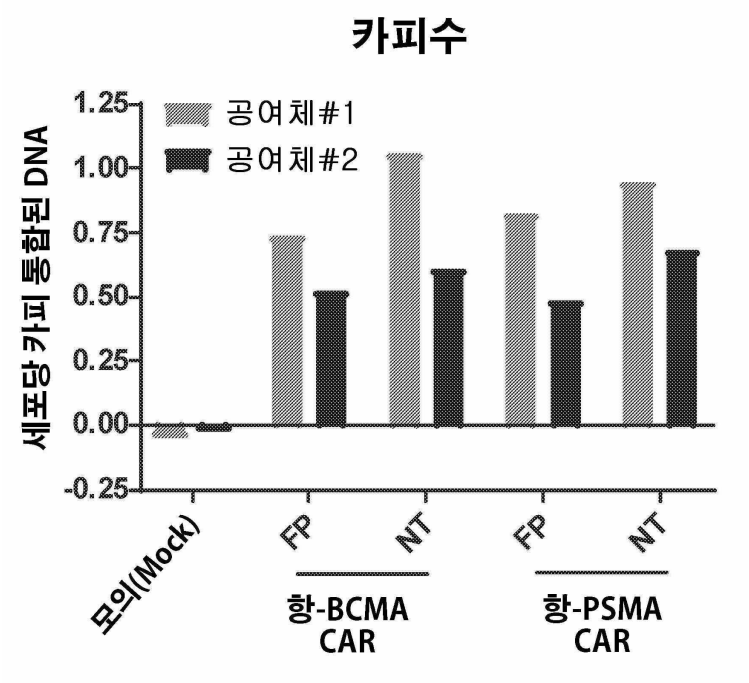
도면6


공여체 #1
공여체 #2
K562. BCMA
K562. PSMA
% 세포독성
E:T 비율
40:1
20:1
10:1
5:1
2.5:1
1.25:1
0.625:1
Mock
항-BCMA CAR
항-PSMA CAR
FP
MT

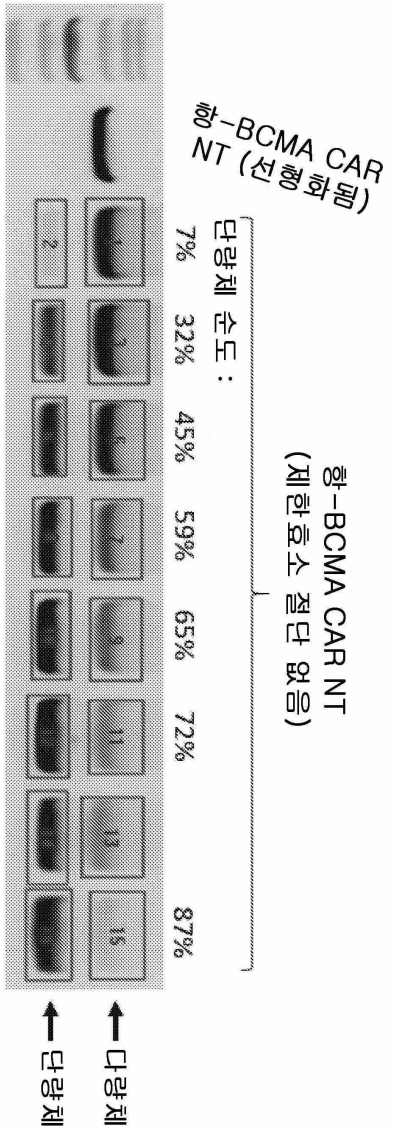
도면7


CD8+ (%)
0
10
20
30
40
50
60
70
80
90
100
항-BCMA CAR
항-PSMA CAR
FR
NT
FR
NT
공여체 #1
CD8+ (%)
항-BCMA CAR
항-PSMA CAR
FR
NT
FR
NT
공여체 #2
Teff
Tem
Tcm
Tscm

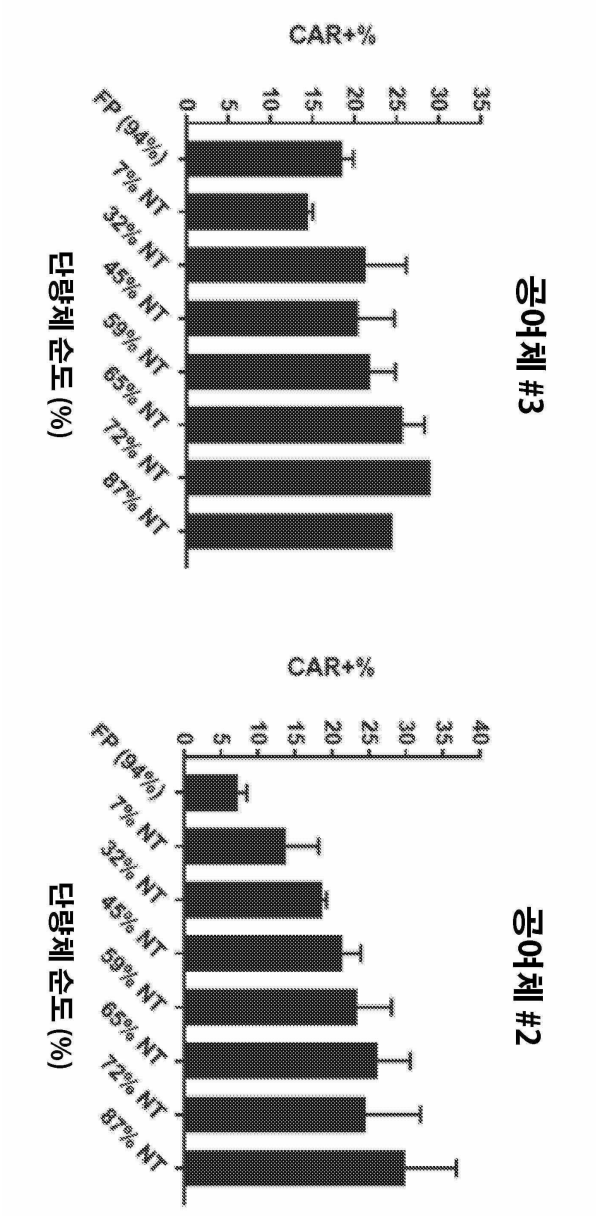
도면8


카피수
세포당 카피 통합된 DNA
1.25
1.00
0.75
0.50
0.25
0.00
-0.25
공여체#1
공여체#2
모의(Mock)
FP
NT
FP
NT
항-BCMA CAR
항-PSMA CAR

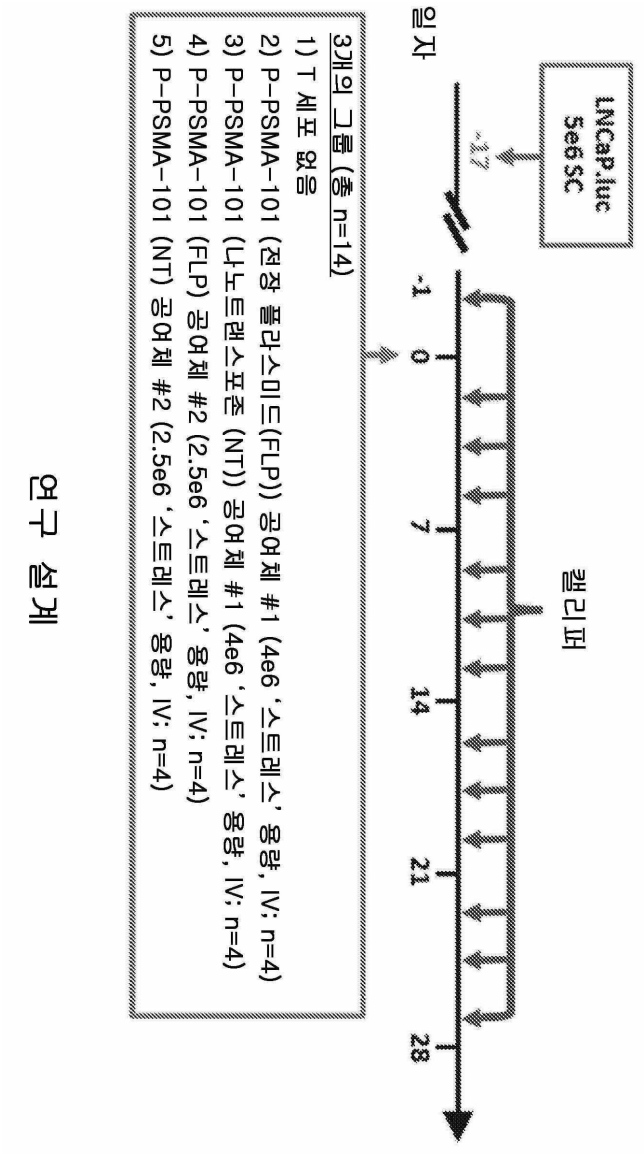
도면9


항-BCMA CAR NT (선형화됨)
항-BCMA CAR NT (제한효소 절단 없음)
단량체 순도 :
7%
32%
45%
59%
65%
72%
87%
2
4
6
8
10
11
13
15
단량체
다량체

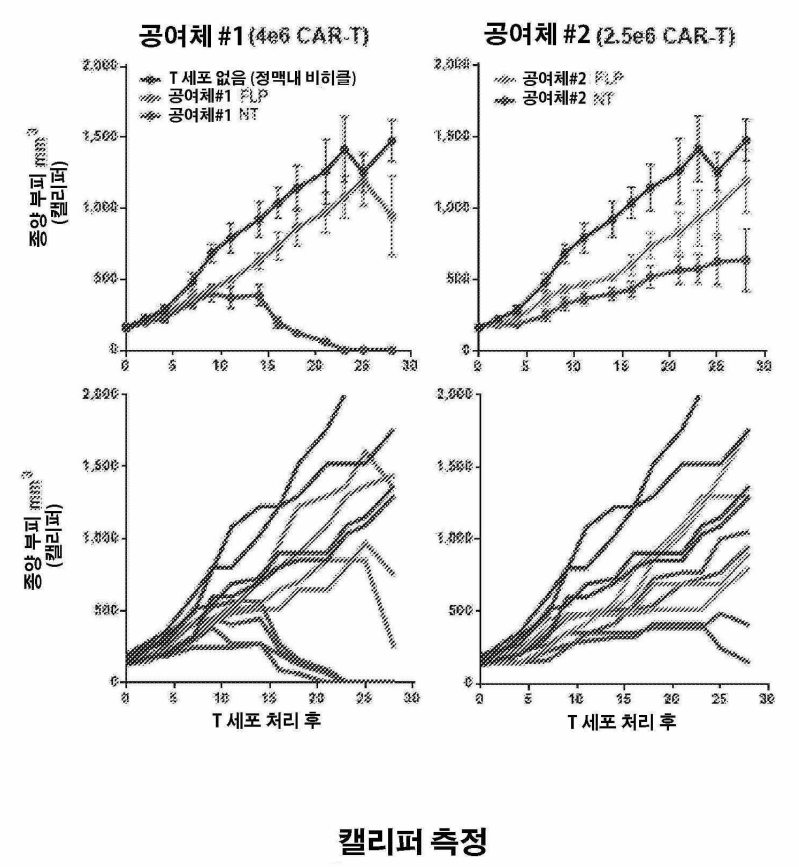
도면10


CAR+%
0
5
10
15
20
25
30
35
FP (94%)
7% NT
32% NT
45% NT
59% NT
65% NT
72% NT
87% NT
단량체 순도 (%)
공여체 #3
CAR+%
0
5
10
15
20
25
30
35
40
FP (94%)
7% NT
32% NT
45% NT
59% NT
65% NT
72% NT
87% NT
단량체 순도 (%)
공여체 #2

도면11

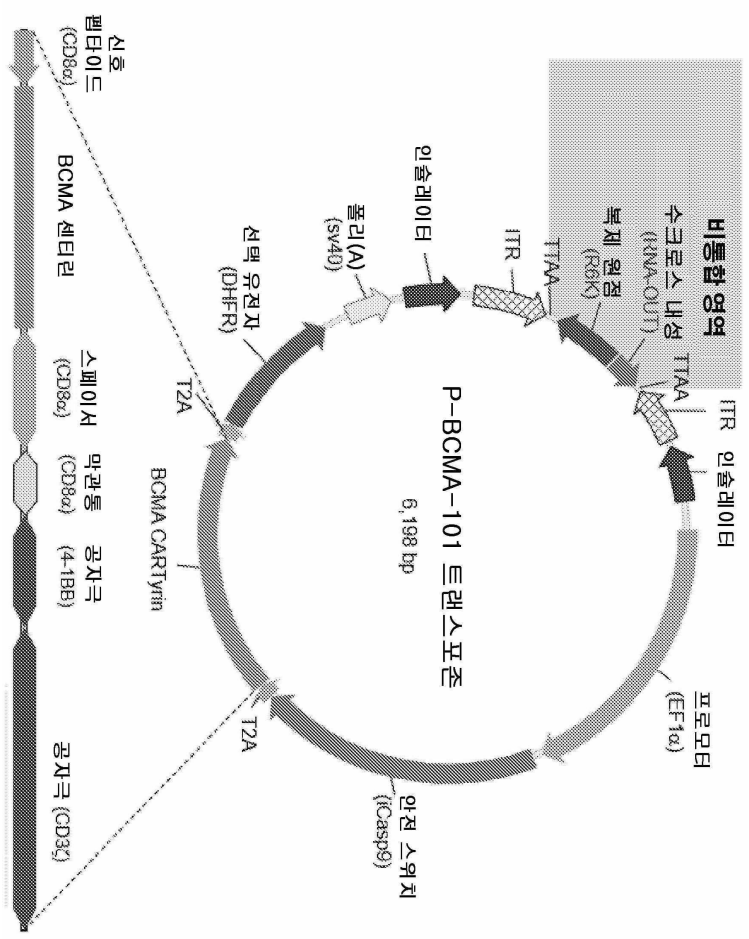

연구 설계
LNCaP.luc
5e6 SC
-17
일자
-1
0
7
14
21
28
캘리퍼
3개의 그룹 (총 n=14)
1) T 세포 없음
2) P-PSMA-101 (전장 플라스미드(FLP)) 공여체 #1 (4e6 '스트레스' 용량, IV; n=4)
3) P-PSMA-101 (나노트랜스포존 (NT)) 공여체 #1 (4e6 '스트레스' 용량, IV; n=4)
4) P-PSMA-101 (FLP) 공여체 #2 (2.5e6 '스트레스' 용량, IV; n=4)
5) P-PSMA-101 (NT) 공여체 #2 (2.5e6 '스트레스' 용량, IV; n=4)

도면12

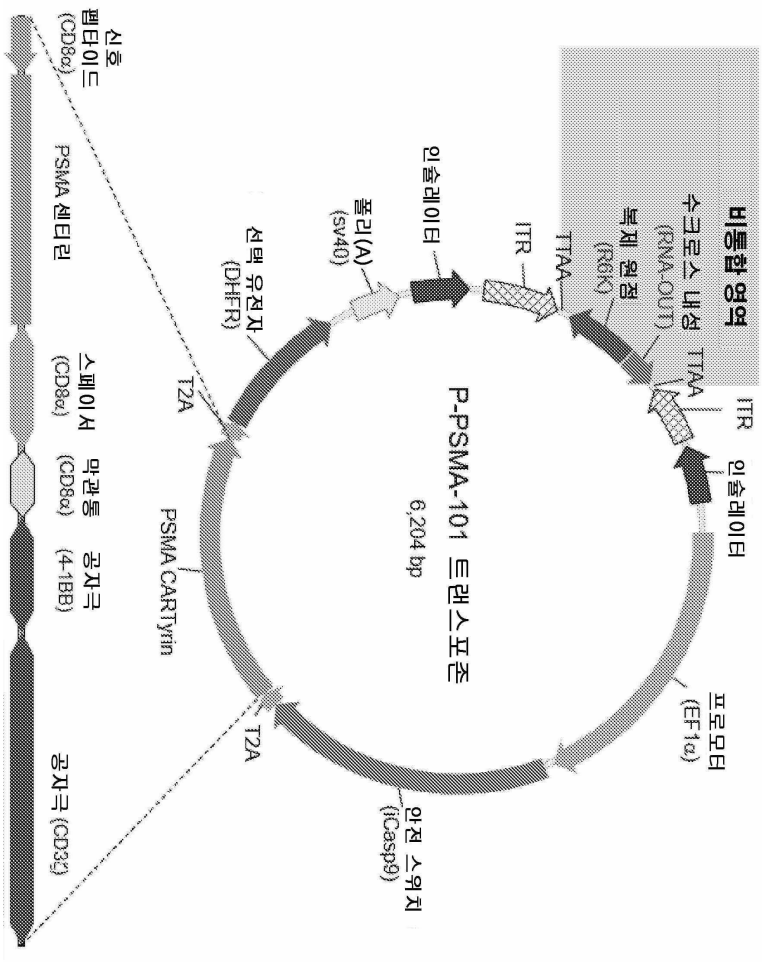

공여체 #1 (4e6 CAR-T)
공여체 #2 (2.5e6 CAR-T)
T 세포 없음 (정맥내 비히클)
공여체#1 FLP
공여체#1 NT
공여체#2 FLP
공여체#2 NT
종양 부피 (캘리퍼)
T 세포 처리 후

캘리퍼 측정

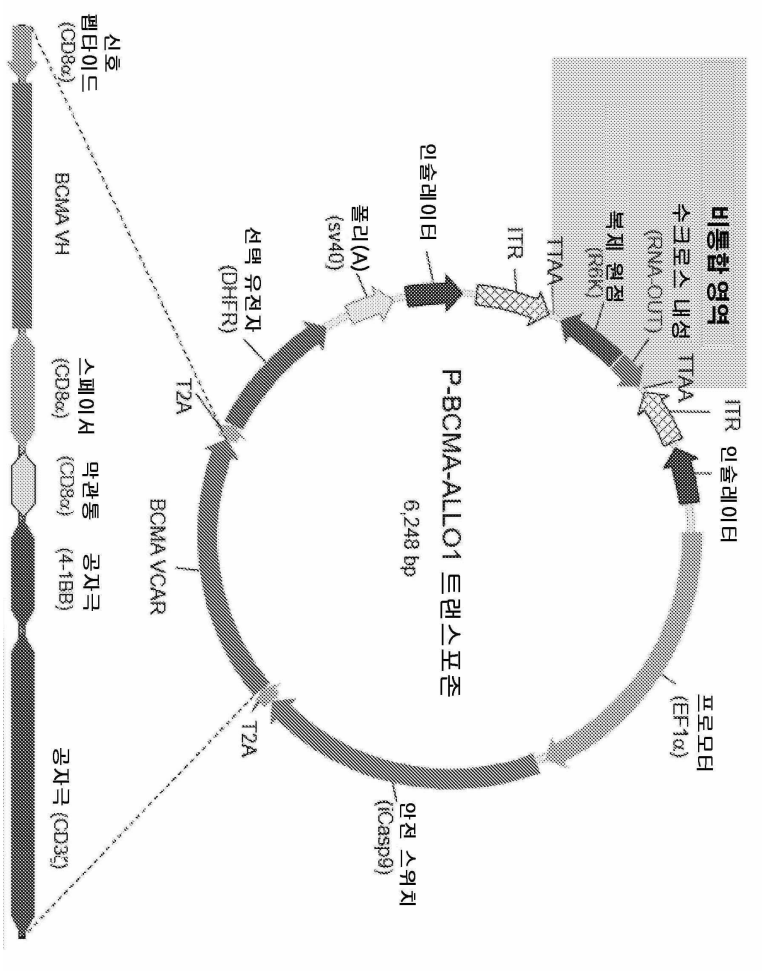
도면13


P-BCMA-101 트랜스포존
6,198 bp
비통합 영역
수크로스 내성 (RNA-OUT)
복제 원점 (R6K)
TTAA
ITR
인슐레이터
폴리(A) (sv40)
선택 유전자 (DHFR)
T2A
BCMA CARTyrin
T2A
안전 스위치 (iCasp9)
프로모터 (EF1α)
ITR
인슐레이터
신호 펩타이드 (CD8α)
BCMA 센티린
스페이서 (CD8α)
막관통 (CD8α)
공자극 (4-1BB)
공자극 (CD3ζ)

도면14


비통합 영역
수크로스 내성 (RNA-OUT)
복제 원점 (R6K)
TTAA
ITR
인슐레이터
프로모터 (EF1α)
안전 스위치 (iCasp9)
T2A
PSMA CARTyrin
선택 유전자 (DHFR)
폴리(A) (sv40)
인슐레이터
ITR
TTAA
P-PSMA-101 트랜스포존
6,204 bp
신호 펩타이드 (CD8α)
PSMA 센티린
스페이서 (CD8α)
막관통 (CD8α)
공자극 (4-1BB)
공자극 (CD3ζ)

도면15




## 서 열 목 록

SEQUENCE LISTING

<110> Poseida Therapeutics, Inc.

<120> NANOTRANSPOSON COMPOSITIONS AND METHODS OF USE

<130> POTH-047/001WO (325002-2401)

<150> 62/783,133

<151> 2018-12-20

<150> 62/815,335

<151> 2019-03-07

<150> 62/815,845

<151> 2019-03-08

<160> 98

<170> PatentIn version 3.5

<210> 1

<211> 6198

<212> DNA

<213> Artificial Sequence

<220><223> P-BCMA-101 nanotransposon expressing a BCMA CARTyrin

<400> 1

tgtacataga ttaaccctag aaagataatc atattgtgac gtacgttaaa gataatcatg 60

cgtaaaattg acgcatgtgt tttatcggtc tgtatatcga ggtttattta ttaatttgaa 120

tagatattaa gttttattat atttacactt acatactaat aataaattca acaaacaatt 180

tatttatgtt tatttattta ttaaaaaaaa acaaaaactc aaaatttctt ctataaagta 240

acaaaacttt tatcgaatac ctgcagcccg ggggatgcag agggacagcc ccccccaaaa 300

gcccccaggg atgtaattac gtccctcccc cgctaggggg cagcagcgag ccgcccgggg 360

ctccgctccg gtccggcgct cccccgcat ccccgagccg gcagcgtgcg gggacagccc 420

gggcacgggg aaggtggcac gggatcgctt tcctctgaac gcttctcgct gctctttgag 480

cctgcagaca cctgggggga tacggggaaa agttgactgt gcctttcgat cgaaccatgg 540

acagttagct ttgcaaagat ggataaagtt ttaaacagag aggaatcttt gcagctaatg 600

gaccttctag gtcttgaaag gagtgggaat tggctccggt gcccgtcagt gggcagagcg 660

cacatcgccc acagtccccg agaagttggg gggaggggtc ggcaattgaa ccggtgccta 720

gagaaggtgg cgcggggtaa actgggaaag tgatgtcgtg tactggctcc gccttttttcc 780

cgagggtggg ggagaaccgt atataagtgc agtagtcgcc gtgaacgttc tttttcgcaa 840

cgggtttgcc gccagaacac aggtaagtgc cgtgtgtggt tcccgcgggc ctggcctctt 900

tacgggttat ggcccttgcg tgccttgaat tacttccacc tggctgcagt acgtgattct 960

tgatcccgag cttcgggttg gaagtgggtg ggagagttcg aggccttgcg cttaaggagc 1020

cccttcgcct cgtgcttgag ttgaggcctg gcctgggcgc tggggccgcc gcgtgcgaat 1080

ctggtggcac cttcgcgcct gtctcgctgc tttcgataag tctctagcca tttaaaattt 1140

ttgatgacct gctgcgacgc tttttttctg gcaagatagt cttgtaaatg cgggccaaga 1200

tctgcacact ggtatttcgg tttttggggc cgcgggcggc gacggggccc gtgcgtccca 1260

gcgcacatgt tcggcgaggc ggggcctgcg agcgcggcca ccgagaatcg gacgggggta 1320

gtctcaagct ggccggcctg ctctggtgcc tggcctcgcg ccgccgtgta tcgccccgcc 1380

ctgggcggca aggctggccc ggtcggcacc agttgcgtga gcggaaagat ggccgcttcc 1440

cggccctgct gcagggagct caaaatggag gacgcggcgc tcgggagagc gggcgggtga 1500

gtcacccaca caaaggaaaa gggcctttcc gtcctcagcc gtcgcttcat gtgactccac 1560

ggagtaccgg gcgccgtcca ggcacctcga ttagttctcg agcttttgga gtacgtcgtc 1620

tttaggttgg ggggaggggt tttatgcgat ggagtttccc cacactgagt gggtggagac 1680

tgaagttagg ccagcttggc acttgatgta attctccttg gaatttgccc tttttgagtt 1740

tggatcttgg ttcattctca agcctcagac agtggttcaa agtttttttc ttccatttca 1800

ggtgtcgtga gaattctaat acgactcact atagggtgtg ctgtctcatc attttggcaa 1860

agattggcca ccaagcttgc caccatgggg gtccaggtcg agactatttc accaggggat 1920

gggcgaacat ttccaaaaag gggccagact tgcgtcgtgc attacaccgg gatgctggag 1980

gacgggaaga aagtggacag ctccagggat cgcaacaagc ccttcaagtt catgctggga 2040

aagcaggaag tgatccgagg atgggaggaa ggcgtggcac agatgtcagt cggccagcgg 2100

gccaaactga ccattagccc tgactacgct tatggagcaa caggccaccc agggatcatt 2160

ccccctcatg ccaccctggt cttcgatgtg gaactgctga agctggaggg aggaggagga 2220

tccggatttg gggacgtggg ggccctggag tctctgcgag gaaatgccga tctggcttac 2280

atcctgagca tggaaccctg cggccactgt ctgatcatta acaatgtgaa cttctgcaga 2340

gaaagcggac tgcgaacacg gactggctcc aatattgact gtgagaagct gcggagaagg 2400

ttctctagtc tgcactttat ggtcgaagtg aaaggggatc tgaccgccaa gaaaatggtg 2460

ctggccctgc tggagctggc tcagcaggac catggagctc tggattgctg cgtggtcgtg 2520

atcctgtccc acgggtgcca ggcttctcat ctgcagttcc ccggagcagt gtacggaaca 2580

gacggctgtc ctgtcagcgt ggagaagatc gtcaacatct tcaacggcac ttcttgccct 2640

agtctggggg gaaagccaaa actgttcttt atccaggcct gtggcgggga acagaaagat 2700

cacggcttcg aggtggccag caccagccct gaggacgaat caccagggag caaccctgaa 2760

ccagatgcaa ctccattcca ggagggactg aggacctttg accagctgga tgctatctca 2820

agcctgccca ctcctagtga cattttcgtg tcttacagta ccttcccagg ctttgtctca 2880

tggcgcgatc ccaagtcagg gagctggtac gtggagacac tggacgacat ctttgaacag 2940

tgggcccatt cagaggacct gcagagcctg ctgctgcgag tggcaaacgc tgtctctgtg 3000

aagggcatct acaaacagat gcccgggtgc ttcaattttc tgagaaagaa actgttcttt 3060

aagacttccg gatctggaga gggaagggga agcctgctga cctgtggaga cgtggaggaa 3120

aacccaggac caatggcact gccagtcacc gccctgctgc tgcctctggc tctgctgctg 3180

cacgcagcta gaccaatgct gcctgcacca aagaacctgg tggtgagccg gatcacagag 3240

gactccgcca gactgtcttg gaccgcccct gacgccgcct tcgattcctt tccaatccgg 3300

tacatcgaga cactgatctg gggcgaggcc atctggctgg acgtgcccgg ctctgagagg 3360

agctacgatc tgacaggcct gaagcctggc accgagtatg cagtggtcat cacaggagtg 3420

aagggcggca ggttcagctc ccctctggtg gcctctttta ccacaaccac aacccctgcc 3480
cccagacctc ccacacccgc ccctaccatc gcgagtcagc ccctgagtct gagacctgag 3540
gcctgcaggc cagctgcagg aggagctgtg cacaccaggg gcctggactt cgcctgcgac 3600
atctacattt gggcaccact ggccgggacc tgtggagtgc tgctgctgag cctggtcatc 3660
acactgtact gcaagagagg caggaagaaa ctgctgtata ttttcaaaca gcccttcatg 3720
cgccccgtgc agactaccca ggaggaagac gggtgctcct gtcgattccc tgaggaagag 3780
gaaggcgggt gtgagctgcg cgtgaagttt agtcgatcag cagatgcccc agcttacaaa 3840

cagggacaga accagctgta taacgagctg aatctgggcc gccgagagga atatgacgtg 3900
ctggataagc ggagaggacg cgaccccgaa atgggaggca agcccaggcg caaaaaccct 3960
caggaaggcc tgtataacga gctgcagaag gacaaaatgg cagaagccta ttctgagatc 4020
ggcatgaagg gggagcgacg gagaggcaaa gggcacgatg ggctgtacca gggactgagc 4080
accgccacaa aggacaccta tgatgctctg catatgcagg cactgcctcc aaggggaagt 4140
ggagaaggac gaggatcact gctgacatgc ggcgacgtgg aggaaaaccc tggcccaatg 4200
gtcgggtctc tgaattgtat cgtcgccgtg agtcagaaca tgggcattgg gaagaatggc 4260

gatttcccat ggccacctct gcgcaacgag tcccgatact ttcagcggat gacaactacc 4320
tcctctgtgg aagggaaaca gaatctggtc atcatgggaa agaaaacttg gttcagcatt 4380
ccagagaaga accggcccct gaaaggcaga atcaatctgg tgctgtcccg agaactgaag 4440
gagccaccac agggagctca ctttctgagc cggtccctgg acgatgcact gaagctgaca 4500
gaacagcctg agctggccaa caaagtcgat atggtgtgga tcgtcggggg aagttcagtg 4560
tataaggagg ccatgaatca ccccggccat ctgaaactgt tcgtcacacg gatcatgcag 4620
gactttgaga gcgatacttt ctttcctgaa attgacctgg agaagtacaa actgctgccc 4680

gaatatcctg gcgtgctgtc cgatgtccag gaagagaaag gcatcaaata caagttcgag 4740
gtctatgaga agaatgacta ataaggtacc gatcacatat gcctttaatt aaacactagt 4800
tctatagtgt cacctaaatt ccctttagtg agggttaatg gccgtaggcc gccagaattg 4860
ggtccagaca tgataagata cattgatgag tttggacaaa ccacaactag aatgcagtga 4920
aaaaaatgct ttatttgtga aatttgtgat gctattgctt tatttgtaac cattataagc 4980
tgcaataaac aagttaacaa caacaattgc attcatttta tgtttcaggt tcaggggggag 5040
gtgtgggagg ttttttcgga ctctaggacc tgcgcatgcg cttggcgtaa tcatggtcat 5100

agctgtttcc tgttttcccc gtatcccccc aggtgtctgc aggctcaaag agcagcgaga 5160

agcgttcaga ggaaagcgat cccgtgccac cttccccgtg cccgggctgt ccccgcacgc 5220

tgccggctcg gggatgcggg gggagcgccg gaccggagcg gagccccggg cggctcgctg 5280

ctgcccccta gcgggggagg gacgtaatta catccctggg ggctttgggg gggggctgtc 5340

cctctcaccg cggtggagct ccagcttttg ttcgaattgg ggccccccct cgagggtatc 5400

gatgatatct ataacaagaa aatatatata taataagtta tcacgtaagt agaacatgaa 5460

ataacaatat aattatcgta tgagttaaat cttaaaagtc acgtaaaaga taatcatgcg 5520

tcattttgac tcacgcggtc gttatagttc aaaatcagtg acacttaccg cattgacaag 5580

cacgcctcac gggagctcca agcggcgact gagatgtcct aaatgcacag cgacggattc 5640

gcgctattta gaaagagaga gcaatatttc aagaatgcat gcgtcaattt tacgcagact 5700

atctttctag ggttaatcta gctagcctta agggcgcagc ccgcctaatg agcgggcttt 5760

tttttggctt gttgtccaca accgttaaac cttaaaagct ttaaaagcct tatatattct 5820

tttttttctt ataaaactta aaaccttaga ggctatttaa gttgctgatt tatattaatt 5880

ttattgttca aacatgagag cttagtacgt gaaacatgag agcttagtac gttagccatg 5940

agagcttagt acgttagcca tgagggttta gttcgttaaa catgagagct tagtacgtta 6000

aacatgagag cttagtacgt actatcaaca ggttgaactg ctgatccacg ttgtggtaga 6060

attggtaaag agagtcgtgt aaaatatcga gttcgcacat cttgttgtct gattattgat 6120

ttttggcgaa accatttgat catatgacaa gatgtgtatc taccttaact taatgatttt 6180

gataaaaatc attaggta 6198

<210> 2

<211> 238

<212> DNA

<213> Artificial Sequence

<220><223> ITR

<400> 2

ccctagaaag ataatcatat tgtgacgtac gttaaagata atcatgcgta aaattgacgc 60

atgtgtttta tcggtctgta tatcgaggtt tatttattaa tttgaataga tattaagttt 120

tattatattt acacttacat actaataata aattcaacaa acaatttatt tatgtttatt 180

tatttattaa aaaaaaacaa aaactcaaaa tttcttctat aaagtaacaa aactttta 238

<210> 3

<211> 232

<212> DNA

<213> Artificial Sequence

<220><223> Insulator

<400> 3

gagggacagc ccccccccaa agcccccagg gatgtaatta cgtccctccc ccgctagggg 60

gcagcagcga gccgcccggg gctccgctcc ggtccggcgc tccccccgca tccccgagcc 120

ggcagcgtgc ggggacagcc cgggcacggg gaaggtggca cgggatcgct ttcctctgaa 180

cgcttctcgc tgctctttga gcctgcagac acctgggggg atacggggaa aa 232

<210> 4

<211> 1264

<212> DNA

<213> Artificial Sequence

<220><223> EFlalpha promoter

<400> 4

agctttgcaa agatggataa agttttaaac agagaggaat ctttgcagct aatggacctt 60

ctaggtcttg aaaggagtgg gaattggctc cggtgcccgt cagtgggcag agcgcacatc 120

gcccacagtc cccgagaagt tgggggggagg ggtcggcaat tgaaccggtg cctagagaag 180

gtggcgcggg gtaaactggg aaagtgatgt cgtgtactgg ctccgccttt ttcccgaggg 240

tgggggagaa ccgtatataa gtgcagtagt cgccgtgaac gttcttttttc gcaacgggtt 300

tgccgccaga acacaggtaa gtgccgtgtg tggttcccgc gggcctggcc tctttacggg 360

ttatggccct tgcgtgcctt gaattacttc cacctggctg cagtacgtga ttcttgatcc 420

cgagcttcgg gttggaagtg ggtgggagag ttcgaggcct tgcgcttaag gagcccctttc 480

gcctcgtgct tgagttgagg cctggcctgg gcgctggggc cgccgcgtgc gaatctggtg 540

gcaccttcgc gcctgtctcg ctgctttcga taagtctcta gccatttaaa atttttgatg 600

acctgctgcg acgcttttttt tctggcaaga tagtcttgta aatgcgggcc aagatctgca 660

cactggtatt tcggtttttg gggccgcggg cggcgacggg gcccgtgcgt cccagcgcac 720

atgttcggcg aggcggggcc tgcgagcgcg gccaccgaga atcggacggg ggtagtctca 780

agctggccgg cctgctctgg tgcctggcct cgcgccgccg tgtatcgccc cgccctgggc 840

ggcaaggctg gcccggtcgg caccagttgc gtgagcggaa agatggccgc ttcccggccc 900

tgctgcaggg agctcaaaat ggaggacgcg gcgctcggga gagcgggcgg gtgagtcacc 960

cacacaaagg aaaagggcct ttccgtcctc agccgtcgct tcatgtgact ccacggagta 1020

ccgggcgccg tccaggcacc tcgattagtt ctcgagcttt tggagtacgt cgtctttagg 1080

ttgggggggag gggttttatg cgatggagtt tccccacact gagtgggtgg agactgaagt 1140

taggccagct tggcacttga tgtaattctc cttggaattt gccctttttg agtttggatc 1200

ttggttcatt ctcaagcctc agacagtggt tcaaagtttt tttcttccat ttcaggtgtc 1260

gtga 1264

<210> 5

<211> 1185

<212> DNA

<213> Artificial Sequence

<220><223> Inducible proapoptotic polypeptides

<400> 5

atgggggtcc aggtcgagac tatttcacca ggggatgggc gaacatttcc aaaaaggggc 60

cagacttgcg tcgtgcatta caccgggatg ctggaggacg ggaagaaagt ggacagctcc 120

agggatcgca acaagccctt caagttcatg ctgggaaagc aggaagtgat ccgaggatgg 180

gaggaaggcg tggcacagat gtcagtcggc cagcgggcca aactgaccat tagccctgac 240

tacgcttatg gagcaacagg ccacccaggg atcattcccc ctcatgccac cctggtcttc 300

gatgtggaac tgctgaagct ggagggagga ggaggatccg gatttgggga cgtgggggcc 360

ctggagtctc tgcgaggaaa tgccgatctg gcttacatcc tgagcatgga accctgcggc 420

cactgtctga tcattaacaa tgtgaacttc tgcagagaaa gcggactgcg aacacggact 480

ggctccaata ttgactgtga gaagctgcgg agaaggttct ctagtctgca ctttatggtc 540

gaagtgaaag gggatctgac cgccaagaaa atggtgctgg ccctgctgga gctggctcag 600

caggaccatg gagctctgga ttgctgcgtg gtcgtgatcc tgtcccacgg gtgccaggct 660

tctcatctgc agttccccgg agcagtgtac ggaacagacg gctgtcctgt cagcgtggag 720

aagatcgtca acatcttcaa cggcacttct tgccctagtc tgggggggaaa gccaaaactg 780

ttctttatcc aggcctgtgg cggggaacag aaagatcacg gcttcgaggt ggccagcacc 840

agccctgagg acgaatcacc agggagcaac cctgaaccag atgcaactcc attccaggag 900

ggactgagga cctttgacca gctggatgct atctcaagcc tgcccactcc tagtgacatt 960

ttcgtgtctt acagtacctt cccaggcttt gtctcatggc gcgatcccaa gtcagggagc 1020

tggtacgtgg agacactgga cgacatcttt gaacagtggg cccattcaga ggacctgcag 1080

agcctgctgc tgcgagtggc aaacgctgtc tctgtgaagg gcatctacaa acagatgccc 1140

gggtgcttca attttctgag aaagaaactg ttctttaaga cttcc 1185

<210> 6

<211> 1185

<212> DNA

<213> Artificial Sequence

<220><223> Inducible proapoptotic polypeptides

<400> 6

atgggggtcc aggtcgagac tatttcacca ggggatgggc gaacatttcc aaaaaggggc 60

cagacttgcg tcgtgcatta caccgggatg ctggaggacg ggaagaaagt ggacagctcc 120

agggatcgca acaagccctt caagttcatg ctgggaaagc aggaagtgat ccgaggatgg 180

gaggaaggcg tggcacagat gtcagtcggc cagcgggcca aactgaccat tagccctgac 240

tacgcttatg gagcaacagg ccacccaggg atcattcccc ctcatgccac cctggtcttc 300

gatgtggaac tgctgaagct ggagggagga ggaggatccg gatttgggga cgtgggggcc 360

ctggagtctc tgcgaggaaa tgccgatctg gcttacatcc tgagcatgga accctgcggc 420

cactgtctga tcattaacaa tgtgaacttc tgcagagaaa gcggactgcg aacacggact 480

ggctccaata ttgactgtga gaagctgcgg agaaggttct ctagtctgca ctttatggtc 540

gaagtgaaag gggatctgac cgccaagaaa atggtgctgg ccctgctgga gctggctcag 600

caggaccatg gagctctgga ttgctgcgtg gtcgtgatcc tgtcccacgg gtgccaggct 660

tctcatctgc agttccccgg agcagtgtac ggaacagacg gctgtcctgt cagcgtggag 720

aagatcgtca acatcttcaa cggcacttct tgccctagtc tgggggggaaa gccaaaactg 780

ttctttatcc aggcctgtgg cggggaacag aaagatcacg gcttcgaggt ggccagcacc 840

agccctgagg acgaatcacc agggagcaac cctgaaccag atgcaactcc attccaggag 900

ggactgagga cctttgacca gctggatgct atctcaagcc tgcccactcc tagtgacatt 960

ttcgtgtctt acagtacctt cccaggcttt gtctcatggc gcgatcccaa gtcagggagc 1020

tggtacgtgg agacactgga cgacatcttt gaacagtggg cccattcaga ggacctgcag 1080

agcctgctgc tgcgagtggc aaacgctgtc tctgtgaagg gcatctacaa acagatgccc 1140

gggtgcttca attttctgag aaagaaactg ttctttaaga cttcc 1185

<210> 7

<211> 63

<212> DNA

<213> Artificial Sequence

<220><223> T2A Sequence

<400> 7

ggatctggag agggaagggg aagcctgctg acctgtggag acgtggagga aaacccagga 60

cca 63

<210> 8

<211> 63

<212> DNA

<213> Artificial Sequence

<220><223> T2A Sequence

<400> 8

ggatctggag agggaagggg aagcctgctg acctgtggag acgtggagga aaacccagga 60

cca 63

<210> 9

<211> 1002

<212> DNA

<213> Artificial Sequence

<220><223> BCMA CARTyrin

<400> 9

atggcactgc cagtcaccgc cctgctgctg cctctggctc tgctgctgca cgcagctaga 60

ccaatgctgc ctgcaccaaa gaacctggtg gtgagccgga tcacagagga ctccgccaga 120

ctgtcttgga ccgcccctga cgccgccttc gattcctttc caatccggta catcgagaca 180

ctgatctggg gcgaggccat ctggctggac gtgcccggct ctgagaggag ctacgatctg 240

acaggcctga agcctggcac cgagtatgca gtggtcatca caggagtgaa gggcggcagg 300

ttcagctccc ctctggtggc ctcttttacc acaaccacaa cccctgcccc cagacctccc 360

acacccgccc ctaccatcgc gagtcagccc ctgagtctga gacctgaggc ctgcaggcca 420

gctgcaggag gagctgtgca caccaggggc ctggacttcg cctgcgacat ctacatttgg 480

gcaccactgg ccgggacctg tggagtgctg ctgctgagcc tggtcatcac actgtactgc 540

aagagaggca ggaagaaact gctgtatatt ttcaaacagc ccttcatgcg ccccgtgcag 600

actacccagg aggaagacgg gtgctcctgt cgattccctg aggaagagga aggcgggtgt 660

gagctgcgcg tgaagtttag tcgatcagca gatgccccag cttacaaaca gggacagaac 720

cagctgtata acgagctgaa tctgggccgc cgagaggaat atgacgtgct ggataagcgg 780

agaggacgcg accccgaaat gggaggcaag cccaggcgca aaaaccctca ggaaggcctg 840

tataacgagc tgcagaagga caaaatggca gaagcctatt ctgagatcgg catgaagggg 900

gagcgacgga gaggcaaagg gcacgatggg ctgtaccagg gactgagcac cgccacaaag 960

gacacctatg atgctctgca tatgcaggca ctgcctccaa gg 1002

<210> 10

<211> 63

<212> DNA

<213> Artificial Sequence

<220><223> T2A Sequence

<400> 10

ggaagtggag aaggacgagg atcactgctg acatgcggcg acgtggagga aaaccctggc 60

cca 63

<210> 11

<211> 561

<212> DNA

<213> Artificial Sequence

<220><223> DHFR selection

<400

> 11

atggtcgggt ctctgaattg tatcgtcgcc gtgagtcaga acatgggcat tgggaagaat 60

ggcgatttcc catggccacc tctgcgcaac gagtcccgat actttcagcg gatgacaact 120

acctcctctg tggaagggaa acagaatctg gtcatcatgg gaaagaaaac ttggttcagc 180

attccagaga agaaccggcc cctgaaaggc agaatcaatc tggtgctgtc ccgagaactg 240

aaggagccac cacagggagc tcactttctg agccggtccc tggacgatgc actgaagctg 300

acagaacagc ctgagctggc caacaaagtc gatatggtgt ggatcgtcgg gggaagttca 360

gtgtataagg aggccatgaa tcaccccggc catctgaaac tgttcgtcac acggatcatg 420

caggactttg agagcgatac tttctttcct gaaattgacc tggagaagta caaactgctg 480

cccgaatatc ctggcgtgct gtccgatgtc caggaagaga aaggcatcaa atacaagttc 540

gaggtctatg agaagaatga c 561

<210> 12

<211> 127

<212> DNA

<213> Artificial Sequence

<220><223> PolyA sv40

<400> 12

cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa 60

aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca 120

ataaaca 127

<210> 13

<211> 231

<212> DNA

<213> Artificial Sequence

<220><223> Insulator

<400> 13

ttttccccgt atccccccag gtgtctgcag gctcaaagag cagcgagaag cgttcagagg 60

aaagcgatcc cgtgccacct tccccgtgcc cgggctgtcc ccgcacgctg ccggctcggg 120

gatgcggggg gagcgccgga ccggagcgga gccccgggcg gctcgctgct gccccctagc 180

gggggaggga cgtaattaca tccctggggg ctttgggggg gggctgtccc t 231

<210> 14

<211> 309

<212> DNA

<213> Artificial Sequence

<220><223> ITR

<400> 14

gatatctata acaagaaaat atatatataa taagttatca cgtaagtaga acatgaaata 60

acaatataat tatcgtatga gttaaatctt aaaagtcacg taaaagataa tcatgcgtca 120

ttttgactca cgcggtcgtt atagttcaaa atcagtgaca cttaccgcat tgacaagcac 180

gcctcacggg agctccaagc ggcgactgag atgtcctaaa tgcacagcga cggattcgcg 240

ctatttagaa agagagagca atatttcaag aatgcatgcg tcaattttac gcagactatc 300

tttctaggg 309

<210> 15

<211> 281

<212> DNA

<213> Artificial Sequence

<220><223> R6K Origin of Replication

<400> 15

ggcttgttgt ccacaaccgt taaaccttaa aagctttaaa agccttatat attcttttt 60

ttcttataaa acttaaaacc ttagaggcta tttaagttgc tgatttatat taattttatt 120

gttcaaacat gagagcttag tacgtgaaac atgagagctt agtacgttag ccatgagagc 180

ttagtacgtt agccatgagg gtttagttcg ttaaacatga gagcttagta cgttaaacat 240

gagagcttag tacgtactat caacaggttg aactgctgat c 281

<210> 16

<211> 139

<212> DNA

<213> Artificial Sequence

<220><223> RNA-OUT

<400> 16

gtagaattgg taaagagagt cgtgtaaaat atcgagttcg cacatcttgt tgtctgatta 60

ttgatttttg gcgaaaccat ttgatcatat gacaagatgt gtatctacct taacttaatg 120

attttgataa aaatcatta 139

<210> 17

<211> 958

<212> PRT

<213> Artificial Sequence

<220><223> P-BCMA-101 amino acid sequence

<400> 17

Met Gly Val Gln Val Glu Thr Ile Ser Pro Gly Asp Gly Arg Thr Phe

1 5 10 15

Pro Lys Arg Gly Gln Thr Cys Val Val His Tyr Thr Gly Met Leu Glu

20 25 30

Asp Gly Lys Lys Val Asp Ser Ser Arg Asp Arg Asn Lys Pro Phe Lys

35 40 45

Phe Met Leu Gly Lys Gln Glu Val Ile Arg Gly Trp Glu Glu Gly Val

50 55 60

Ala Gln Met Ser Val Gly Gln Arg Ala Lys Leu Thr Ile Ser Pro Asp

65 70 75 80

Tyr Ala Tyr Gly Ala Thr Gly His Pro Gly Ile Ile Pro Pro His Ala

85 90 95

Thr Leu Val Phe Asp Val Glu Leu Leu Lys Leu Glu Gly Gly Gly Gly

100 105 110

Ser Gly Phe Gly Asp Val Gly Ala Leu Glu Ser Leu Arg Gly Asn Ala
115 120 125
Asp Leu Ala Tyr Ile Leu Ser Met Glu Pro Cys Gly His Cys Leu Ile
130 135 140
Ile Asn Asn Val Asn Phe Cys Arg Glu Ser Gly Leu Arg Thr Arg Thr
145 150 155 160
Gly Ser Asn Ile Asp Cys Glu Lys Leu Arg Arg Arg Phe Ser Ser Leu

165 170 175
His Phe Met Val Glu Val Lys Gly Asp Leu Thr Ala Lys Lys Met Val
180 185 190
Leu Ala Leu Leu Glu Leu Ala Gln Gln Asp His Gly Ala Leu Asp Cys
195 200 205
Cys Val Val Val Ile Leu Ser His Gly Cys Gln Ala Ser His Leu Gln
210 215 220
Phe Pro Gly Ala Val Tyr Gly Thr Asp Gly Cys Pro Val Ser Val Glu

225 230 235 240
Lys Ile Val Asn Ile Phe Asn Gly Thr Ser Cys Pro Ser Leu Gly Gly
245 250 255
Lys Pro Lys Leu Phe Phe Ile Gln Ala Cys Gly Gly Glu Gln Lys Asp
260 265 270
His Gly Phe Glu Val Ala Ser Thr Ser Pro Glu Asp Glu Ser Pro Gly
275 280 285
Ser Asn Pro Glu Pro Asp Ala Thr Pro Phe Gln Glu Gly Leu Arg Thr

290 295 300
Phe Asp Gln Leu Asp Ala Ile Ser Ser Leu Pro Thr Pro Ser Asp Ile
305 310 315 320
Phe Val Ser Tyr Ser Thr Phe Pro Gly Phe Val Ser Trp Arg Asp Pro
325 330 335
Lys Ser Gly Ser Trp Tyr Val Glu Thr Leu Asp Asp Ile Phe Glu Gln
340 345 350
Trp Ala His Ser Glu Asp Leu Gln Ser Leu Leu Leu Arg Val Ala Asn

355 360 365

Ala Val Ser Val Lys Gly Ile Tyr Lys Gln Met Pro Gly Cys Phe Asn

370 375 380

Phe Leu Arg Lys Lys Leu Phe Phe Lys Thr Ser Gly Ser Gly Glu Gly

385 390 395 400

Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro Gly Pro

405 410 415

Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu

420 425 430

His Ala Ala Arg Pro Met Leu Pro Ala Pro Lys Asn Leu Val Val Ser

435 440 445

Arg Ile Thr Glu Asp Ser Ala Arg Leu Ser Trp Thr Ala Pro Asp Ala

450 455 460

Ala Phe Asp Ser Phe Pro Ile Arg Tyr Ile Glu Thr Leu Ile Trp Gly

465 470 475 480

Glu Ala Ile Trp Leu Asp Val Pro Gly Ser Glu Arg Ser Tyr Asp Leu

485 490 495

Thr Gly Leu Lys Pro Gly Thr Glu Tyr Ala Val Val Ile Thr Gly Val

500 505 510

Lys Gly Gly Arg Phe Ser Ser Pro Leu Val Ala Ser Phe Thr Thr Thr

515 520 525

Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala Ser

530 535 540

Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly Gly

545 550 555 560

Ala Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr Ile Trp

565 570 575

Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu Ser Leu Val Ile

580 585 590

Thr Leu Tyr Cys Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe Lys

595 600 605

Gln Pro Phe Met Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly Cys

610 615 620

Ser Cys Arg Phe Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu Arg Val

625 630 635 640

Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly Gln Asn

645 650 655

Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp Val

660 665 670

Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro Arg

675 680 685

Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp Lys

690 695 700

Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg Arg

705 710 715 720

Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys

725 730 735

Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg Gly Ser

740 745 750

Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn

755 760 765

Pro Gly Pro Met Val Gly Ser Leu Asn Cys Ile Val Ala Val Ser Gln

770 775 780

Asn Met Gly Ile Gly Lys Asn Gly Asp Phe Pro Trp Pro Pro Leu Arg

785 790 795 800

Asn Glu Ser Arg Tyr Phe Gln Arg Met Thr Thr Thr Ser Ser Val Glu

805 810 815

Gly Lys Gln Asn Leu Val Ile Met Gly Lys Lys Thr Trp Phe Ser Ile

820 825 830

Pro Glu Lys Asn Arg Pro Leu Lys Gly Arg Ile Asn Leu Val Leu Ser

835 840 845

Arg Glu Leu Lys Glu Pro Pro Gln Gly Ala His Phe Leu Ser Arg Ser

850 855 860

Leu Asp Asp Ala Leu Lys Leu Thr Glu Gln Pro Glu Leu Ala Asn Lys

865 870 875 880

Val Asp Met Val Trp Ile Val Gly Gly Ser Ser Val Tyr Lys Glu Ala

885 890 895

Met Asn His Pro Gly His Leu Lys Leu Phe Val Thr Arg Ile Met Gln

900 905 910

Asp Phe Glu Ser Asp Thr Phe Phe Pro Glu Ile Asp Leu Glu Lys Tyr

915 920 925

Lys Leu Leu Pro Glu Tyr Pro Gly Val Leu Ser Asp Val Gln Glu Glu

930 935 940

Lys Gly Ile Lys Tyr Lys Phe Glu Val Tyr Glu Lys Asn Asp

945 950 955

<210> 18

<211> 6204

<212> DNA

<213> Artificial Sequence

<220><223> P-PSMA-101 nanotransposon expressing a PSMA CARTyrin

<400> 18

tgtacataga ttaaccctag aaagataatc atattgtgac gtacgttaaa gataatcatg 60

cgtaaaattg acgcatgtgt tttatcggtc tgtatatcga ggtttattta ttaatttgaa 120

tagatattaa gttttattat atttacactt acatactaat aataaattca acaaacaatt 180

tatttatgtt tatttattta ttaaaaaaaa acaaaaactc aaaatttctt ctataaagta 240

acaaaacttt tatcgaatac ctgcagcccg ggggatgcag agggacagcc ccccccaaaa 300

gcccccaggg atgtaattac gtccctcccc cgctaggggg cagcagcgag ccgcccgggg 360

ctccgctccg gtccggcgct cccccccgcat ccccgagccg gcagcgtgcg gggacagccc 420

gggcacgggg aaggtggcac gggatcgctt tcctctgaac gcttctcgct gctctttgag 480

cctgcagaca cctgggggga tacggggaaa agttgactgt gcctttcgat cgaaccatgg 540

acagttagct ttgcaaagat ggataaagtt ttaaacagag aggaatcttt gcagctaatg 600

gaccttctag gtcttgaaag gagtgggaat tggctccggt gcccgtcagt gggcagagcg 660

cacatcgccc acagtccccg agaagttggg gggaggggtc ggcaattgaa ccggtgccta 720

gagaaggtgg cgcggggtaa actgggaaag tgatgtcgtg tactggctcc gccttttttcc 780

cgagggtggg ggagaaccgt atataagtgc agtagtcgcc gtgaacgttc tttttcgcaa 840

cgggtttgcc gccagaacac aggtaagtgc cgtgtgtggt tcccgcgggc ctggcctctt 900

tacgggttat ggcccttgcg tgccttgaat tacttccacc tggctgcagt acgtgattct 960

tgatcccgag cttcgggttg gaagtgggtg ggagagttcg aggccttgcg cttaaggagc 1020

cccttcgcct cgtgcttgag ttgaggcctg gcctgggcgc tggggccgcc gcgtgcgaat 1080

ctggtggcac cttcgcgcct gtctcgctgc tttcgataag tctctagcca tttaaaattt 1140

ttgatgacct gctgcgacgc tttttttctg gcaagatagt cttgtaaatg cgggccaaga 1200

tctgcacact ggtatttcgg tttttggggc cgcgggcggc gacggggccc gtgcgtccca 1260

gcgcacatgt tcggcgaggc ggggcctgcg agcgcggcca ccgagaatcg gacgggggta 1320

gtctcaagct ggccggcctg ctctggtgcc tggcctcgcg ccgccgtgta tcgccccgcc 1380

ctgggcggca aggctggccc ggtcggcacc agttgcgtga gcggaaagat ggccgcttcc 1440

cggccctgct gcagggagct caaaatggag gacgcggcgc tcgggagagc gggcgggtga 1500

gtcacccaca caaaggaaaa gggcctttcc gtcctcagcc gtcgcttcat gtgactccac 1560

ggagtaccgg gcgccgtcca ggcacctcga ttagttctcg agcttttgga gtacgtcgtc 1620

tttaggttgg ggggaggggt tttatgcgat ggagtttccc cacactgagt gggtggagac 1680

tgaagttagg ccagcttggc acttgatgta attctccttg gaatttgccc tttttgagtt 1740

tggatcttgg ttcattctca agcctcagac agtggttcaa agttttttttc ttccatttca 1800

ggtgtcgtga gaattctaat acgactcact atagggtgtg ctgtctcatc attttggcaa 1860

agattggcca ccaagcttgc caccatgggg gtccaggtcg agactatttc accaggggat 1920

gggcgaacat ttccaaaaag gggccagact tgcgtcgtgc attacaccgg gatgctggag 1980

gacgggaaga aagtggacag ctccagggat cgcaacaagc ccttcaagtt catgctggga 2040

aagcaggaag tgatccgagg atgggaggaa ggcgtggcac agatgtcagt cggccagcgg 2100

gccaaactga ccattagccc tgactacgct tatggagcaa caggccaccc agggatcatt 2160

ccccctcatg ccaccctggt cttcgatgtg gaactgctga agctggaggg aggaggagga 2220

tccggatttg gggacgtggg ggccctggag tctctgcgag gaaatgccga tctggcttac 2280

atcctgagca tggaaccctg cggccactgt ctgatcatta acaatgtgaa cttctgcaga 2340

gaaagcggac tgcgaacacg gactggctcc aatattgact gtgagaagct gcggagaagg 2400

ttctctagtc tgcactttat ggtcgaagtg aaaggggatc tgaccgccaa gaaaatggtg 2460

ctggccctgc tggagctggc tcagcaggac catggagctc tggattgctg cgtggtcgtg 2520

atcctgtccc acgggtgcca ggcttctcat ctgcagttcc ccggagcagt gtacggaaca 2580

gacggctgtc ctgtcagcgt ggagaagatc gtcaacatct tcaacggcac ttcttgccct 2640

agtctggggg gaaagccaaa actgttcttt atccaggcct gtggcgggga acagaaagat 2700

cacggcttcg aggtggccag caccagccct gaggacgaat caccagggag caaccctgaa 2760

ccagatgcaa ctccattcca ggagggactg aggacctttg accagctgga tgctatctca 2820

agcctgccca ctcctagtga cattttcgtg tcttacagta ccttcccagg ctttgtctca 2880

tggcgcgatc ccaagtcagg gagctggtac gtggagacac tggacgacat ctttgaacag 2940

tgggcccatt cagaggacct gcagagcctg ctgctgcgag tggcaaacgc tgtctctgtg 3000

aagggcatct acaaacagat gcccgggtgc ttcaattttc tgagaaagaa actgttcttt 3060

aagacttccg gatctggaga gggaaggggа agcctgctga cctgtggaga cgtggaggaa 3120

aacccaggac caatggcact gccagtcacc gccctgctgc tgcctctggc tctgctgctg 3180

cacgcagcta gaccaatgct gcctgcacca aagaacctgg tggtgtctcg ggtgaccgag 3240

gactctgcca gactgagctg ggccatcgac gagcagaggg attggttcga gagctttctg 3300

atccagtatc aggagtccga gaaagtgggc gaggccatcg tgctgacagt gcctggcagc 3360

gagcggtcct atgatctgac cggcctgaag ccaggcacag agtacaccgt gtccatctac 3420

ggcgtgtatc acgtgtacag gtccaatcct ctgtctgcca tcttcaccac aaccacaacc 3480

cctgccccca gacctcccac acccgccccт accatcgcga gtcagcccct gagtctgaga 3540

cctgaggcct gcaggccagc tgcaggagga gctgtgcaca ccaggggcct ggacttcgcc 3600

tgcgacatct acatttgggc accactggcc gggacctgtg gagtgctgct gctgagcctg 3660

gtcatcacac tgtactgcaa gagaggcagg aagaaactgc tgtatatttt caaacagccc 3720

ttcatgcgcc ccgtgcagac tacccaggag gaagacgggt gctcctgtcg attccctgag 3780

gaagaggaag gcgggtgtga gctgcgcgtg aagtttagtc gatcagcaga tgccccagct 3840

tacaaacagg gacagaacca gctgtataac gagctgaatc tgggccgccg agaggaatat 3900

gacgtgctgg ataagcggag aggacgcgac cccgaaatgg gaggcaagcc caggcgcaaa 3960

aaccctcagg aaggcctgta taacgagctg cagaaggaca aaatggcaga agcctattct 4020

gagatcggca tgaaggggga gcgacggaga ggcaaagggc acgatgggct gtaccaggga 4080

ctgagcaccg ccacaaagga cacctatgat gctctgcata tgcaggcact gcctccaagg 4140

ggaagtggag aaggacgagg atcactgctg acatgcggcg acgtggagga aaaccctggc 4200

ccaatggtcg ggtctctgaa ttgtatcgtc gccgtgagtc agaacatggg cattgggaag 4260

aatggcgatt tcccatggcc acctctgcgc aacgagtccc gatactttca gcggatgaca 4320

actacctcct ctgtggaagg gaaacagaat ctggtcatca tgggaaagaa aacttggttc 4380

agcattccag agaagaaccg gcccctgaaa ggcagaatca atctggtgct gtcccgagaa 4440
ctgaaggagc caccacaggg agctcacttt ctgagccggt ccctggacga tgcactgaag 4500
ctgacagaac agcctgagct ggccaacaaa gtcgatatgg tgtggatcgt cgggggaagt 4560
tcagtgtata aggaggccat gaatcacccc ggccatctga aactgttcgt cacacggatc 4620
atgcaggact ttgagagcga tactttcttt cctgaaattg acctggagaa gtacaaactg 4680
ctgcccgaat atcctggcgt gctgtccgat gtccaggaag agaaaggcat caaatacaag 4740
ttcgaggtct atgagaagaa tgactaataa ggtaccgatc acatatgcct ttaattaaac 4800

actagttcta tagtgtcacc taaattccct ttagtgaggg ttaatggccg taggccgcca 4860
gaattgggtc cagacatgat aagatacatt gatgagtttg gacaaaccac aactagaatg 4920
cagtgaaaaa aatgctttat ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt 4980
ataagctgca ataaacaagt taacaacaac aattgcattc attttatgtt tcaggttcag 5040
ggggaggtgt gggaggtttt ttcggactct aggacctgcg catgcgcttg gcgtaatcat 5100
ggtcatagct gtttcctgtt ttccccgtat ccccccaggt gtctgcaggc tcaaagagca 5160
gcgagaagcg ttcagaggaa agcgatcccg tgccaccttc cccgtgcccg ggctgtcccc 5220

gcacgctgcc ggctcgggga tgcgggggga gcgccggacc ggagcggagc cccgggcggc 5280
tcgctgctgc cccctagcgg gggagggacg taattacatc cctgggggct ttgggggggg 5340
gctgtccctc tcaccgcggt ggagctccag cttttgttcg aattggggcc ccccctcgag 5400
ggtatcgatg atatctataa caagaaaata tatatataat aagttatcac gtaagtagaa 5460
catgaaataa caatataatt atcgtatgag ttaaatctta aaagtcacgt aaaagataat 5520
catgcgtcat tttgactcac gcggtcgtta tagttcaaaa tcagtgacac ttaccgcatt 5580
gacaagcacg cctcacggga gctccaagcg gcgactgaga tgtcctaaat gcacagcgac 5640

ggattcgcgc tatttagaaa gagagagcaa tatttcaaga atgcatgcgt caattttacg 5700
cagactatct ttctagggtt aatctagcta gccttaaggg cgcagcccgc ctaatgagcg 5760
ggctttttt tggcttgttg tccacaaccg ttaaacctta aaagctttaa aagccttata 5820
tattcttttt tttcttataa aacttaaaac cttagaggct atttaagttg ctgatttata 5880
ttaattttat tgttcaaaca tgagagctta gtacgtgaaa catgagagct tagtacgtta 5940
gccatgagag cttagtacgt tagccatgag ggtttagttc gttaaacatg agagcttagt 6000
acgttaaaca tgagagctta gtacgtacta tcaacaggtt gaactgctga tccacgttgt 6060

ggtagaattg gtaaagagag tcgtgtaaaa tatcgagttc gcacatcttg ttgtctgatt 6120

attgattttt ggcgaaacca tttgatcata tgacaagatg tgtatctacc ttaacttaat 6180

gattttgata aaaatcatta ggta 6204

<210> 19

<211> 1008

<212> DNA

<213> Artificial Sequence

<220><223> PSMA CARTyrin

<400> 19

atggcactgc cagtcaccgc cctgctgctg cctctggctc tgctgctgca cgcagctaga 60

ccaatgctgc ctgcaccaaa gaacctggtg gtgtctcggg tgaccgagga ctctgccaga 120

ctgagctggg ccatcgacga gcagagggat tggttcgaga gctttctgat ccagtatcag 180

gagtccgaga aagtgggcga ggccatcgtg ctgacagtgc ctggcagcga gcggtcctat 240

gatctgaccg gcctgaagcc aggcacagag tacaccgtgt ccatctacgg cgtgtatcac 300

gtgtacaggt ccaatcctct gtctgccatc ttcaccacaa ccacaacccc tgcccccaga 360

cctcccacac ccgcccctac catcgcgagt cagcccctga gtctgagacc tgaggcctgc 420

aggccagctg caggaggagc tgtgcacacc aggggcctgg acttcgcctg cgacatctac 480

atttgggcac cactggccgg gacctgtgga gtgctgctgc tgagcctggt catcacactg 540

tactgcaaga gaggcaggaa gaaactgctg tatattttca aacagccctt catgcgcccc 600

gtgcagacta cccaggagga agacgggtgc tcctgtcgat tccctgagga agaggaaggc 660

gggtgtgagc tgcgcgtgaa gtttagtcga tcagcagatg ccccagctta caaacaggga 720

cagaaccagc tgtataacga gctgaatctg ggccgccgag aggaatatga cgtgctggat 780

aagcggagag gacgcgaccc cgaaatggga ggcaagccca ggcgcaaaaa ccctcaggaa 840

ggcctgtata acgagctgca gaaggacaaa atggcagaag cctattctga gatcggcatg 900

aagggggagc gacggagagg caaagggcac gatgggctgt accagggact gagcaccgcc 960

acaaaggaca cctatgatgc tctgcatatg caggcactgc ctccaagg 1008

<210> 20

<211> 960

<212> PRT

<213> Artificial Sequence

<220><223> P-PSMA-101 amino acid sequence

<400> 20

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

Phe Met Leu Gly Lys Gln Glu Val Ile Arg Gly Trp Glu Glu Gly Val
50 55 60
Ala Gln Met Ser Val Gly Gln Arg Ala Lys Leu Thr Ile Ser Pro Asp
65 70 75 80
Tyr Ala Tyr Gly Ala Thr Gly His Pro Gly Ile Ile Pro Pro His Ala
85 90 95
Thr Leu Val Phe Asp Val Glu Leu Leu Lys Leu Glu Gly Gly Gly Gly
100 105 110

Ser Gly Phe Gly Asp Val Gly Ala Leu Glu Ser Leu Arg Gly Asn Ala
115 120 125
Asp Leu Ala Tyr Ile Leu Ser Met Glu Pro Cys Gly His Cys Leu Ile
130 135 140
Ile Asn Asn Val Asn Phe Cys Arg Glu Ser Gly Leu Arg Thr Arg Thr
145 150 155 160
Gly Ser Asn Ile Asp Cys Glu Lys Leu Arg Arg Arg Phe Ser Ser Leu
165 170 175

His Phe Met Val Glu Val Lys Gly Asp Leu Thr Ala Lys Lys Met Val
180 185 190
Leu Ala Leu Leu Glu Leu Ala Gln Gln Asp His Gly Ala Leu Asp Cys
195 200 205
Cys Val Val Val Ile Leu Ser His Gly Cys Gln Ala Ser His Leu Gln
210 215 220
Phe Pro Gly Ala Val Tyr Gly Thr Asp Gly Cys Pro Val Ser Val Glu
225 230 235 240

Lys Ile Val Asn Ile Phe Asn Gly Thr Ser Cys Pro Ser Leu Gly Gly

245 250 255

Lys Pro Lys Leu Phe Phe Ile Gln Ala Cys Gly Gly Glu Gln Lys Asp

260 265 270

His Gly Phe Glu Val Ala Ser Thr Ser Pro Glu Asp Glu Ser Pro Gly

275 280 285

Ser Asn Pro Glu Pro Asp Ala Thr Pro Phe Gln Glu Gly Leu Arg Thr

290 295 300

Phe Asp Gln Leu Asp Ala Ile Ser Ser Leu Pro Thr Pro Ser Asp Ile

305 310 315 320

Phe Val Ser Tyr Ser Thr Phe Pro Gly Phe Val Ser Trp Arg Asp Pro

325 330 335

Lys Ser Gly Ser Trp Tyr Val Glu Thr Leu Asp Asp Ile Phe Glu Gln

340 345 350

Trp Ala His Ser Glu Asp Leu Gln Ser Leu Leu Leu Arg Val Ala Asn

355 360 365

Ala Val Ser Val Lys Gly Ile Tyr Lys Gln Met Pro Gly Cys Phe Asn

370 375 380

Phe Leu Arg Lys Lys Leu Phe Phe Lys Thr Ser Gly Ser Gly Glu Gly

385 390 395 400

Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro Gly Pro

405 410 415

Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu

420 425 430

His Ala Ala Arg Pro Met Leu Pro Ala Pro Lys Asn Leu Val Val Ser

435 440 445

Arg Val Thr Glu Asp Ser Ala Arg Leu Ser Trp Ala Ile Asp Glu Gln

450 455 460

Arg Asp Trp Phe Glu Ser Phe Leu Ile Gln Tyr Gln Glu Ser Glu Lys

465 470 475 480

Val Gly Glu Ala Ile Val Leu Thr Val Pro Gly Ser Glu Arg Ser Tyr

485 490 495

Asp Leu Thr Gly Leu Lys Pro Gly Thr Glu Tyr Thr Val Ser Ile Tyr
500 505 510

Gly Val Tyr His Val Tyr Arg Ser Asn Pro Leu Ser Ala Ile Phe Thr
515 520 525

Thr Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile
530 535 540

Ala Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala
545 550 555 560

Gly Gly Ala Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr
565 570 575

Ile Trp Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu Ser Leu
580 585 590

Val Ile Thr Leu Tyr Cys Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile
595 600 605

Phe Lys Gln Pro Phe Met Arg Pro Val Gln Thr Thr Gln Glu Glu Asp
610 615 620

Gly Cys Ser Cys Arg Phe Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu
625 630 635 640

Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly
645 650 655

Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr
660 665 670

Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys
675 680 685

Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys
690 695 700

Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg
705 710 715 720

Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala
725 730 735

Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg

740 745 750

Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu
755 760 765
Glu Asn Pro Gly Pro Met Val Gly Ser Leu Asn Cys Ile Val Ala Val
770 775 780
Ser Gln Asn Met Gly Ile Gly Lys Asn Gly Asp Phe Pro Trp Pro Pro
785 790 795 800
Leu Arg Asn Glu Ser Arg Tyr Phe Gln Arg Met Thr Thr Thr Ser Ser
805 810 815

Val Glu Gly Lys Gln Asn Leu Val Ile Met Gly Lys Lys Thr Trp Phe
820 825 830
Ser Ile Pro Glu Lys Asn Arg Pro Leu Lys Gly Arg Ile Asn Leu Val
835 840 845
Leu Ser Arg Glu Leu Lys Glu Pro Pro Gln Gly Ala His Phe Leu Ser
850 855 860
Arg Ser Leu Asp Asp Ala Leu Lys Leu Thr Glu Gln Pro Glu Leu Ala
865 870 875 880

Asn Lys Val Asp Met Val Trp Ile Val Gly Gly Ser Ser Val Tyr Lys
885 890 895
Glu Ala Met Asn His Pro Gly His Leu Lys Leu Phe Val Thr Arg Ile
900 905 910
Met Gln Asp Phe Glu Ser Asp Thr Phe Phe Pro Glu Ile Asp Leu Glu
915 920 925
Lys Tyr Lys Leu Leu Pro Glu Tyr Pro Gly Val Leu Ser Asp Val Gln
930 935 940

Glu Glu Lys Gly Ile Lys Tyr Lys Phe Glu Val Tyr Glu Lys Asn Asp
945 950 955 960

<210> 21
<211> 6248
<212> DNA
<213> Artificial Sequence

<220><223> P-BCMA-ALLO1 nanotransposon expressing a BCMA VCAR

<400> 21

tgtacataga ttaaccctag aaagataatc atattgtgac gtacgttaaa gataatcatg 60

cgtaaaattg acgcatgtgt tttatcggtc tgtatatcga ggtttattta ttaatttgaa 120

tagatattaa gttttattat atttacactt acatactaat aataaattca acaaacaatt 180

tatttatgtt tatttattta ttaaaaaaaa acaaaaactc aaaatttctt ctataaagta 240

acaaaacttt tatcgaatac ctgcagcccg ggggatgcag agggacagcc ccccccaaa 300

gcccccaggg atgtaattac gtccctcccc cgctaggggg cagcagcgag ccgcccgggg 360

ctccgctccg gtccggcgct ccccccgcat ccccgagccg gcagcgtgcg gggacagccc 420

gggcacgggg aaggtggcac gggatcgctt tcctctgaac gcttctcgct gctctttgag 480

cctgcagaca cctgggggga tacggggaaa agttgactgt gcctttcgat cgaaccatgg 540

acagttagct ttgcaaagat ggataaagtt ttaaacagag aggaatcttt gcagctaatg 600

gaccttctag gtcttgaaag gagtgggaat tggctccggt gcccgtcagt gggcagagcg 660

cacatcgccc acagtccccg agaagttggg gggaggggtc ggcaattgaa ccggtgccta 720

gagaaggtgg cgcggggtaa actgggaaag tgatgtcgtg tactggctcc gccttttcc 780

cgagggtggg ggagaaccgt atataagtgc agtagtcgcc gtgaacgttc tttttcgcaa 840

cgggtttgcc gccagaacac aggtaagtgc cgtgtgtggt tcccgcgggc ctggcctctt 900

tacgggttat ggcccttgcg tgccttgaat tacttccacc tggctgcagt acgtgattct 960

tgatcccgag cttcgggttg gaagtgggtg ggagagttcg aggccttgcg cttaaggagc 1020

cccttcgcct cgtgcttgag ttgaggcctg gcctgggcgc tggggccgcc gcgtgcgaat 1080

ctggtggcac cttcgcgcct gtctcgctgc tttcgataag tctctagcca tttaaaattt 1140

ttgatgacct gctgcgacgc tttttttctg gcaagatagt cttgtaaatg cgggccaaga 1200

tctgcacact ggtatttcgg tttttggggc cgcgggcggc gacggggccc gtgcgtccca 1260

gcgcacatgt tcggcgaggc ggggcctgcg agcgcggcca ccgagaatcg gacggggta 1320

gtctcaagct ggccggcctg ctctggtgcc tggcctcgcg ccgccgtgta tcgccccgcc 1380

ctgggcggca aggctggccc ggtcggcacc agttgcgtga gcggaaagat ggccgcttcc 1440

cggccctgct gcagggagct caaaatggag gacgcggcgc tcgggagagc gggcgggtga 1500

gtcacccaca caaaggaaaa gggcctttcc gtcctcagcc gtcgcttcat gtgactccac 1560

ggagtaccgg gcgccgtcca ggcacctcga ttagttctcg agcttttgga gtacgtcgtc 1620

tttaggttgg ggggaggggt tttatgcgat ggagtttccc cacactgagt gggtggagac 1680

tgaagttagg ccagcttggc acttgatgta attctccttg gaatttgccc tttttgagtt 1740

tggatcttgg ttcattctca agcctcagac agtggttcaa agtttttttc ttccatttca 1800

ggtgtcgtga gaattctaat acgactcact atagggtgaa gcttgccacc atgggggtcc 1860

aggtggaaac aatctctccg ggggatgggc ggacattccc taaaaggggc cagacctgcg 1920

tggtgcatta caccggcatg ctggaagatg gcaagaaggt ggacagcagc cgggacagaa 1980

acaagccctt caagttcatg ctgggcaagc aagaagtgat cagaggctgg gaagagggcg 2040

tcgcccagat gtctgttgga cagagagcca agctgacaat cagccccgat tacgcctatg 2100

gcgccacagg acaccctggc atcattcctc cacatgccac actggtgttc gacgtggaac 2160

tgctgaagct ggaaggcggc ggaggatctg gctttggaga tgtgggagcc ctggaaagcc 2220

tgagaggcaa tgccgatctg gcctacatcc tgagcatgga accttgcggc cactgcctga 2280

ttatcaacaa cgtgaacttc tgtagagaga gcggcctgcg gaccagaacc ggcagcaata 2340

tcgattgcga gaagctgcgg cggagattca gcagcctgca cttcatggtg gaagtgaagg 2400

gcgacctgac cgccaagaaa atggtgctgg ctctgctgga actggcccag caagatcatg 2460

gcgccctgga ttgctgtgtg gtcgtgatcc tgtctcacgg ctgtcaggcc agccaccttc 2520

aattccctgg cgccgtgtat ggcacagatg gctgtcctgt gtccgtggaa aagatcgtga 2580

acatcttcaa cggcaccagc tgtcctagcc tcggcggaaa gcccaagctg ttcttcatcc 2640

aagcctgtgg cggcgagcag aaggatcacg gatttgaggt ggccagcaca agccccgagg 2700

atgagtctcc tggaagcaac cctgagcctg acgccacacc tttccaagag ggcctgagaa 2760

ccttcgacca gctggacgct atcagctccc tgcctacacc tagcgacatc ttcgtgtcct 2820

acagcacatt ccccggcttt gtgtcttggc gggaccctaa gtctggctct tggtacgtgg 2880

aaaccctgga cgacatcttt gagcagtggg ctcacagcga ggacctccag tctctgctgc 2940

tgagagtggc caatgccgtg tccgtgaagg gcatctacaa gcagatgcct ggctgcttca 3000

acttcctgcg gaagaagctg tttttcaaga ccagcggcag cggcgaaggc agaggatccc 3060

ttttgacatg cggcgatgtg gaagagaacc ccggacctat ggctctgcct gtgacagctc 3120

tgcttctgcc tctggcactg cttcttcatg cggcgcgccc tgaagttcag ctgcttgaat 3180

ctggcggagg cctggttcaa cctggcggat ctctgagact gagctgtgcc gccagcggct 3240

tcaccttcag caattacgcc atgacctgga tcagacaggc ccctggcaaa ggcctggaat 3300

gggtgtccgg aattacaggc gacggcggca gcacctttta cgccgattct gtgaagggca 3360

gattcaccat cagccgggac aacagcaaga acaccctgta cctgcagatg aacagcctga 3420

gagccgagga caccgccgtg tactactgcg tgaaggactg gaacaccacc atgatcaccg 3480

agagaggcca gggcacactg gtcaccgtgt cctctacaac aacaccggcg cctagacctc 3540

caacaccagc tcctacaatc gcgagtcagc ccctgtctct cagacccgaa gcctgcaggc 3600

cagctgcagg aggagctgtg cacaccaggg gcctggactt cgcctgcgac atctacattt 3660
gggcaccact ggccgggacc tgtggagtgc tgctgctgag cctggtcatc acactgtact 3720
gcaagagagg caggaagaaa ctgctgtata ttttcaaaca gcccttcatg cgccccgtgc 3780
agactaccca ggaggaagac gggtgctcct gtcgattccc tgaggaagag gaaggcgggt 3840
gtgagctgcg cgtgaagttt agtcgatcag cagatgcccc agcttacaaa cagggacaga 3900
accagctgta taacgagctg aatctgggcc gccgagagga atatgacgtg ctggataagc 3960
ggagaggacg cgaccccgaa atgggaggca agcccaggcg caaaaaccct caggaaggcc 4020

tgtataacga gctgcagaag gacaaaatgg cagaagccta ttctgagatc ggcatgaagg 4080
gggagcgacg gagaggcaaa gggcacgatg ggctgtacca gggactgagc accgccacaa 4140
aggacaccta tgatgctctg catatgcagg cactgcctcc aaggggaagt ggagaaggac 4200
gaggatcact gctgacatgc ggcgacgtgg aggaaaaccc tggcccaatg gtcgggtctc 4260
tgaattgtat cgtcgccgtg agtcagaaca tgggcattgg gaagaatggc gatttcccat 4320
ggccacctct gcgcaacgag tcccgatact ttcagcggat gacaactacc tcctctgtgg 4380
aagggaaaca gaatctggtc atcatgggaa agaaaacttg gttcagcatt ccagagaaga 4440

accggcccct gaaaggcaga atcaatctgg tgctgtcccg agaactgaag gagccaccac 4500
agggagctca ctttctgagc cggtccctgg acgatgcact gaagctgaca gaacagcctg 4560
agctggccaa caaagtcgat atggtgtgga tcgtcggggg aagttcagtg tataaggagg 4620
ccatgaatca ccccggccat ctgaaactgt tcgtcacacg gatcatgcag gactttgaga 4680
gcgatacttt ctttcctgaa attgacctgg agaagtacaa actgctgccc gaatatcctg 4740
gcgtgctgtc cgatgtccag gaagagaaag gcatcaaata caagttcgag gtctatgaga 4800
agaatgacta ataaggtacc gatcacatat gcctttaatt aaacactagt tctatagtgt 4860

cacctaaatt ccctttagtg agggttaatg gccgtaggcc gccagaattg ggtccagaca 4920
tgataagata cattgatgag tttggacaaa ccacaactag aatgcagtga aaaaaatgct 4980
ttatttgtga aatttgtgat gctattgctt tatttgtaac cattataagc tgcaataaac 5040
aagttaacaa caacaattgc attcatttta tgtttcaggt tcaggggggag gtgtgggagg 5100
ttttttcgga ctctaggacc tgcgcatgcg cttggcgtaa tcatggtcat agctgtttcc 5160
tgttttcccc gtatcccccc aggtgtctgc aggctcaaag agcagcgaga agcgttcaga 5220
ggaaagcgat cccgtgccac cttccccgtg cccgggctgt ccccgcacgc tgccggctcg 5280

gggatgcggg gggagcgccg gaccggagcg gagccccggg cggctcgctg ctgccccta 5340

gcgggggagg gacgtaatta catccctggg ggctttgggg gggggctgtc cctctcaccg 5400

cggtggagct ccagcttttg ttcgaattgg ggcccccccct cgagggtatc gatgatatct 5460

ataacaagaa aatatatata taataagtta tcacgtaagt agaacatgaa ataacaatat 5520

aattatcgta tgagttaaat cttaaaagtc acgtaaaaga taatcatgcg tcattttgac 5580

tcacgcggtc gttatagttc aaaatcagtg acacttaccg cattgacaag cacgcctcac 5640

gggagctcca agcggcgact gagatgtcct aaatgcacag cgacggattc gcgctattta 5700

gaaagagaga gcaatatttc aagaatgcat gcgtcaattt tacgcagact atctttctag 5760

ggttaatcta gctagcctta agggcgcagc ccgcctaatg agcgggcttt tttttggctt 5820

gttgtccaca accgttaaac cttaaaagct ttaaaagcct tatatattct tttttttctt 5880

ataaaactta aaaccttaga ggctatttaa gttgctgatt tatattaatt ttattgttca 5940

aacatgagag cttagtacgt gaaacatgag agcttagtac gttagccatg agagcttagt 6000

acgttagcca tgagggttta gttcgttaaa catgagagct tagtacgtta aacatgagag 6060

cttagtacgt actatcaaca ggttgaactg ctgatccacg ttgtggtaga attggtaaag 6120

agagtcgtgt aaaatatcga gttcgcacat cttgttgtct gattattgat ttttggcgaa 6180

accatttgat catatgacaa gatgtgtatc taccttaact taatgatttt gataaaaatc 6240

attaggta 6248

<210> 22

<211> 1086

<212> DNA

<213> Artificial Sequence

<220><223> BCMA VCAR

<400> 22

atggctctgc ctgtgacagc tctgcttctg cctctggcac tgcttcttca tgcggcgcgc 60

cctgaagttc agctgcttga atctggcgga ggcctggttc aacctggcgg atctctgaga 120

ctgagctgtg ccgccagcgg cttcaccttc agcaattacg ccatgacctg gatcagacag 180

gcccctggca aaggcctgga atgggtgtcc ggaattacag gcgacggcgg cagcaccttt 240

tacgccgatt ctgtgaaggg cagattcacc atcagccggg acaacagcaa gaacaccctg 300

tacctgcaga tgaacagcct gagagccgag gacaccgccg tgtactactg cgtgaaggac 360

tggaacacca ccatgatcac cgagagaggc cagggcacac tggtcaccgt gtcctctaca 420

acaacaccgg cgcctagacc tccaacacca gctcctacaa tcgcgagtca gcccctgtct 480

ctcagacccg aagcctgcag gccagctgca ggaggagctg tgcacaccag gggcctggac 540

ttcgcctgcg acatctacat ttgggcacca ctggccggga cctgtggagt gctgctgctg 600

agcctggtca tcacactgta ctgcaagaga ggcaggaaga aactgctgta tattttcaaa 660
cagcccttca tgcgccccgt gcagactacc caggaggaag acgggtgctc ctgtcgattc 720
cctgaggaag aggaaggcgg gtgtgagctg cgcgtgaagt ttagtcgatc agcagatgcc 780
ccagcttaca aacagggaca gaaccagctg tataacgagc tgaatctggg ccgccgagag 840
gaatatgacg tgctggataa gcggagagga cgcgaccccg aaatgggagg caagcccagg 900
cgcaaaaacc ctcaggaagg cctgtataac gagctgcaga aggacaaaat ggcagaagcc 960
tattctgaga tcggcatgaa gggggagcga cggagaggca aagggcacga tgggctgtac 1020

cagggactga gcaccgccac aaaggacacc tatgatgctc tgcatatgca ggcactgcct 1080
ccaagg 1086

<210> 23

<211> 570

<212> PRT

<213> Artificial Sequence

<220><223> P-BCMA-ALLO1 amino acid sequence

<400> 23

Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu
1 5 10 15
His Ala Ala Arg Pro Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu
20 25 30

Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe
35 40 45
Thr Phe Ser Asn Tyr Ala Met Thr Trp Ile Arg Gln Ala Pro Gly Lys
50 55 60
Gly Leu Glu Trp Val Ser Gly Ile Thr Gly Asp Gly Gly Ser Thr Phe
65 70 75 80
Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser
85 90 95

Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr
100 105 110
Ala Val Tyr Tyr Cys Val Lys Asp Trp Asn Thr Thr Met Ile Thr Glu

115 120 125

Arg Gly Gln Gly Thr Leu Val Thr Val Ser Ser Thr Thr Thr Pro Ala

130 135 140

Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala Ser Gln Pro Leu Ser

145 150 155 160

Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly Gly Ala Val His Thr

165 170 175

Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr Ile Trp Ala Pro Leu Ala

180 185 190

Gly Thr Cys Gly Val Leu Leu Leu Ser Leu Val Ile Thr Leu Tyr Cys

195 200 205

Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe Lys Gln Pro Phe Met

210 215 220

Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly Cys Ser Cys Arg Phe

225 230 235 240

Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu Arg Val Lys Phe Ser Arg

245 250 255

Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly Gln Asn Gln Leu Tyr Asn

260 265 270

Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp Val Leu Asp Lys Arg

275 280 285

Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro Arg Arg Lys Asn Pro

290 295 300

Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp Lys Met Ala Glu Ala

305 310 315 320

Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg Arg Gly Lys Gly His

325 330 335

Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys Asp Thr Tyr Asp

340 345 350

Ala Leu His Met Gln Ala Leu Pro Pro Arg Gly Ser Gly Glu Gly Arg

355 360 365

Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro Gly Pro Met
370 375 380
Val Gly Ser Leu Asn Cys Ile Val Ala Val Ser Gln Asn Met Gly Ile
385 390 395 400
Gly Lys Asn Gly Asp Phe Pro Trp Pro Pro Leu Arg Asn Glu Ser Arg
405 410 415

Tyr Phe Gln Arg Met Thr Thr Thr Ser Ser Val Glu Gly Lys Gln Asn
420 425 430
Leu Val Ile Met Gly Lys Lys Thr Trp Phe Ser Ile Pro Glu Lys Asn
435 440 445
Arg Pro Leu Lys Gly Arg Ile Asn Leu Val Leu Ser Arg Glu Leu Lys
450 455 460
Glu Pro Pro Gln Gly Ala His Phe Leu Ser Arg Ser Leu Asp Asp Ala
465 470 475 480

Leu Lys Leu Thr Glu Gln Pro Glu Leu Ala Asn Lys Val Asp Met Val
485 490 495
Trp Ile Val Gly Gly Ser Ser Val Tyr Lys Glu Ala Met Asn His Pro
500 505 510
Gly His Leu Lys Leu Phe Val Thr Arg Ile Met Gln Asp Phe Glu Ser
515 520 525
Asp Thr Phe Phe Pro Glu Ile Asp Leu Glu Lys Tyr Lys Leu Leu Pro
530 535 540

Glu Tyr Pro Gly Val Leu Ser Asp Val Gln Glu Glu Lys Gly Ile Lys
545 550 555 560
Tyr Lys Phe Glu Val Tyr Glu Lys Asn Asp
565 570

<210> 24
<211> 35
<212> DNA
<213> Artificial Sequence
<220><223> ITR
<400> 24

ccctagaaag atagtctgcg taaaattgac gcatg 35

<210> 25

<211> 63

<212> DNA

<213> Artificial Sequence

<220><223> ITR

<400> 25

ccctagaaag ataatcatat tgtgacgtac gttaaagata atcatgcgta aaattgacgc 60

atg 63

<210> 26

<211> 63

<212> DNA

<213> Artificial Sequence

<220><223> ITR

<400> 26

ccctagaaag ataatcatat tgtgacgtac gttaaagata atcatgcgta aaattgacgc 60

atg 63

<210> 27

<211> 241

<212> DNA

<213> Artificial Sequence

<220><223> ITR

<400> 27

ttaaccctag aaagataatc atattgtgac gtacgttaaa gataatcatg tgtaaaattg 60

acgcatgtgt tttatcggtc tgtatatcga ggtttattta ttaatttgaa tagatattaa 120

gttttattat atttacactt acatactaat aataaattca acaaacaatt tatttatgtt 180

tatttattta ttaaaaaaaa caaaaactca aaatttcttc tataaagtaa caaaactttt 240

a 241

<210> 28

<211> 270

<212> PRT

<213> Artificial Sequence

<220><223> BCMA centyrin

<400> 28

Ala Thr Gly Cys Thr Gly Cys Cys Thr Gly Cys Ala Cys Cys Ala Ala
1 5 10 15
Ala Gly Ala Ala Cys Cys Thr Gly Gly Thr Gly Gly Thr Gly Ala Gly

20 25 30
Cys Cys Gly Gly Ala Thr Cys Ala Cys Ala Gly Ala Gly Gly Ala Cys
35 40 45
Thr Cys Cys Gly Cys Cys Ala Gly Ala Cys Thr Gly Thr Cys Thr Thr
50 55 60
Gly Gly Ala Cys Cys Gly Cys Cys Cys Cys Thr Gly Ala Cys Gly Cys
65 70 75 80
Cys Gly Cys Cys Thr Thr Cys Gly Ala Thr Thr Cys Cys Thr Thr Thr

85 90 95
Cys Cys Ala Ala Thr Cys Cys Gly Gly Thr Ala Cys Ala Thr Cys Gly
100 105 110
Ala Gly Ala Cys Ala Cys Thr Gly Ala Thr Cys Thr Gly Gly Gly Gly
115 120 125
Cys Gly Ala Gly Gly Cys Cys Ala Thr Cys Thr Gly Gly Cys Thr Gly
130 135 140
Gly Ala Cys Gly Thr Gly Cys Cys Cys Gly Gly Cys Thr Cys Thr Gly

145 150 155 160
Ala Gly Ala Gly Gly Ala Gly Cys Thr Ala Cys Gly Ala Thr Cys Thr
165 170 175
Gly Ala Cys Ala Gly Gly Cys Cys Thr Gly Ala Ala Gly Cys Cys Thr
180 185 190
Gly Gly Cys Ala Cys Cys Gly Ala Gly Thr Ala Thr Gly Cys Ala Gly
195 200 205
Thr Gly Gly Thr Cys Ala Thr Cys Ala Cys Ala Gly Gly Ala Gly Thr

210 215 220
Gly Ala Ala Gly Gly Gly Cys Gly Gly Cys Ala Gly Gly Thr Thr Cys
225 230 235 240

Ala Gly Cys Thr Cys Cys Cys Cys Thr Cys Thr Gly Gly Thr Gly Gly

245 250 255

Cys Cys Thr Cys Thr Thr Thr Thr Ala Cys Cys Ala Cys Ala

260 265 270

<210> 29

<211> 90

<212> PRT

<213> Artificial Sequence

<220><223> BCMA centyrin

<400> 29

Met Leu Pro Ala Pro Lys Asn Leu Val Val Ser Arg Ile Thr Glu Asp

1 5 10 15

Ser Ala Arg Leu Ser Trp Thr Ala Pro Asp Ala Ala Phe Asp Ser Phe

20 25 30

Pro Ile Arg Tyr Ile Glu Thr Leu Ile Trp Gly Glu Ala Ile Trp Leu

35 40 45

Asp Val Pro Gly Ser Glu Arg Ser Tyr Asp Leu Thr Gly Leu Lys Pro

50 55 60

Gly Thr Glu Tyr Ala Val Val Ile Thr Gly Val Lys Gly Gly Arg Phe

65 70 75 80

Ser Ser Pro Leu Val Ala Ser Phe Thr Thr

85 90

<210> 30

<211> 334

<212> PRT

<213> Artificial Sequence

<220><223> BCMA CARTyrin

<400> 30

Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu

1 5 10 15

His Ala Ala Arg Pro Met Leu Pro Ala Pro Lys Asn Leu Val Val Ser

20 25 30

Arg Ile Thr Glu Asp Ser Ala Arg Leu Ser Trp Thr Ala Pro Asp Ala

35 40 45

Ala Phe Asp Ser Phe Pro Ile Arg Tyr Ile Glu Thr Leu Ile Trp Gly

50 55 60

Glu Ala Ile Trp Leu Asp Val Pro Gly Ser Glu Arg Ser Tyr Asp Leu

65 70 75 80

Thr Gly Leu Lys Pro Gly Thr Glu Tyr Ala Val Val Ile Thr Gly Val

85 90 95

Lys Gly Gly Arg Phe Ser Ser Pro Leu Val Ala Ser Phe Thr Thr Thr

100 105 110

Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala Ser

115 120 125

Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly Gly

130 135 140

Ala Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr Ile Trp

145 150 155 160

Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu Ser Leu Val Ile

165 170 175

Thr Leu Tyr Cys Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe Lys

180 185 190

Gln Pro Phe Met Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly Cys

195 200 205

Ser Cys Arg Phe Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu Arg Val

210 215 220

Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly Gln Asn

225 230 235 240

Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp Val

245 250 255

Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro Arg

260 265 270

Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp Lys

275 280 285

Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg Arg

290 295 300

Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys

305 310 315 320

Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg

325 330

<210> 31

<211> 21

<212> PRT

<213> Artificial Sequence

<220><223> CD8 signal peptide

<400> 31

Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu

1 5 10 15

His Ala Ala Arg Pro

20

<210> 32

<211> 63

<212> DNA

<213> Artificial Sequence

<220><223> CD8 Signal Peptide

<400> 32

atggctctgc ctgtgacagc tctgcttctg cctctggcac tgcttcttca tgcggcgcgc 60

cct 63

<210> 33

<211> 15

<

212> PRT

<213> Artificial Sequence

<220><223> linker

<400> 33

Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser

1 5 10 15

<210> 34

<211> 45

<212> DNA

<213> Artificial Sequence

<220><223> linker

<400> 34

ggcggaggcg gtagcggtgg cggaggtagc ggaggtggtg gatct 45

<210> 35

<211> 45

<212> PRT

<213> Artificial Sequence

<220><223> CD8a Hinge

<400> 35

Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala

1 5 10 15

Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly

20 25 30

Gly Ala Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp

35 40 45

<210> 36

<211> 135

<212> DNA

<213> Artificial Sequence

<220><223> CD8a Hinge

<400> 36

accacaacac cggcgcctag acctccaaca ccagctccta caatcgcgag tcagcccctg 60

tctctcagac ccgaagcctg caggccagct gcaggaggag ctgtgcacac caggggcctg 120

gacttcgcct gcgac 135

<210> 37

<211> 24

<212> PRT

<213> Artificial Sequence

<220><223> CD8a TM

<400> 37

Ile Tyr Ile Trp Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu
1 5 10 15

Ser Leu Val Ile Thr Leu Tyr Cys
20

<210> 38

<211> 72

<212> DNA

<213> Artificial Sequence

<220><223> CD8a TM

<400> 38

atctacattt gggcaccact ggccgggacc tgtggagtgc tgctgctgag cctggtcatc 60

acactgtact gc 72

<210> 39

<211> 42

<212> PRT

<213> Artificial Sequence

<220><223> 41BB ICS

<400> 39

Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe Lys Gln Pro Phe Met
1 5 10 15

Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly Cys Ser Cys Arg Phe
20 25 30

Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu
35 40

<210> 40

<211> 126

<212> DNA

<213> Artificial Sequence

<220><223> 41BB ICS

<400> 40

aagaggaggca ggaagaaact gctgtatatt ttcaaacagc ccttcatgcg ccccgtgcag 60

actacccagg aggaagacgg gtgctcctgt cgattccctg aggaagagga aggcgggtgt 120

gagctg 126

<210> 41

<211> 112

<212> PRT

<213> Artificial Sequence

<220><223> CD3z ICS

<400> 41

Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly
1 5 10 15

Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr
20 25 30

Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys
35 40 45

Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys
50 55 60

Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg
65 70 75 80

Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala
85 90 95

Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
100 105 110

<210> 42

<211> 336

<212> DNA

<213> Artificial Sequence

<220><223> 41BB ICS

<400> 42

cgcgtgaagt ttagtcgatc agcagatgcc ccagcttaca aacagggaca gaaccagctg 60

tataacgagc tgaatctggg ccgccgagag gaatatgacg tgctggataa gcggagagga 120

cgcgaccccg aaatgggagg caagcccagg cgcaaaaacc ctcaggaagg cctgtataac 180

gagctgcaga aggacaaaat ggcagaagcc tattctgaga tcggcatgaa gggggagcga 240

cggagaggca aagggcacga tgggctgtac cagggactga gcaccgccac aaaggacacc 300

tatgatgctc tgcatatgca ggcactgcct ccaagg 336

<210> 43

<211> 1053

<212> PRT

<213> Artificial Sequence

<220><223> SCas9 - 40-18163

<400> 43

Met Lys Arg Asn Tyr Ile Leu Gly Leu Asp Ile Gly Ile Thr Ser Val
1 5 10 15

Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly
20 25 30

Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg
35 40 45

Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile
50 55 60

Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His
65 70 75 80

Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu
85 90 95

Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu
100 105 110

Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr
115 120 125

Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala
130 135 140

Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys
145 150 155 160

Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr
165 170 175

Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln
180 185 190

Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg

195 200 205

Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys
210 215 220
Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe
225 230 235 240
Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr
245 250 255
Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn
260 265 270

Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe
275 280 285
Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu
290 295 300
Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys
305 310 315 320
Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr
325 330 335

Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala
340 345 350
Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu
355 360 365
Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser
370 375 380
Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile
385 390 395 400

Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala
405 410 415
Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln
420 425 430
Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro
435 440 445

Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile
450 455 460

Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg
465 470 475 480
Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys
485 490 495
Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr
500 505 510
Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp
515 520 525

Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu
530 535 540
Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro
545 550 555 560
Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys
565 570 575
Gln Glu Glu Asn Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu
580 585 590

Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile
595 600 605
Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu
610 615 620
Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
625 630 635 640
Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu
645 650 655

Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys
660 665 670
Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp
675 680 685
Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp

690 695 700

Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
705 710 715 720

Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys
725 730 735

Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu
740 745 750

Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp
755 760 765

Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile
770 775 780

Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
785 790 795 800

Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu
805 810 815

Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
820 825 830

Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly
835 840 845

Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr
850 855 860

Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile
865 870 875 880

Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp
885 890 895

Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr
900 905 910

Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val
915 920 925

Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser
930 935 940

Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala
945 950 955 960
Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly
965 970 975

Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile
980 985 990
Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met
995 1000 1005
Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys
1010 1015 1020
Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu
1025 1030 1035

Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly
1040 1045 1050

<210> 44
<211> 1053
<212> PRT
<213> Artificial Sequence
<220><223> dSaCas9 - 40-18164
<400> 44

Met Lys Arg Asn Tyr Ile Leu Gly Leu Ala Ile Gly Ile Thr Ser Val
1 5 10 15
Gly Tyr Gly Ile Ile Asp Tyr Glu Thr Arg Asp Val Ile Asp Ala Gly
20 25 30
Val Arg Leu Phe Lys Glu Ala Asn Val Glu Asn Asn Glu Gly Arg Arg

35 40 45
Ser Lys Arg Gly Ala Arg Arg Leu Lys Arg Arg Arg Arg His Arg Ile
50 55 60
Gln Arg Val Lys Lys Leu Leu Phe Asp Tyr Asn Leu Leu Thr Asp His
65 70 75 80
Ser Glu Leu Ser Gly Ile Asn Pro Tyr Glu Ala Arg Val Lys Gly Leu
85 90 95

Ser Gln Lys Leu Ser Glu Glu Glu Phe Ser Ala Ala Leu Leu His Leu

100 105 110

Ala Lys Arg Arg Gly Val His Asn Val Asn Glu Val Glu Glu Asp Thr

115 120 125

Gly Asn Glu Leu Ser Thr Lys Glu Gln Ile Ser Arg Asn Ser Lys Ala

130 135 140

Leu Glu Glu Lys Tyr Val Ala Glu Leu Gln Leu Glu Arg Leu Lys Lys

145 150 155 160

Asp Gly Glu Val Arg Gly Ser Ile Asn Arg Phe Lys Thr Ser Asp Tyr

165 170 175

Val Lys Glu Ala Lys Gln Leu Leu Lys Val Gln Lys Ala Tyr His Gln

180 185 190

Leu Asp Gln Ser Phe Ile Asp Thr Tyr Ile Asp Leu Leu Glu Thr Arg

195 200 205

Arg Thr Tyr Tyr Glu Gly Pro Gly Glu Gly Ser Pro Phe Gly Trp Lys

210 215 220

Asp Ile Lys Glu Trp Tyr Glu Met Leu Met Gly His Cys Thr Tyr Phe

225 230 235 240

Pro Glu Glu Leu Arg Ser Val Lys Tyr Ala Tyr Asn Ala Asp Leu Tyr

245 250 255

Asn Ala Leu Asn Asp Leu Asn Asn Leu Val Ile Thr Arg Asp Glu Asn

260 265 270

Glu Lys Leu Glu Tyr Tyr Glu Lys Phe Gln Ile Ile Glu Asn Val Phe

275 280 285

Lys Gln Lys Lys Lys Pro Thr Leu Lys Gln Ile Ala Lys Glu Ile Leu

290 295 300

Val Asn Glu Glu Asp Ile Lys Gly Tyr Arg Val Thr Ser Thr Gly Lys

305 310 315 320

Pro Glu Phe Thr Asn Leu Lys Val Tyr His Asp Ile Lys Asp Ile Thr

325 330 335

Ala Arg Lys Glu Ile Ile Glu Asn Ala Glu Leu Leu Asp Gln Ile Ala

340 345 350

Lys Ile Leu Thr Ile Tyr Gln Ser Ser Glu Asp Ile Gln Glu Glu Leu

355 360 365

Thr Asn Leu Asn Ser Glu Leu Thr Gln Glu Glu Ile Glu Gln Ile Ser

370 375 380

Asn Leu Lys Gly Tyr Thr Gly Thr His Asn Leu Ser Leu Lys Ala Ile

385 390 395 400

Asn Leu Ile Leu Asp Glu Leu Trp His Thr Asn Asp Asn Gln Ile Ala

405 410 415

Ile Phe Asn Arg Leu Lys Leu Val Pro Lys Lys Val Asp Leu Ser Gln

420 425 430

Gln Lys Glu Ile Pro Thr Thr Leu Val Asp Asp Phe Ile Leu Ser Pro

435 440 445

Val Val Lys Arg Ser Phe Ile Gln Ser Ile Lys Val Ile Asn Ala Ile

450 455 460

Ile Lys Lys Tyr Gly Leu Pro Asn Asp Ile Ile Ile Glu Leu Ala Arg

465 470 475 480

Glu Lys Asn Ser Lys Asp Ala Gln Lys Met Ile Asn Glu Met Gln Lys

485 490 495

Arg Asn Arg Gln Thr Asn Glu Arg Ile Glu Glu Ile Ile Arg Thr Thr

500 505 510

Gly Lys Glu Asn Ala Lys Tyr Leu Ile Glu Lys Ile Lys Leu His Asp

515 520 525

Met Gln Glu Gly Lys Cys Leu Tyr Ser Leu Glu Ala Ile Pro Leu Glu

530 535 540

Asp Leu Leu Asn Asn Pro Phe Asn Tyr Glu Val Asp His Ile Ile Pro

545 550 555 560

Arg Ser Val Ser Phe Asp Asn Ser Phe Asn Asn Lys Val Leu Val Lys

565 570 575

Gln Glu Glu Ala Ser Lys Lys Gly Asn Arg Thr Pro Phe Gln Tyr Leu

580 585 590

Ser Ser Ser Asp Ser Lys Ile Ser Tyr Glu Thr Phe Lys Lys His Ile
595 600 605
Leu Asn Leu Ala Lys Gly Lys Gly Arg Ile Ser Lys Thr Lys Lys Glu

610 615 620
Tyr Leu Leu Glu Glu Arg Asp Ile Asn Arg Phe Ser Val Gln Lys Asp
625 630 635 640
Phe Ile Asn Arg Asn Leu Val Asp Thr Arg Tyr Ala Thr Arg Gly Leu
645 650 655
Met Asn Leu Leu Arg Ser Tyr Phe Arg Val Asn Asn Leu Asp Val Lys
660 665 670
Val Lys Ser Ile Asn Gly Gly Phe Thr Ser Phe Leu Arg Arg Lys Trp

675 680 685
Lys Phe Lys Lys Glu Arg Asn Lys Gly Tyr Lys His His Ala Glu Asp
690 695 700
Ala Leu Ile Ile Ala Asn Ala Asp Phe Ile Phe Lys Glu Trp Lys Lys
705 710 715 720
Leu Asp Lys Ala Lys Lys Val Met Glu Asn Gln Met Phe Glu Glu Lys
725 730 735
Gln Ala Glu Ser Met Pro Glu Ile Glu Thr Glu Gln Glu Tyr Lys Glu

740 745 750
Ile Phe Ile Thr Pro His Gln Ile Lys His Ile Lys Asp Phe Lys Asp
755 760 765
Tyr Lys Tyr Ser His Arg Val Asp Lys Lys Pro Asn Arg Glu Leu Ile
770 775 780
Asn Asp Thr Leu Tyr Ser Thr Arg Lys Asp Asp Lys Gly Asn Thr Leu
785 790 795 800
Ile Val Asn Asn Leu Asn Gly Leu Tyr Asp Lys Asp Asn Asp Lys Leu

805 810 815
Lys Lys Leu Ile Asn Lys Ser Pro Glu Lys Leu Leu Met Tyr His His
820 825 830
Asp Pro Gln Thr Tyr Gln Lys Leu Lys Leu Ile Met Glu Gln Tyr Gly

835 840 845

Asp Glu Lys Asn Pro Leu Tyr Lys Tyr Tyr Glu Glu Thr Gly Asn Tyr

850 855 860

Leu Thr Lys Tyr Ser Lys Lys Asp Asn Gly Pro Val Ile Lys Lys Ile

865 870 875 880

Lys Tyr Tyr Gly Asn Lys Leu Asn Ala His Leu Asp Ile Thr Asp Asp

885 890 895

Tyr Pro Asn Ser Arg Asn Lys Val Val Lys Leu Ser Leu Lys Pro Tyr

900 905 910

Arg Phe Asp Val Tyr Leu Asp Asn Gly Val Tyr Lys Phe Val Thr Val

915 920 925

Lys Asn Leu Asp Val Ile Lys Lys Glu Asn Tyr Tyr Glu Val Asn Ser

930 935 940

Lys Cys Tyr Glu Glu Ala Lys Lys Leu Lys Lys Ile Ser Asn Gln Ala

945 950 955 960

Glu Phe Ile Ala Ser Phe Tyr Asn Asn Asp Leu Ile Lys Ile Asn Gly

965 970 975

Glu Leu Tyr Arg Val Ile Gly Val Asn Asn Asp Leu Leu Asn Arg Ile

980 985 990

Glu Val Asn Met Ile Asp Ile Thr Tyr Arg Glu Tyr Leu Glu Asn Met

995 1000 1005

Asn Asp Lys Arg Pro Pro Arg Ile Ile Lys Thr Ile Ala Ser Lys

1010 1015 1020

Thr Gln Ser Ile Lys Lys Tyr Ser Thr Asp Ile Leu Gly Asn Leu

1025 1030 1035

Tyr Glu Val Lys Ser Lys Lys His Pro Gln Ile Ile Lys Lys Gly

1040 1045 1050

<210> 45

<211> 1368

<212> PRT

<213> Artificial Sequence

<220><223> dCas9 - 40-18165

<220><221> misc_feature

<222> (1)..(1)

<223> Xaa can be any naturally occurring amino acid

<400> 45

Xaa Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val
1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe
20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile
35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu
50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys
65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser
85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys
100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr
115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp
130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His
145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro
165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr
180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala
195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn

210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn
225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe
245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp
260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
275 280 285

Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
290 295 300

Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser
305 310 315 320

Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
325 330 335

Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
340 345 350

Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
355 360 365

Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp
370 375 380

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

Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415

Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430

Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile
435 440 445

Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460

Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480
Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495

Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser
500 505 510
Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys
515 520 525
Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln
530 535 540
Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr
545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp
565 570 575
Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly
580 585 590
Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp
595 600 605
Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr
610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala
625 630 635 640
His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr
645 650 655
Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp
660 665 670
Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe
675 680 685

Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe
690 695 700
Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu

705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly

725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly

740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln

755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile

770 775 780

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

785 790 795 800

Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu

805 810 815

Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg

820 825 830

Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys

835 840 845

Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg

850 855 860

Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys

865 870 875 880

Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys

885 890 895

Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp

900 905 910

Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr

915 920 925

Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp

930 935 940

Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser

945 950 955 960

Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
965 970 975
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
980 985 990
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
995 1000 1005

Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala
1010 1015 1020
Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe
1025 1030 1035
Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala
1040 1045 1050
Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu
1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val
1070 1075 1080
Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr
1085 1090 1095
Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys
1100 1105 1110
Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro
1115 1120 1125

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val
1130 1135 1140
Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys
1145 1150 1155
Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser
1160 1165 1170
Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys
1175 1180 1185

Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu

1190 1195 1200

Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly

1205 1210 1215

Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val

1220 1225 1230

Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser

1235 1240 1245

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys

1250 1255 1260

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys

1265 1270 1275

Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala

1280 1285 1290

Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn

1295 1300 1305

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala

1310 1315 1320

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser

1325 1330 1335

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr

1340 1345 1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp

1355 1360 1365

<210> 46

<211> 1368

<212> PRT

<213> Artificial Sequence

<220><223> dCas9 - 40-18166

<400> 46

Met Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val

1 5 10 15

Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe

20 25 30

Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile

35 40 45

Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu

50 55 60

Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys

65 70 75 80

Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser

85 90 95

Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys

100 105 110

His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr

115 120 125

His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp

130 135 140

Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His

145 150 155 160

Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro

165 170 175

Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr

180 185 190

Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala

195 200 205

Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn

210 215 220

Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn

225 230 235 240

Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe

245 250 255

Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp

260 265 270

Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp
275 280 285
Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp
290 295 300
Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser

305 310 315 320
Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys
325 330 335
Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe
340 345 350
Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser
355 360 365
Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp

370 375 380
Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg
385 390 395 400
Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu
405 410 415
Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe
420 425 430
Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile

435 440 445
Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp
450 455 460
Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu
465 470 475 480
Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr
485 490 495
Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser

500 505 510
Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys

515 520 525

Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln

530 535 540

Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr

545 550 555 560

Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp

565 570 575

Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly

580 585 590

Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp

595 600 605

Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr

610 615 620

Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala

625 630 635 640

His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr

645 650 655

Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp

660 665 670

Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe

675 680 685

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

690 695 700

Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu

705 710 715 720

His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly

725 730 735

Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly

740 745 750

Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln

755 760 765

Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile
770 775 780
Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro
785 790 795 800
Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu
805 810 815
Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg

820 825 830
Leu Ser Asp Tyr Asp Val Asp Ala Ile Val Pro Gln Ser Phe Leu Lys
835 840 845
Asp Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg
850 855 860
Gly Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys
865 870 875 880
Asn Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys

885 890 895
Phe Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp
900 905 910
Lys Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr
915 920 925
Lys His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp
930 935 940
Glu Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser

945 950 955 960
Lys Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
965 970 975
Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
980 985 990
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe
995 1000 1005
Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala

1010 1015 1020

Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe

1025 1030 1035

Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala

1040 1045 1050

Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu

1055 1060 1065

Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val

1070 1075 1080

Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr

1085 1090 1095

Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys

1100 1105 1110

Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro

1115 1120 1125

Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val

1130 1135 1140

Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys

1145 1150 1155

Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser

1160 1165 1170

Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys

1175 1180 1185

Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu

1190 1195 1200

Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly

1205 1210 1215

Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val

1220 1225 1230

Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser

1235 1240 1245

Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys

1250 1255 1260

His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys

1265 1270 1275

Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala

1280 1285 1290

Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn

1295 1300 1305

Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala

1310 1315 1320

Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser

1325 1330 1335

Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr

1340 1345 1350

Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp

1355 1360 1365

<210> 47

<211> 199

<212> PRT

<213> Artificial Sequence

<220><223> Clo051 - 40-18030

<400> 47

Glu Gly Ile Lys Ser Asn Ile Ser Leu Leu Lys Asp Glu Leu Arg Gly

1 5 10 15

Gln Ile Ser His Ile Ser His Glu Tyr Leu Ser Leu Ile Asp Leu Ala

20 25 30

Phe Asp Ser Lys Gln Asn Arg Leu Phe Glu Met Lys Val Leu Glu Leu

35 40 45

Leu Val Asn Glu Tyr Gly Phe Lys Gly Arg His Leu Gly Gly Ser Arg

50 55 60

Lys Pro Asp Gly Ile Val Tyr Ser Thr Thr Leu Glu Asp Asn Phe Gly

65 70 75 80

Ile Ile Val Asp Thr Lys Ala Tyr Ser Glu Gly Tyr Ser Leu Pro Ile

85 90 95

Ser Gln Ala Asp Glu Met Glu Arg Tyr Val Arg Glu Asn Ser Asn Arg

100 105 110

Asp Glu Glu Val Asn Pro Asn Lys Trp Trp Glu Asn Phe Ser Glu Glu

115 120 125

Val Lys Lys Tyr Tyr Phe Val Phe Ile Ser Gly Ser Phe Lys Gly Lys

130 135 140

Phe Glu Glu Gln Leu Arg Arg Leu Ser Met Thr Thr Gly Val Asn Gly

145 150 155 160

Ser Ala Val Asn Val Val Asn Leu Leu Leu Gly Ala Glu Lys Ile Arg

165 170 175

Ser Gly Glu Met Thr Ile Glu Glu Leu Glu Arg Ala Met Phe Asn Asn

180 185 190

Ser Glu Phe Ile Leu Lys Tyr

195

<210> 48

<211> 1591

<212> PRT

<213> Artificial Sequence

<220><223> dCas9-Clo051 - 40-18168

<400> 48

Met Ala Pro Lys Lys Lys Arg Lys Val Glu Gly Ile Lys Ser Asn Ile

1 5 10 15

Ser Leu Leu Lys Asp Glu Leu Arg Gly Gln Ile Ser His Ile Ser His

20 25 30

Glu Tyr Leu Ser Leu Ile Asp Leu Ala Phe Asp Ser Lys Gln Asn Arg

35 40 45

Leu Phe Glu Met Lys Val Leu Glu Leu Leu Val Asn Glu Tyr Gly Phe

50 55 60

Lys Gly Arg His Leu Gly Gly Ser Arg Lys Pro Asp Gly Ile Val Tyr

65 70 75 80

Ser Thr Thr Leu Glu Asp Asn Phe Gly Ile Ile Val Asp Thr Lys Ala

85 90 95

Tyr Ser Glu Gly Tyr Ser Leu Pro Ile Ser Gln Ala Asp Glu Met Glu

100 105 110

Arg Tyr Val Arg Glu Asn Ser Asn Arg Asp Glu Glu Val Asn Pro Asn

115 120 125

Lys Trp Trp Glu Asn Phe Ser Glu Glu Val Lys Lys Tyr Tyr Phe Val

130 135 140

Phe Ile Ser Gly Ser Phe Lys Gly Lys Phe Glu Glu Gln Leu Arg Arg

145 150 155 160

Leu Ser Met Thr Thr Gly Val Asn Gly Ser Ala Val Asn Val Val Asn

165 170 175

Leu Leu Leu Gly Ala Glu Lys Ile Arg Ser Gly Glu Met Thr Ile Glu

180 185 190

Glu Leu Glu Arg Ala Met Phe Asn Asn Ser Glu Phe Ile Leu Lys Tyr

195 200 205

Gly Gly Gly Gly Ser Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly

210 215 220

Thr Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro

225 230 235 240

Ser Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys

245 250 255

Lys Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu

260 265 270

Ala Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys

275 280 285

Asn Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys

290 295 300

Val Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu

305 310 315 320

Glu Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp
325 330 335
Glu Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys
340 345 350
Lys Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu
355 360 365

Ala Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly
370 375 380
Asp Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu
385 390 395 400
Val Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser
405 410 415
Gly Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg
420 425 430

Arg Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly
435 440 445
Leu Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe
450 455 460
Lys Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys
465 470 475 480
Asp Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp
485 490 495

Gln Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile
500 505 510
Leu Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro
515 520 525
Leu Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu
530 535 540
Thr Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys
545 550 555 560

Glu Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp

565 570 575

Gly Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu

580 585 590

Glu Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu

595 600 605

Asp Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His

610 615 620

Gln Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp

625 630 635 640

Phe Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu

645 650 655

Thr Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser

660 665 670

Arg Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp

675 680 685

Asn Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile

690 695 700

Glu Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu

705 710 715 720

Pro Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu

725 730 735

Thr Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu

740 745 750

Ser Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn

755 760 765

Arg Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile

770 775 780

Glu Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn

785 790 795 800

Ala Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys

805 810 815

Asp Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val
820 825 830
Leu Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu
835 840 845
Lys Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys
850 855 860
Arg Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn
865 870 875 880

Gly Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys
885 890 895
Ser Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp
900 905 910
Ser Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln
915 920 925
Gly Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala
930 935 940

Ile Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val
945 950 955 960
Lys Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala
965 970 975
Arg Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg
980 985 990
Met Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu
995 1000 1005

Lys Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu
1010 1015 1020
Tyr Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp Gln
1025 1030 1035
Glu Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp Ala Ile
1040 1045 1050
Val Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val

1055 1060 1065

Leu Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro
1070 1075 1080
Ser Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu
1085 1090 1095
Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr
1100 1105 1110
Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe
1115 1120 1125

Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val
1130 1135 1140
Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn
1145 1150 1155
Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys
1160 1165 1170
Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg
1175 1180 1185

Glu Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala
1190 1195 1200
Val Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser
1205 1210 1215
Glu Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met
1220 1225 1230
Ile Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr
1235 1240 1245

Phe Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr
1250 1255 1260
Leu Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn
1265 1270 1275
Gly Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala
1280 1285 1290

Thr Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys
1295 1300 1305

Lys Thr Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu
1310 1315 1320
Pro Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp
1325 1330 1335
Asp Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr
1340 1345 1350
Ser Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys
1355 1360 1365

Leu Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg
1370 1375 1380
Ser Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly
1385 1390 1395
Tyr Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr
1400 1405 1410
Ser Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser
1415 1420 1425

Ala Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys
1430 1435 1440
Tyr Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys
1445 1450 1455
Gly Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln
1460 1465 1470
His Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe
1475 1480 1485

Ser Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu
1490 1495 1500
Ser Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala
1505 1510 1515
Glu Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro

1520 1525 1530

Ala Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr

1535 1540 1545

Thr Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser

1550 1555 1560

Ile Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly

1565 1570 1575

Gly Asp Gly Ser Pro Lys Lys Lys Arg Lys Val Ser Ser

1580 1585 1590

<210> 49

<211> 4776

<212> DNA

<213> Artificial Sequence

<220><223> dCas9-Clo051 - 40-18169

<400> 49

atggcaccaa agaagaaaag aaaagtggag ggcatcaagt caaacatcag cctgctgaaa 60

gacgaactgc ggggacagat tagtcacatc agtcacgagt acctgtcact gattgatctg 120

gccttcgaca gcaagcagaa tagactgttt gagatgaaag tgctggaact gctggtcaac 180

gagtatggct tcaagggcag acatctgggc gggtctagga aacctgacgg catcgtgtac 240

agtaccacac tggaagacaa cttcggaatc attgtcgata ccaaggctta ttccgagggc 300

tactctctgc caattagtca ggcagatgag atggaaaggt acgtgcgcga aaactcaaat 360

agggacgagg aagtcaaccc caataagtgg tgggagaatt tcagcgagga agtgaagaaa 420

tactacttcg tctttatctc aggcagcttc aaagggaagt ttgaggaaca gctgcggaga 480

ctgtccatga ctaccggggt gaacggatct gctgtcaacg tggtcaatct gctgctgggc 540

gcagaaaaga tcaggtccgg ggagatgaca attgaggaac tggaacgcgc catgttcaac 600

aattctgagt ttatcctgaa gtatggaggc gggggaagcg ataagaaata ctccatcgga 660

ctggccattg gcaccaattc cgtgggctgg gctgtcatca cagacgagta caaggtgcca 720

agcaagaagt tcaaggtcct ggggaacacc gatcgccaca gtatcaagaa aaatctgatt 780

ggagccctgc tgttcgactc aggcgagact gctgaagcaa cccgactgaa gcggactgct 840

aggcgccgat atacccggag aaaaaatcgg atctgctacc tgcaggaaat tttcagcaac 900

gagatggcca aggtggacga tagtttcttt caccgcctgg aggaatcatt cctggtggag 960

gaagataaga aacacgagcg gcatcccatc tttggcaaca ttgtggacga agtcgcttat 1020

cacgagaagt accctactat ctatcatctg aggaagaaac tggtggactc caccgataag 1080

gcagacctgc gcctgatcta tctggccctg gctcacatga tcaagttccg ggggcatttt 1140

ctgatcgagg gagatctgaa ccctgacaat tctgatgtgg acaagctgtt catccagctg 1200

gtccagacat acaatcagct gtttgaggaa aacccaatta atgcctcagg cgtggacgca 1260

aaggccatcc tgagcgccag actgtccaaa tctaggcgcc tggaaaacct gatcgctcag 1320

ctgccaggag agaagaaaaa cggcctgttt gggaatctga ttgcactgtc cctgggcctg 1380

acacccaact tcaagtctaa ttttgatctg gccgaggacg ctaagctgca gctgtccaaa 1440

gacacttatg acgatgacct ggataacctg ctggctcaga tcggcgatca gtacgcagac 1500

ctgttcctgg ccgctaagaa tctgagtgac gccatcctgc tgtcagatat tctgcgcgtg 1560

aacacagaga ttactaaggc cccactgagt gcttcaatga tcaaaagata tgacgagcac 1620

catcaggatc tgaccctgct gaaggctctg gtgaggcagc agctgcccga gaaatacaag 1680

gaaatcttct ttgatcagag caagaatgga tacgccggct atattgacgg cggggcttcc 1740

caggaggagt tctacaagtt catcaagccc attctggaaa agatggacgg caccgaggaa 1800

ctgctggtga agctgaatcg ggaggacctg ctgagaaaac agaggacatt tgataacgga 1860

agcatccctc accagattca tctgggcgaa ctgcacgcca tcctgcgacg gcaggaggac 1920

ttctacccat ttctgaagga taaccgcgag aaaatcgaaa agatcctgac cttcagaatc 1980

ccctactatg tggggcctct ggcacgggga aatagtagat ttgcctggat gacaagaaag 2040

tcagaggaaa ctatcacccc ctggaacttc gaggaagtgg tcgataaagg cgctagcgca 2100

cagtccttca ttgaaaggat gacaaatttt gacaagaacc tgccaaatga gaaggtgctg 2160

cccaaacaca gcctgctgta cgaatatttc acagtgtata acgagctgac taaagtgaag 2220

tacgtcaccg aagggatgcg caagcccgca ttcctgtccg gagagcagaa gaaagccatc 2280

gtggacctgc tgtttaagac aaatcggaaa gtgactgtca aacagctgaa ggaagactat 2340

ttcaagaaaa ttgagtgttt cgattcagtg gaaatcagcg gcgtcgagga caggtttaac 2400

gcctccctgg ggacctacca cgatctgctg aagatcatca aggataagga cttcctggac 2460

aacgaggaaa atgaggacat cctggaggac attgtgctga cactgactct gtttgaggat 2520

cgcgaaatga tcgaggaacg actgaagact tatgcccatc tgttcgatga caaagtgatg 2580

aagcagctga aaagaaggcg ctacaccgga tggggacgcc tgagccgaaa actgatcaat 2640

gggattagag acaagcagag cggaaaaact atcctggact ttctgaagtc cgatggcttc 2700

gccaacagga acttcatgca gctgattcac gatgactctc tgaccttcaa ggaggacatc 2760

cagaaagcac aggtgtctgg ccagggggac agtctgcacg agcatatcgc aaacctggcc 2820

ggcagccccg ccatcaagaa agggattctg cagaccgtga aggtggtgga cgaactggtc 2880

aaggtcatgg gacgacacaa acctgagaac atcgtgattg agatggcccg cgaaaatcag 2940

acaactcaga agggccagaa aaacagtcga gaacggatga agagaatcga ggaaggcatc 3000

aaggagctgg ggtcacagat cctgaaggag catcctgtgg aaaacactca gctgcagaat 3060

gagaaactgt atctgtacta tctgcagaat ggacgggata tgtacgtgga ccaggagctg 3120

gatattaaca gactgagtga ttatgacgtg gatgccatcg tccctcagag cttcctgaag 3180

gatgactcca ttgacaacaa ggtgctgacc aggtccgaca agaaccgcgg caaatcagat 3240

aatgtgccaa gcgaggaagt ggtcaagaaa atgaagaact actggaggca gctgctgaat 3300

gccaagctga tcacacagcg gaaatttgat aacctgacta aggcagaaag aggaggcctg 3360

tctgagctgg acaaggccgg cttcatcaag cggcagctgg tggagacaag acagatcact 3420

aagcacgtcg ctcagattct ggatagcaga atgaacacaa agtacgatga aaacgacaag 3480

ctgatcaggg aggtgaaagt cattactctg aaatccaagc tggtgtctga ctttagaaag 3540

gatttccagt tttataaagt cagggagatc aacaactacc accatgctca tgacgcatac 3600

ctgaacgcag tggtcgggac cgccctgatt aagaaatacc ccaagctgga gtccgagttc 3660

gtgtacggag actataaagt gtacgatgtc cggaagatga tcgccaaatc tgagcaggaa 3720

attggcaagg ccaccgctaa gtatttcttt tacagtaaca tcatgaattt ctttaagacc 3780

gaaatcacac tggcaaatgg ggagatcaga aaaaggcctc tgattgagac caacggggag 3840

acaggagaaa tcgtgtggga caagggaagg gattttgcta ccgtgcgcaa agtcctgtcc 3900

atgccccaag tgaatattgt caagaaaact gaagtgcaga ccgggggatt ctctaaggag 3960

agtattctgc ctaagcgaaa ctctgataaa ctgatcgccc ggaagaaaga ctgggacccc 4020

aagaagtatg gcgggttcga ctctccaaca gtggcttaca gtgtcctggt ggtcgcaaag 4080

gtggaaaagg ggaagtccaa gaaactgaag tctgtcaaag agctgctggg aatcactatt 4140

atggaacgca gctccttcga gaagaatcct atcgattttc tggaagccaa gggctataaa 4200

gaggtgaaga aagacctgat cattaagctg ccaaaatact cactgtttga gctggaaaac 4260

ggacgaaagc gaatgctggc aagcgccgga gaactgcaga agggcaatga gctggccctg 4320

ccctccaaat acgtgaactt cctgtatctg gctagccact acgagaaact gaaggggtcc 4380

cctgaggata acgaacagaa gcagctgttt gtggagcagc acaaacatta tctggacgag 4440

atcattgaac agatttcaga gttcagcaag agagtgatcc tggctgacgc aaatctggat 4500

aaagtcctga gcgcatacaa caagcaccga gacaaaccaa tccgggagca ggccgaaaat 4560

atcattcatc tgttcaccct gacaaacctg ggcgcccctg cagccttcaa gtattttgac 4620

accacaatcg atcggaagag atacacttct accaaagagg tgctggatgc taccctgatc 4680

caccagagta ttaccggcct gtatgagaca cgcatcgacc tgtcacagct gggaggcgat 4740

gggagcccca agaaaaagcg gaaggtgtct agttaa 4776

<210> 50

<211> 1588

<212> PRT

<213> Artificial Sequence

<220><223> dCas9-Clo051 - 40-18170

<400> 50

Met Pro Lys Lys Lys Arg Lys Val Glu Gly Ile Lys Ser Asn Ile Ser
1 5 10 15

Leu Leu Lys Asp Glu Leu Arg Gly Gln Ile Ser His Ile Ser His Glu
20 25 30

Tyr Leu Ser Leu Ile Asp Leu Ala Phe Asp Ser Lys Gln Asn Arg Leu
35 40 45

Phe Glu Met Lys Val Leu Glu Leu Leu Val Asn Glu Tyr Gly Phe Lys
50 55 60

Gly Arg His Leu Gly Gly Ser Arg Lys Pro Asp Gly Ile Val Tyr Ser
65 70 75 80

Thr Thr Leu Glu Asp Asn Phe Gly Ile Ile Val Asp Thr Lys Ala Tyr
85 90 95

Ser Glu Gly Tyr Ser Leu Pro Ile Ser Gln Ala Asp Glu Met Glu Arg
100 105 110

Tyr Val Arg Glu Asn Ser Asn Arg Asp Glu Glu Val Asn Pro Asn Lys
115 120 125

Trp Trp Glu Asn Phe Ser Glu Glu Val Lys Lys Tyr Tyr Phe Val Phe
130 135 140

Ile Ser Gly Ser Phe Lys Gly Lys Phe Glu Glu Gln Leu Arg Arg Leu
145 150 155 160

Ser Met Thr Thr Gly Val Asn Gly Ser Ala Val Asn Val Val Asn Leu
165 170 175

Leu Leu Gly Ala Glu Lys Ile Arg Ser Gly Glu Met Thr Ile Glu Glu
180 185 190
Leu Glu Arg Ala Met Phe Asn Asn Ser Glu Phe Ile Leu Lys Tyr Gly
195 200 205
Gly Gly Gly Ser Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr
210 215 220

Asn Ser Val Gly Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser
225 230 235 240
Lys Lys Phe Lys Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys
245 250 255
Asn Leu Ile Gly Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala
260 265 270
Thr Arg Leu Lys Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn
275 280 285

Arg Ile Cys Tyr Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val
290 295 300
Asp Asp Ser Phe Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu
305 310 315 320
Asp Lys Lys His Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu
325 330 335
Val Ala Tyr His Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys
340 345 350

Leu Val Asp Ser Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala
355 360 365
Leu Ala His Met Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp
370 375 380
Leu Asn Pro Asp Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val
385 390 395 400
Gln Thr Tyr Asn Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly
405 410 415

Val Asp Ala Lys Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg

420 425 430

Leu Glu Asn Leu Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu

435 440 445

Phe Gly Asn Leu Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys

450 455 460

Ser Asn Phe Asp Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp

465 470 475 480

Thr Tyr Asp Asp Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln

485 490 495

Tyr Ala Asp Leu Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu

500 505 510

Leu Ser Asp Ile Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu

515 520 525

Ser Ala Ser Met Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr

530 535 540

Leu Leu Lys Ala Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu

545 550 555 560

Ile Phe Phe Asp Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly

565 570 575

Gly Ala Ser Gln Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu

580 585 590

Lys Met Asp Gly Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp

595 600 605

Leu Leu Arg Lys Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln

610 615 620

Ile His Leu Gly Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe

625 630 635 640

Tyr Pro Phe Leu Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr

645 650 655

Phe Arg Ile Pro Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg

660 665 670

Phe Ala Trp Met Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn
675 680 685

Phe Glu Glu Val Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu
690 695 700

Arg Met Thr Asn Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro
705 710 715 720

Lys His Ser Leu Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr
725 730 735

Lys Val Lys Tyr Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser
740 745 750

Gly Glu Gln Lys Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg
755 760 765

Lys Val Thr Val Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu
770 775 780

Cys Phe Asp Ser Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala
785 790 795 800

Ser Leu Gly Thr Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp
805 810 815

Phe Leu Asp Asn Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu
820 825 830

Thr Leu Thr Leu Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys
835 840 845

Thr Tyr Ala His Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg
850 855 860

Arg Arg Tyr Thr Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly
865 870 875 880

Ile Arg Asp Lys Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser
885 890 895

Asp Gly Phe Ala Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser
900 905 910

Leu Thr Phe Lys Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly

915 920 925

Asp Ser Leu His Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile
930 935 940
Lys Lys Gly Ile Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys
945 950 955 960
Val Met Gly Arg His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg
965 970 975
Glu Asn Gln Thr Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met
980 985 990

Lys Arg Ile Glu Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys
995 1000 1005
Glu His Pro Val Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr
1010 1015 1020
Leu Tyr Tyr Leu Gln Asn Gly Arg Asp Met Tyr Val Asp Gln Glu
1025 1030 1035
Leu Asp Ile Asn Arg Leu Ser Asp Tyr Asp Val Asp Ala Ile Val
1040 1045 1050

Pro Gln Ser Phe Leu Lys Asp Asp Ser Ile Asp Asn Lys Val Leu
1055 1060 1065
Thr Arg Ser Asp Lys Asn Arg Gly Lys Ser Asp Asn Val Pro Ser
1070 1075 1080
Glu Glu Val Val Lys Lys Met Lys Asn Tyr Trp Arg Gln Leu Leu
1085 1090 1095
Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe Asp Asn Leu Thr Lys
1100 1105 1110

Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys Ala Gly Phe Ile
1115 1120 1125
Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys His Val Ala
1130 1135 1140
Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu Asn Asp
1145 1150 1155

Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys Leu
1160 1165 1170

Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu
1175 1180 1185
Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val
1190 1195 1200
Val Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu
1205 1210 1215
Phe Val Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile
1220 1225 1230

Ala Lys Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe
1235 1240 1245
Phe Tyr Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu
1250 1255 1260
Ala Asn Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly
1265 1270 1275
Glu Thr Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr
1280 1285 1290

Val Arg Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys
1295 1300 1305
Thr Glu Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro
1310 1315 1320
Lys Arg Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp
1325 1330 1335
Pro Lys Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser
1340 1345 1350

Val Leu Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu
1355 1360 1365
Lys Ser Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser
1370 1375 1380
Ser Phe Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr

1385 1390 1395

Lys Glu Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser

1400 1405 1410

Leu Phe Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala

1415 1420 1425

Gly Glu Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr

1430 1435 1440

Val Asn Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly

1445 1450 1455

Ser Pro Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His

1460 1465 1470

Lys His Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser

1475 1480 1485

Lys Arg Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser

1490 1495 1500

Ala Tyr Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu

1505 1510 1515

Asn Ile Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala

1520 1525 1530

Ala Phe Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr

1535 1540 1545

Ser Thr Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile

1550 1555 1560

Thr Gly Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly

1565 1570 1575

Asp Gly Ser Pro Lys Lys Lys Arg Lys Val

1580 1585

<210> 51

<211> 4767

<212>

DNA

<213> Artificial Sequence

<220><223> dCas9-Clo051 - 40-18171

<400> 51

atgcctaaga agaagcggaa ggtggaaggc atcaaaagca acatctccct cctgaaagac 60

gaactccggg ggcagattag ccacattagt cacgaatacc tctccctcat cgacctggct 120

ttcgatagca agcagaacag gctctttgag atgaaagtgc tggaactgct cgtcaatgag 180

tacgggttca agggtcgaca cctcggcgga tctaggaaac cagacggcat cgtgtatagt 240

accacactgg aagacaactt tgggatcatt gtggatacca aggcatactc tgagggttat 300

agtctgccca tttcacaggc cgacgagatg gaacggtacg tgcgcgagaa ctcaaataga 360

gatgaggaag tcaaccctaa caagtggtgg gagaacttct ctgaggaagt gaagaaatac 420

tacttcgtct ttatcagcgg gtccttcaag ggtaaatttg aggaacagct caggagactg 480

agcatgacta ccggcgtgaa tggcagcgcc gtcaacgtgg tcaatctgct cctgggcgct 540

gaaaagattc ggagcggaga gatgaccatc gaagagctgg agagggcaat gtttaataat 600

agcgagttta tcctgaaata cggtggcggt ggatccgata aaaagtattc tattggttta 660

gccatcggca ctaattccgt tggatgggct gtcataaccg atgaatacaa agtaccttca 720

aagaaattta aggtgttggg gaacacagac cgtcattcga ttaaaaagaa tcttatcggt 780

gccctcctat tcgatagtgg cgaaacggca gaggcgactc gcctgaaacg aaccgctcgg 840

agaaggtata cacgtcgcaa gaaccgaata tgttacttac aagaaatttt tagcaatgag 900

atggccaaag ttgacgattc tttctttcac cgtttggaag agtccttcct tgtcgaagag 960

gacaagaaac atgaacggca ccccatcttt ggaaacatag tagatgaggt ggcatatcat 1020

gaaaagtacc caacgattta tcacctcaga aaaaagctag ttgactcaac tgataaagcg 1080

gacctgaggt taatctactt ggctcttgcc catatgataa agttccgtgg gcactttctc 1140

attgagggtg atctaaatcc ggacaactcg gatgtcgaca aactgttcat ccagttagta 1200

caaacctata atcagttgtt tgaagagaac cctataaatg caagtggcgt ggatgcgaag 1260

gctattctta gcgcccgcct ctctaaatcc cgacggctag aaaacctgat cgcacaatta 1320

cccggagaga agaaaaatgg gttgttcggt aaccttatag cgctctcact aggcctgaca 1380

ccaaatttta agtcgaactt cgacttagct gaagatgcca aattgcagct tagtaaggac 1440

acgtacgatg acgatctcga caatctactg gcacaaattg gagatcagta tgcggactta 1500

tttttggctg ccaaaaacct tagcgatgca atcctcctat ctgacatact gagagttaat 1560

actgagatta ccaaggcgcc gttatccgct tcaatgatca aaaggtacga tgaacatcac 1620

caagacttga cacttctcaa ggccctagtc cgtcagcaac tgcctgagaa atataaggaa 1680

atattctttg atcagtcgaa aaacgggtac gcaggttata ttgacggcgg agcgagtcaa 1740

gaggaattct acaagtttat caaacccata ttagagaaga tggatgggac ggaagagttg 1800

cttgtaaaac tcaatcgcga agatctactg cgaaagcagc ggactttcga caacggtagc 1860

attccacatc aaatccactt aggcgaattg catgctatac ttagaaggca ggaggatttt 1920

tatccgttcc tcaaagacaa tcgtgaaaag attgagaaaa tcctaacctt tcgcatacct 1980

tactatgtgg gacccctggc ccgagggaac tctcggttcg catggatgac aagaaagtcc 2040

gaagaaacga ttactccatg gaattttgag gaagttgtcg ataaaggtgc gtcagctcaa 2100

tcgttcatcg agaggatgac caactttgac aagaatttac cgaacgaaaa agtattgcct 2160

aagcacagtt tactttacga gtatttcaca gtgtacaatg aactcacgaa agttaagtat 2220

gtcactgagg gcatgcgtaa acccgccttt ctaagcggag aacagaagaa agcaatagta 2280

gatctgttat tcaagaccaa ccgcaaagtg acagttaagc aattgaaaga ggactacttt 2340

aagaaaattg aatgcttcga ttctgtcgag atctccgggg tagaagatcg atttaatgcg 2400

tcacttggta cgtatcatga cctcctaaag ataattaaag ataaggactt cctggataac 2460

gaagagaatg aagatatctt agaagatata gtgttgactc ttaccctctt tgaagatcgg 2520

gaaatgattg aggaaagact aaaaacatac gctcacctgt tcgacgataa ggttatgaaa 2580

cagttaaaga ggcgtcgcta tacgggctgg ggacgattgt cgcggaaact tatcaacggg 2640

ataagagaca agcaaagtgg taaaactatt ctcgattttc taaagagcga cggcttcgcc 2700

aataggaact ttatgcagct gatccatgat gactctttaa ccttcaaaga ggatatacaa 2760

aaggcacagg tttccggaca aggggactca ttgcacgaac atattgcgaa tcttgctggt 2820

tcgccagcca tcaaaaaggg catactccag acagtcaaag tagtggatga gctagttaag 2880

gtcatgggac gtcacaaacc ggaaaacatt gtaatcgaga tggcacgcga aaatcaaacg 2940

actcagaagg ggcaaaaaaa cagtcgagag cggatgaaga gaatagaaga gggtattaaa 3000

gaactgggca gccagatctt aaaggagcat cctgtggaaa atacccaatt gcagaacgag 3060

aaactttacc tctattacct acaaaatgga agggacatgt atgttgatca ggaactggac 3120

ataaaccgtt tatctgatta cgacgtcgat gccattgtac cccaatcctt tttgaaggac 3180

gattcaatcg acaataaagt gcttacacgc tcggataaga accgagggaa aagtgacaat 3240

gttccaagcg aggaagtcgt aaagaaaatg aagaactatt ggcggcagct cctaaatgcg 3300

aaactgataa cgcaaagaaa gttcgataac ttaactaaag ctgagagggg tggcttgtct 3360

gaacttgaca aggccggatt tattaaacgt cagctcgtgg aaacccgcca aatcacaaag 3420

catgttgcac agatactaga ttcccgaatg aatacgaaat acgacgagaa cgataagctg 3480

attcgggaag tcaaagtaat cactttaaag tcaaaattgg tgtcggactt cagaaaggat 3540

tttcaattct ataaagttag ggagataaat aactaccacc atgcgcacga cgcttatctt 3600

aatgccgtcg tagggaccgc actcattaag aaatacccga agctagaaag tgagtttgtg 3660

tatggtgatt acaaagttta tgacgtccgt aagatgatcg cgaaaagcga acaggagata 3720

ggcaaggcta cagccaaata cttcttttat tctaacatta tgaatttctt taagacggaa 3780

atcactctgg caaacggaga gatacgcaaa cgacctttaa ttgaaaccaa tggggagaca 3840

ggtgaaatcg tatgggataa gggccgggac ttcgcgacgg tgagaaaagt tttgtccatg 3900

ccccaagtca acatagtaaa gaaaactgag gtgcagaccg gagggttttc aaaggaatcg 3960

attcttccaa aaaggaatag tgataagctc atcgctcgta aaaaggactg ggacccgaaa 4020

aagtacggtg gcttcgatag ccctacagtt gcctattctg tcctagtagt ggcaaaagtt 4080

gagaagggaa aatccaagaa actgaagtca gtcaaagaat tattggggat aacgattatg 4140

gagcgctcgt cttttgaaaa gaaccccatc gacttccttg aggcgaaagg ttacaaggaa 4200

gtaaaaaagg atctcataat taaactacca aagtatagtc tgtttgagtt agaaaatggc 4260

cgaaaacgga tgttggctag cgccggagag cttcaaaagg ggaacgaact cgcactaccg 4320

tctaaatacg tgaatttcct gtatttagcg tcccattacg agaagttgaa aggttcacct 4380

gaagataacg aacagaagca acttttttgtt gagcagcaca aacattatct cgacgaaatc 4440

atagagcaaa tttcggaatt cagtaagaga gtcatcctag ctgatgccaa tctggacaaa 4500

gtattaagcg catacaacaa gcacagggat aaacccatac gtgagcaggc ggaaaatatt 4560

atccatttgt ttactcttac caacctcggc gctccagccg cattcaagta ttttgacaca 4620

acgatagatc gcaaacgata cacttctacc aaggaggtgc tagacgcgac actgattcac 4680

caatccatca cgggattata tgaaactcgg atagatttgt cacagcttgg gggtgacgga 4740

tcccccaaga agaagaggaa agtctga 4767

<210> 52

<211> 187

<212> PRT

<213> Artificial Sequence

<220><223> DHFR mutein - 40-17012

<400> 52

Met Val Gly Ser Leu Asn Cys Ile Val Ala Val Ser Gln Asn Met Gly

1 5 10 15

Ile Gly Lys Asn Gly Asp Phe Pro Trp Pro Pro Leu Arg Asn Glu Ser

20 25 30

Arg Tyr Phe Gln Arg Met Thr Thr Thr Ser Ser Val Glu Gly Lys Gln

35 40 45

Asn Leu Val Ile Met Gly Lys Lys Thr Trp Phe Ser Ile Pro Glu Lys

50 55 60

Asn Arg Pro Leu Lys Gly Arg Ile Asn Leu Val Leu Ser Arg Glu Leu

65 70 75 80

Lys Glu Pro Pro Gln Gly Ala His Phe Leu Ser Arg Ser Leu Asp Asp

85 90 95

Ala Leu Lys Leu Thr Glu Gln Pro Glu Leu Ala Asn Lys Val Asp Met

100 105 110

Val Trp Ile Val Gly Gly Ser Ser Val Tyr Lys Glu Ala Met Asn His

115 120 125

Pro Gly His Leu Lys Leu Phe Val Thr Arg Ile Met Gln Asp Phe Glu

130 135 140

Ser Asp Thr Phe Phe Pro Glu Ile Asp Leu Glu Lys Tyr Lys Leu Leu

145 150 155 160

Pro Glu Tyr Pro Gly Val Leu Ser Asp Val Gln Glu Glu Lys Gly Ile

165 170 175

Lys Tyr Lys Phe Glu Val Tyr Glu Lys Asn Asp

180 185

<210> 53

<211> 561

<212> DNA

<213> Artificial Sequence

<220><223> DHFR mutein - 40-17095

<400> 53

atggtcgggt ctctgaattg tatcgtcgcc gtgagtcaga acatgggcat tgggaagaat 60

ggcgatttcc catggccacc tctgcgcaac gagtcccgat actttcagcg gatgacaact 120

acctcctctg tggaagggaa acagaatctg gtcatcatgg gaaagaaaac ttggttcagc 180

attccagaga agaaccggcc cctgaaaggc agaatcaatc tggtgctgtc ccgagaactg 240

aaggagccac cacagggagc tcactttctg agccggtccc tggacgatgc actgaagctg 300

acagaacagc ctgagctggc caacaaagtc gatatggtgt ggatcgtcgg gggaagttca 360

gtgtataagg aggccatgaa tcaccccggc catctgaaac tgttcgtcac acggatcatg 420

caggactttg agagcgatac tttctttcct gaaattgacc tggagaagta caaactgctg 480

cccgaatatc ctggcgtgct gtccgatgtc caggaagaga aaggcatcaa atacaagttc 540

gaggtctatg agaagaatga c 561

<210> 54

<211> 18

<212> PRT

<213> Homo Sapiens

<400> 54

Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro

1 5 10 15

Gly Pro

<210> 55

<211> 21

<212> PRT

<213> Homo Sapiens

<400> 55

Gly Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu

1 5 10 15

Glu Asn Pro Gly Pro

20

<210> 56

<211> 63

<212> DNA

<213> Homo Sapiens

<400> 56

ggatctggag agggaagggg aagcctgctg acctgtggag acgtggagga aaacccagga 60

cca 63

<210> 57

<211> 20

<212> PRT

<213> Homo Sapiens

<400> 57

Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly Asp Val Glu Ser

1 5 10 15

Asn Pro Gly Pro

20

<210> 58

<211> 23

<212> PRT

<213> Homo Sapiens

<400> 58

Gly Ser Gly Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly Asp

1 5 10 15

Val Glu Ser Asn Pro Gly Pro

20

<210> 59

<211> 22

<212> PRT

<213> Homo Sapiens

<400> 59

Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val

1 5 10 15

Glu Ser Asn Pro Gly Pro

20

<210> 60

<211> 25

<212> PRT

<213> Homo Sapiens

<400> 60

Gly Ser Gly Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala

1 5 10 15

Gly Asp Val Glu Ser Asn Pro Gly Pro

20 25

<210> 61

<211> 19

<212> PRT

<213> Homo Sapiens

<400> 61

Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn

1 5 10 15

Pro Gly Pro

<210> 62

<211> 22

<212> PRT

<213> Homo Sapiens

<400> 62

Gly Ser Gly Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val

1 5 10 15

Glu Glu Asn Pro Gly Pro

20

<210> 63

<211> 594

<212> PRT

<213> Artificial Sequence

<220><223> T2A - 40-18099

<400> 63

Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln

1 5 10 15

Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Ile Ser Asp

20 25 30

His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile

35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu

50 55 60

Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn

65 70 75 80

Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95
Cys Trp Ser Thr Ser Lys Ser Thr Arg Arg Ser Arg Val Ser Ala Leu
100 105 110
Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
115 120 125
Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
130 135 140

Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
145 150 155 160
Glu Ser Met Thr Gly Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile
165 170 175
Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
180 185 190
His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
195 200 205

Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
210 215 220
Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225 230 235 240
Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
245 250 255
Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
260 265 270

Gly Phe Arg Gly Arg Cys Pro Phe Arg Met Tyr Ile Pro Asn Lys Pro
275 280 285
Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Tyr Lys
290 295 300
Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
305 310 315 320
Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val

325 330 335

His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile
340 345 350
Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
355 360 365
Gly Thr Val Arg Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
370 375 380
Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385 390 395 400

Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
405 410 415
Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
420 425 430
Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445
Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
450 455 460

Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465 470 475 480
Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495
Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Ser Leu Thr Ser
500 505 510
Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525

Arg Asp Asn Ile Ser Asn Ile Leu Pro Asn Glu Val Pro Gly Thr Ser
530 535 540
Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
545 550 555 560
Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Asn Ala Ser Cys Lys Lys
565 570 575

Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
580 585 590

Cys Phe

<210> 64
<211> 594
<212> PRT
<213> Artificial Sequence
<220><223> SPB transposase -40-14484
<400> 64
Met Gly Ser Ser Leu Asp Asp Glu His Ile Leu Ser Ala Leu Leu Gln
1 5 10 15
Ser Asp Asp Glu Leu Val Gly Glu Asp Ser Asp Ser Glu Val Ser Asp
20 25 30
His Val Ser Glu Asp Asp Val Gln Ser Asp Thr Glu Glu Ala Phe Ile
35 40 45

Asp Glu Val His Glu Val Gln Pro Thr Ser Ser Gly Ser Glu Ile Leu
50 55 60
Asp Glu Gln Asn Val Ile Glu Gln Pro Gly Ser Ser Leu Ala Ser Asn
65 70 75 80
Arg Ile Leu Thr Leu Pro Gln Arg Thr Ile Arg Gly Lys Asn Lys His
85 90 95
Cys Trp Ser Thr Ser Lys Ser Thr Arg Arg Ser Arg Val Ser Ala Leu
100 105 110

Asn Ile Val Arg Ser Gln Arg Gly Pro Thr Arg Met Cys Arg Asn Ile
115 120 125
Tyr Asp Pro Leu Leu Cys Phe Lys Leu Phe Phe Thr Asp Glu Ile Ile
130 135 140
Ser Glu Ile Val Lys Trp Thr Asn Ala Glu Ile Ser Leu Lys Arg Arg
145 150 155 160
Glu Ser Met Thr Ser Ala Thr Phe Arg Asp Thr Asn Glu Asp Glu Ile

165 170 175

Tyr Ala Phe Phe Gly Ile Leu Val Met Thr Ala Val Arg Lys Asp Asn
180 185 190
His Met Ser Thr Asp Asp Leu Phe Asp Arg Ser Leu Ser Met Val Tyr
195 200 205
Val Ser Val Met Ser Arg Asp Arg Phe Asp Phe Leu Ile Arg Cys Leu
210 215 220
Arg Met Asp Asp Lys Ser Ile Arg Pro Thr Leu Arg Glu Asn Asp Val
225 230 235 240

Phe Thr Pro Val Arg Lys Ile Trp Asp Leu Phe Ile His Gln Cys Ile
245 250 255
Gln Asn Tyr Thr Pro Gly Ala His Leu Thr Ile Asp Glu Gln Leu Leu
260 265 270
Gly Phe Arg Gly Arg Cys Pro Phe Arg Val Tyr Ile Pro Asn Lys Pro
275 280 285
Ser Lys Tyr Gly Ile Lys Ile Leu Met Met Cys Asp Ser Gly Thr Lys
290 295 300

Tyr Met Ile Asn Gly Met Pro Tyr Leu Gly Arg Gly Thr Gln Thr Asn
305 310 315 320
Gly Val Pro Leu Gly Glu Tyr Tyr Val Lys Glu Leu Ser Lys Pro Val
325 330 335
His Gly Ser Cys Arg Asn Ile Thr Cys Asp Asn Trp Phe Thr Ser Ile
340 345 350
Pro Leu Ala Lys Asn Leu Leu Gln Glu Pro Tyr Lys Leu Thr Ile Val
355 360 365

Gly Thr Val Arg Ser Asn Lys Arg Glu Ile Pro Glu Val Leu Lys Asn
370 375 380
Ser Arg Ser Arg Pro Val Gly Thr Ser Met Phe Cys Phe Asp Gly Pro
385 390 395 400
Leu Thr Leu Val Ser Tyr Lys Pro Lys Pro Ala Lys Met Val Tyr Leu
405 410 415

Leu Ser Ser Cys Asp Glu Asp Ala Ser Ile Asn Glu Ser Thr Gly Lys
420 425 430

Pro Gln Met Val Met Tyr Tyr Asn Gln Thr Lys Gly Gly Val Asp Thr
435 440 445
Leu Asp Gln Met Cys Ser Val Met Thr Cys Ser Arg Lys Thr Asn Arg
450 455 460
Trp Pro Met Ala Leu Leu Tyr Gly Met Ile Asn Ile Ala Cys Ile Asn
465 470 475 480
Ser Phe Ile Ile Tyr Ser His Asn Val Ser Ser Lys Gly Glu Lys Val
485 490 495

Gln Ser Arg Lys Lys Phe Met Arg Asn Leu Tyr Met Ser Leu Thr Ser
500 505 510
Ser Phe Met Arg Lys Arg Leu Glu Ala Pro Thr Leu Lys Arg Tyr Leu
515 520 525
Arg Asp Asn Ile Ser Asn Ile Leu Pro Lys Glu Val Pro Gly Thr Ser
530 535 540
Asp Asp Ser Thr Glu Glu Pro Val Met Lys Lys Arg Thr Tyr Cys Thr
545 550 555 560

Tyr Cys Pro Ser Lys Ile Arg Arg Lys Ala Asn Ala Ser Cys Lys Lys
565 570 575
Cys Lys Lys Val Ile Cys Arg Glu His Asn Ile Asp Met Cys Gln Ser
580 585 590
Cys Phe

<210> 65
<211> 340
<212> PRT
<213> Artificial Sequence
<220><223> SB100X transposase - 40-14485
<400> 65
Met Gly Lys Ser Lys Glu Ile Ser Gln Asp Leu Arg Lys Lys Ile Val

1 5 10 15

Asp Leu His Lys Ser Gly Ser Ser Leu Gly Ala Ile Ser Lys Arg Leu
20 25 30
Lys Val Pro Arg Ser Ser Val Gln Thr Ile Val Arg Lys Tyr Lys His
35 40 45
His Gly Thr Thr Gln Pro Ser Tyr Arg Ser Gly Arg Arg Arg Val Leu
50 55 60
Ser Pro Arg Asp Glu Arg Thr Leu Val Arg Lys Val Gln Ile Asn Pro
65 70 75 80

Arg Thr Thr Ala Lys Asp Leu Val Lys Met Leu Glu Glu Thr Gly Thr
85 90 95
Lys Val Ser Ile Ser Thr Val Lys Arg Val Leu Tyr Arg His Asn Leu
100 105 110
Lys Gly Arg Ser Ala Arg Lys Lys Pro Leu Leu Gln Asn Arg His Lys
115 120 125
Lys Ala Arg Leu Arg Phe Ala Thr Ala His Gly Asp Lys Asp Arg Thr
130 135 140

Phe Trp Arg Asn Val Leu Trp Ser Asp Glu Thr Lys Ile Glu Leu Phe
145 150 155 160
Gly His Asn Asp His Arg Tyr Val Trp Arg Lys Lys Gly Glu Ala Cys
165 170 175
Lys Pro Lys Asn Thr Ile Pro Thr Val Lys His Gly Gly Gly Ser Ile
180 185 190
Met Leu Trp Gly Cys Phe Ala Ala Gly Gly Thr Gly Ala Leu His Lys
195 200 205

Ile Asp Gly Ile Met Arg Lys Glu Asn Tyr Val Asp Ile Leu Lys Gln
210 215 220
His Leu Lys Thr Ser Val Arg Lys Leu Lys Leu Gly Arg Lys Trp Val
225 230 235 240
Phe Gln Met Asp Asn Asp Pro Lys His Thr Ser Lys Val Val Ala Lys
245 250 255

Trp Leu Lys Asp Asn Lys Val Lys Val Leu Glu Trp Pro Ser Gln Ser
260 265 270

Pro Asp Leu Asn Pro Ile Glu Asn Leu Trp Ala Glu Leu Lys Lys Arg
275 280 285

Val Arg Ala Arg Arg Pro Thr Asn Leu Thr Gln Leu His Gln Leu Cys
290 295 300

Gln Glu Glu Trp Ala Lys Ile His Pro Thr Tyr Cys Gly Lys Leu Val
305 310 315 320

Glu Gly Tyr Pro Lys Arg Leu Thr Gln Val Lys Gln Phe Lys Gly Asn
325 330 335

Ala Thr Lys Tyr
340

<210> 66

<211> 340

<212> PRT

<213> Artificial Sequence

<220><223> hyperSB100X transposase -40-14486

<400> 66

Met Gly Lys Ser Lys Glu Ile Ser Gln Asp Leu Arg Lys Arg Ile Val
1 5 10 15

Asp Leu His Lys Ser Gly Ser Ser Leu Gly Ala Ile Ser Lys Arg Leu
20 25 30

Ala Val Pro Arg Ser Ser Val Gln Thr Ile Val Arg Lys Tyr Lys His
35 40 45

His Gly Thr Thr Gln Pro Ser Tyr Arg Ser Gly Arg Arg Arg Val Leu
50 55 60

Ser Pro Arg Asp Glu Arg Thr Leu Val Arg Lys Val Gln Ile Asn Pro
65 70 75 80

Arg Thr Thr Ala Lys Asp Leu Val Lys Met Leu Glu Glu Thr Gly Thr
85 90 95

Lys Val Ser Ile Ser Thr Val Lys Arg Val Leu Tyr Arg His Asn Leu

100 105 110

Lys Gly His Ser Ala Arg Lys Lys Pro Leu Leu Gln Asn Arg His Lys
115 120 125
Lys Ala Arg Leu Arg Phe Ala Thr Ala His Gly Asp Lys Asp Arg Thr
130 135 140
Phe Trp Arg Asn Val Leu Trp Ser Asp Glu Thr Lys Ile Glu Leu Phe
145 150 155 160
Gly His Asn Asp His Arg Tyr Val Trp Arg Lys Lys Gly Glu Ala Cys
165 170 175

Lys Pro Lys Asn Thr Ile Pro Thr Val Lys His Gly Gly Gly Ser Ile
180 185 190
Met Leu Trp Gly Cys Phe Ala Ala Gly Gly Thr Gly Ala Leu His Lys
195 200 205
Ile Asp Gly Ile Met Asp Ala Val Gln Tyr Val Asp Ile Leu Lys Gln
210 215 220
His Leu Lys Thr Ser Val Arg Lys Leu Lys Leu Gly Arg Lys Trp Val
225 230 235 240

Phe Gln His Asp Asn Asp Pro Lys His Thr Ser Lys Val Val Ala Lys
245 250 255
Trp Leu Lys Asp Asn Lys Val Lys Val Leu Glu Trp Pro Ser Gln Ser
260 265 270
Pro Asp Leu Asn Pro Ile Glu Asn Leu Trp Ala Glu Leu Lys Lys Arg
275 280 285
Val Arg Ala Arg Arg Pro Thr Asn Leu Thr Gln Leu His Gln Leu Cys
290 295 300

Gln Glu Glu Trp Ala Lys Ile His Pro Asn Tyr Cys Gly Lys Leu Val
305 310 315 320
Glu Gly Tyr Pro Lys Arg Leu Thr Gln Val Lys Gln Phe Lys Gly Asn
325 330 335
Ala Thr Lys Tyr
340

<210> 67

<211> 5296

<212> DNA

<213> Artificial Sequence

<220><223> Helraiser transposon - 40-18174

<400> 67

tcctatataa taaaagagaa acatgcaaat tgaccatccc tccgctacgc tcaagccacg 60

cccaccagcc aatcagaagt gactatgcaa attaacccaa caaagatggc agttaaattt 120

gcatacgcag gtgtcaagcg ccccaggagg caacggcggc cgcgggctcc caggaccttc 180

gctggccccg ggaggcgagg ccggccgcgc ctagccacac ccgcgggctc ccgggacctt 240

cgccagcaga gagcagagcg ggagagcggg cggagagcgg gaggtttgga ggacttggca 300

gagcaggagg ccgctggaca tagagcagag cgagagagag ggtggcttgg agggcgtggc 360

tccctctgtc accccagctt cctcatcaca gctgtggaaa ctgacagcag ggaggaggaa 420

gtcccacccc cacagaatca gccagaatca gccgttggtc agacagctct cagcggcctg 480

acagccagga ctctcattca cctgcatctc agaccgtgac agtagagagg tgggactatg 540

tctaaagaac aactgttgat acaacgtagc tctgcagccg aaagatgccg gcgttatcga 600

cagaaaatgt ctgcagagca acgtgcgtct gatcttgaaa gaaggcggcg cctgcaacag 660

aatgtatctg aagagcagct actggaaaaa cgtcgctctg aagccgaaaa acagcggcgt 720

catcgacaga aaatgtctaa agaccaacgt gcctttgaag ttgaaagaag gcggtggcga 780

cgacagaata tgtctagaga acagtcatca acaagtacta ccaataccgg taggaactgc 840

cttctcagca aaaatggagt acatgaggat gcaattctcg aacatagttg tggtggaatg 900

actgttcgat gtgaattttg cctatcacta aatttctctg atgaaaaacc atccgatggg 960

aaatttactc gatgttgtag caaagggaaa gtctgtccaa atgatataca ttttccagat 1020

tacccggcat atttaaaaag attaatgaca aacgaagatt ctgacagtaa aaatttcatg 1080

gaaaatattc gttccataaa tagttctttt gcttttgctt ccatgggtgc aaatattgca 1140

tcgccatcag gatatgggcc atactgtttt agaatacacg gacaagttta tcaccgtact 1200

ggaactttac atccttcgga tggtgtttct cggaagtttg ctcaactcta tattttggat 1260

acagccgaag ctacaagtaa aagattagca atgccagaaa accagggctg ctcagaaaga 1320

ctcatgatca acatcaacaa cctcatgcat gaaataaatg aattaacaaa atcgtacaag 1380

atgctacatg aggtagaaaa ggaagcccaa tctgaagcag cagcaaaagg tattgctccc 1440

acagaagtaa caatggcgat taaatacgat cgtaacagtg acccaggtag atataattct 1500

ccccgtgtaa ccgaggttgc tgtcatattc agaaacgaag atggagaacc tccttttgaa 1560

agggacttgc tcattcattg taaaccagat cccaataatc caaatgccac taaaatgaaa 1620

caaatcagta tcctgtttcc tacattagat gcaatgacat atcctattct ttttccacat 1680

ggtgaaaaag gctggggaac agatattgca ttaagactca gagacaacag tgtaatcgac 1740

aataatacta gacaaaatgt aaggacacga gtcacacaaa tgcagtatta tggatttcat 1800

ctctctgtgc gggacacgtt caatcctatt ttaaatgcag gaaaattaac tcaacagttt 1860

attgtggatt catattcaaa aatggaggcc aatcggataa atttcatcaa agcaaaccaa 1920

tctaagttga gagttgaaaa atatagtggt ttgatggatt atctcaaatc tagatctgaa 1980

aatgacaatg tgccgattgg taaaatgata atacttccat catcttttga gggtagtccc 2040

agaaatatgc agcagcgata tcaggatgct atggcaattg taacgaagta tggcaagccc 2100

gatttattca taaccatgac atgcaacccc aaatgggcag atattacaaa caatttacaa 2160

cgctggcaaa aagttgaaaa cagacctgac ttggtagcca gagtttttaa tattaagctg 2220

aatgctcttt taaatgatat atgtaaattc catttatttg gcaaagtaat agctaaaatt 2280

catgtcattg aatttcagaa acgcggactg cctcacgctc acatattatt gatattagat 2340

agtgagtcca aattacgttc agaagatgac attgaccgta tagttaaggc agaaattcca 2400

gatgaagacc agtgtcctcg acttttttcaa attgtaaaat caaatatggt acatggacca 2460

tgtggaatac aaaatccaaa tagtccatgt atggaaaatg gaaaatgttc aaagggatat 2520

ccaaaagaat ttcaaaatgc gaccattgga aatattgatg gatatcccaa atacaaacga 2580

agatctggta gcaccatgtc tattggaaat aaagttgtcg ataacacttg gattgtccct 2640

tataacccgt atttgtgcct taaatataac tgtcatataa atgttgaagt ctgtgcatca 2700

attaaaagtg tcaaatattt atttaaatac atctataaag ggcacgattg tgcaaatatt 2760

caaatttctg aaaaaaatat tatcaatcat gacgaagtac aggacttcat tgactccagg 2820

tatgtgagcg ctcctgaggc tgtttggaga ctttttgcaa tgcgaatgca tgaccaatct 2880

catgcaatca caagattagc tattcatttg ccaaatgatc agaatttgta ttttcatacc 2940

gatgattttg ctgaagtttt agatagggct aaaaggcata actcgacttt gatggcttgg 3000

ttcttattga atagagaaga ttctgatgca cgtaattatt attattggga gattccacag 3060

cattatgtgt ttaataattc tttgtggaca aaacgccgaa agggtgggaa taaagtatta 3120

ggtagactgt tcactgtgag ctttagagaa ccagaacgat attaccttag acttttgctt 3180

ctgcatgtaa aaggtgcgat aagttttgag gatctgcgaa ctgtaggagg tgtaacttat 3240

gatacatttc atgaagctgc taaacaccga ggattattac ttgatgacac tatctggaaa 3300

gatacgattg acgatgcaat catccttaat atgcccaaac aactacggca acttttgca 3360

tatatatgtg tgtttggatg tccttctgct gcagacaaat tatgggatga gaataaatct 3420

cattttattg aagatttctg ttggaaatta caccgaagag aaggtgcctg tgtgaactgt 3480

gaaatgcatg cccttaacga attcaggag gtattcacat tgcatggaat gaaatgttca 3540

catttcaaac ttccggacta tcctttatta atgaatgcaa atacatgtga tcaattgtac 3600

gagcaacaac aggcagaggt tttgataaat tctctgaatg atgaacagtt ggcagccttt 3660

cagactataa cttcagccat cgaagatcaa actgtacacc ccaaatgctt tttcttggat 3720

ggtccaggtg gtagtggaaa aacatatctg tataaagttt taacacatta tattagaggt 3780

cgtggtggta ctgttttacc cacagcatct acaggaattg ctgcaaattt acttcttggt 3840

ggaagaacct ttcattccca atataaatta ccaattccat taaatgaaac ttcaatttct 3900

agactcgata taaagagtga agttgctaaa accattaaaa aggcccaact tctcattatt 3960

gatgaatgca ccatggcatc cagtcatgct ataaacgcca tagatagatt actaagagaa 4020

attatgaatt tgaatgttgc atttggtggg aaagttctcc ttctcggagg ggattttcga 4080

caatgtctca gtattgtacc acatgctatg cgatcggcca tagtacaaac gagtttaaag 4140

tactgtaatg tttggggatg tttcagaaag ttgtctctta aaacaaatat gagatcagag 4200

gattctgctt atagtgaatg gttagtaaaa cttggagatg gcaaacttga tagcagtttt 4260

catttaggaa tggatattat tgaaatcccc catgaaatga tttgtaacgg atctattatt 4320

gaagctacct ttggaaatag tatatctata gataatatta aaaatatatc taaacgtgca 4380

attctttgtc caaaaaatga gcatgttcaa aaattaaatg aagaaatttt ggatatactt 4440

gatggagatt ttcacacata tttgagtgat gattccattg attcaacaga tgatgctgaa 4500

aaggaaaatt ttcccatcga atttcttaat agtattactc cttcgggaat gccgtgtcat 4560

aaattaaaat tgaaagtggg tgcaatcatc atgctattga gaaatcttaa tagtaaatgg 4620

ggtctttgta atggtactag atttattatc aaaagattac gacctaacat tatcgaagct 4680

gaagtattaa caggatctgc agagggagag gttgttctga ttccaagaat tgatttgtcc 4740

ccatctgaca ctggcctccc atttaaatta attcgaagac agtttcccgt gatgccagca 4800

tttgcgatga ctattaataa atcacaagga caaactctag acagagtagg aatattccta 4860

cctgaacccg ttttcgcaca tggtcagtta tatgttgctt tctctcgagt tcgaagagca 4920

tgtgacgtta aagttaaagt tgtaaatact tcatcacaag ggaaattagt caagcactct 4980

gaaagtgttt ttactcttaa tgtggtatac agggagatat tagaataagt ttaatcactt 5040

tatcagtcat tgtttgcatc aatgttgttt ttatatcatg tttttgttgt ttttatatca 5100

tgtctttgtt gttgttatat catgttgtta ttgtttattt attaataaat ttatgtatta 5160

ttttcatata cattttactc atttcctttc atctctcaca cttctattat agagaaaggg 5220

caaatagcaa tattaaaata tttcctctaa ttaattccct ttcaatgtgc acgaatttcg 5280

tgcaccgggc cactag 5296

<210> 68

<211> 1496

<212> PRT

<213> Artificial Sequence

<220><223> Helitron transposase - 40-14488

<400> 68

Met Ser Lys Glu Gln Leu Leu Ile Gln Arg Ser Ser Ala Ala Glu Arg
1 5 10 15

Cys Arg Arg Tyr Arg Gln Lys Met Ser Ala Glu Gln Arg Ala Ser Asp
20 25 30

Leu Glu Arg Arg Arg Arg Leu Gln Gln Asn Val Ser Glu Glu Gln Leu
35 40 45

Leu Glu Lys Arg Arg Ser Glu Ala Glu Lys Gln Arg Arg His Arg Gln
50 55 60

Lys Met Ser Lys Asp Gln Arg Ala Phe Glu Val Glu Arg Arg Arg Trp
65 70 75 80

Arg Arg Gln Asn Met Ser Arg Glu Gln Ser Ser Thr Ser Thr Thr Asn
85 90 95

Thr Gly Arg Asn Cys Leu Leu Ser Lys Asn Gly Val His Glu Asp Ala
100 105 110

Ile Leu Glu His Ser Cys Gly Gly Met Thr Val Arg Cys Glu Phe Cys
115 120 125

Leu Ser Leu Asn Phe Ser Asp Glu Lys Pro Ser Asp Gly Lys Phe Thr
130 135 140

Arg Cys Cys Ser Lys Gly Lys Val Cys Pro Asn Asp Ile His Phe Pro
145 150 155 160

Asp Tyr Pro Ala Tyr Leu Lys Arg Leu Met Thr Asn Glu Asp Ser Asp

165 170 175

Ser Lys Asn Phe Met Glu Asn Ile Arg Ser Ile Asn Ser Ser Phe Ala

180 185 190

Phe Ala Ser Met Gly Ala Asn Ile Ala Ser Pro Ser Gly Tyr Gly Pro

195 200 205

Tyr Cys Phe Arg Ile His Gly Gln Val Tyr His Arg Thr Gly Thr Leu

210 215 220

His Pro Ser Asp Gly Val Ser Arg Lys Phe Ala Gln Leu Tyr Ile Leu

225 230 235 240

Asp Thr Ala Glu Ala Thr Ser Lys Arg Leu Ala Met Pro Glu Asn Gln

245 250 255

Gly Cys Ser Glu Arg Leu Met Ile Asn Ile Asn Asn Leu Met His Glu

260 265 270

Ile Asn Glu Leu Thr Lys Ser Tyr Lys Met Leu His Glu Val Glu Lys

275 280 285

Glu Ala Gln Ser Glu Ala Ala Ala Lys Gly Ile Ala Pro Thr Glu Val

290 295 300

Thr Met Ala Ile Lys Tyr Asp Arg Asn Ser Asp Pro Gly Arg Tyr Asn

305 310 315 320

Ser Pro Arg Val Thr Glu Val Ala Val Ile Phe Arg Asn Glu Asp Gly

325 330 335

Glu Pro Pro Phe Glu Arg Asp Leu Leu Ile His Cys Lys Pro Asp Pro

340 345 350

Asn Asn Pro Asn Ala Thr Lys Met Lys Gln Ile Ser Ile Leu Phe Pro

355 360 365

Thr Leu Asp Ala Met Thr Tyr Pro Ile Leu Phe Pro His Gly Glu Lys

370 375 380

Gly Trp Gly Thr Asp Ile Ala Leu Arg Leu Arg Asp Asn Ser Val Ile

385 390 395 400

Asp Asn Asn Thr Arg Gln Asn Val Arg Thr Arg Val Thr Gln Met Gln

405 410 415

Tyr Tyr Gly Phe His Leu Ser Val Arg Asp Thr Phe Asn Pro Ile Leu
420 425 430
Asn Ala Gly Lys Leu Thr Gln Gln Phe Ile Val Asp Ser Tyr Ser Lys
435 440 445
Met Glu Ala Asn Arg Ile Asn Phe Ile Lys Ala Asn Gln Ser Lys Leu
450 455 460
Arg Val Glu Lys Tyr Ser Gly Leu Met Asp Tyr Leu Lys Ser Arg Ser

465 470 475 480
Glu Asn Asp Asn Val Pro Ile Gly Lys Met Ile Ile Leu Pro Ser Ser
485 490 495
Phe Glu Gly Ser Pro Arg Asn Met Gln Gln Arg Tyr Gln Asp Ala Met
500 505 510
Ala Ile Val Thr Lys Tyr Gly Lys Pro Asp Leu Phe Ile Thr Met Thr
515 520 525
Cys Asn Pro Lys Trp Ala Asp Ile Thr Asn Asn Leu Gln Arg Trp Gln

530 535 540
Lys Val Glu Asn Arg Pro Asp Leu Val Ala Arg Val Phe Asn Ile Lys
545 550 555 560
Leu Asn Ala Leu Leu Asn Asp Ile Cys Lys Phe His Leu Phe Gly Lys
565 570 575
Val Ile Ala Lys Ile His Val Ile Glu Phe Gln Lys Arg Gly Leu Pro
580 585 590
His Ala His Ile Leu Leu Ile Leu Asp Ser Glu Ser Lys Leu Arg Ser

595 600 605
Glu Asp Asp Ile Asp Arg Ile Val Lys Ala Glu Ile Pro Asp Glu Asp
610 615 620
Gln Cys Pro Arg Leu Phe Gln Ile Val Lys Ser Asn Met Val His Gly
625 630 635 640
Pro Cys Gly Ile Gln Asn Pro Asn Ser Pro Cys Met Glu Asn Gly Lys
645 650 655
Cys Ser Lys Gly Tyr Pro Lys Glu Phe Gln Asn Ala Thr Ile Gly Asn

660 665 670

Ile Asp Gly Tyr Pro Lys Tyr Lys Arg Arg Ser Gly Ser Thr Met Ser

675 680 685

Ile Gly Asn Lys Val Val Asp Asn Thr Trp Ile Val Pro Tyr Asn Pro

690 695 700

Tyr Leu Cys Leu Lys Tyr Asn Cys His Ile Asn Val Glu Val Cys Ala

705 710 715 720

Ser Ile Lys Ser Val Lys Tyr Leu Phe Lys Tyr Ile Tyr Lys Gly His

725 730 735

Asp Cys Ala Asn Ile Gln Ile Ser Glu Lys Asn Ile Ile Asn His Asp

740 745 750

Glu Val Gln Asp Phe Ile Asp Ser Arg Tyr Val Ser Ala Pro Glu Ala

755 760 765

Val Trp Arg Leu Phe Ala Met Arg Met His Asp Gln Ser His Ala Ile

770 775 780

Thr Arg Leu Ala Ile His Leu Pro Asn Asp Gln Asn Leu Tyr Phe His

785 790 795 800

Thr Asp Asp Phe Ala Glu Val Leu Asp Arg Ala Lys Arg His Asn Ser

805 810 815

Thr Leu Met Ala Trp Phe Leu Leu Asn Arg Glu Asp Ser Asp Ala Arg

820 825 830

Asn Tyr Tyr Tyr Trp Glu Ile Pro Gln His Tyr Val Phe Asn Asn Ser

835 840 845

Leu Trp Thr Lys Arg Arg Lys Gly Gly Asn Lys Val Leu Gly Arg Leu

850 855 860

Phe Thr Val Ser Phe Arg Glu Pro Glu Arg Tyr Tyr Leu Arg Leu Leu

865 870 875 880

Leu Leu His Val Lys Gly Ala Ile Ser Phe Glu Asp Leu Arg Thr Val

885 890 895

Gly Gly Val Thr Tyr Asp Thr Phe His Glu Ala Ala Lys His Arg Gly

900 905 910

Leu Leu Leu Asp Asp Thr Ile Trp Lys Asp Thr Ile Asp Asp Ala Ile

915 920 925

Ile Leu Asn Met Pro Lys Gln Leu Arg Gln Leu Phe Ala Tyr Ile Cys

930 935 940

Val Phe Gly Cys Pro Ser Ala Ala Asp Lys Leu Trp Asp Glu Asn Lys

945 950 955 960

Ser His Phe Ile Glu Asp Phe Cys Trp Lys Leu His Arg Arg Glu Gly

965 970 975

Ala Cys Val Asn Cys Glu Met His Ala Leu Asn Glu Ile Gln Glu Val

980 985 990

Phe Thr Leu His Gly Met Lys Cys Ser His Phe Lys Leu Pro Asp Tyr

995 1000 1005

Pro Leu Leu Met Asn Ala Asn Thr Cys Asp Gln Leu Tyr Glu Gln

1010 1015 1020

Gln Gln Ala Glu Val Leu Ile Asn Ser Leu Asn Asp Glu Gln Leu

1025 1030 1035

Ala Ala Phe Gln Thr Ile Thr Ser Ala Ile Glu Asp Gln Thr Val

1040 1045 1050

His Pro Lys Cys Phe Phe Leu Asp Gly Pro Gly Gly Ser Gly Lys

1055 1060 1065

Thr Tyr Leu Tyr Lys Val Leu Thr His Tyr Ile Arg Gly Arg Gly

1070 1075 1080

Gly Thr Val Leu Pro Thr Ala Ser Thr Gly Ile Ala Ala Asn Leu

1085 1090 1095

Leu Leu Gly Gly Arg Thr Phe His Ser Gln Tyr Lys Leu Pro Ile

1100 1105 1110

Pro Leu Asn Glu Thr Ser Ile Ser Arg Leu Asp Ile Lys Ser Glu

1115 1120 1125

Val Ala Lys Thr Ile Lys Lys Ala Gln Leu Leu Ile Ile Asp Glu

1130 1135 1140

Cys Thr Met Ala Ser Ser His Ala Ile Asn Ala Ile Asp Arg Leu

1145 1150 1155

Leu Arg Glu Ile Met Asn Leu Asn Val Ala Phe Gly Gly Lys Val

1160 1165 1170

Leu Leu Leu Gly Gly Asp Phe Arg Gln Cys Leu Ser Ile Val Pro

1175 1180 1185

His Ala Met Arg Ser Ala Ile Val Gln Thr Ser Leu Lys Tyr Cys

1190 1195 1200

Asn Val Trp Gly Cys Phe Arg Lys Leu Ser Leu Lys Thr Asn Met

1205 1210 1215

Arg Ser Glu Asp Ser Ala Tyr Ser Glu Trp Leu Val Lys Leu Gly

1220 1225 1230

Asp Gly Lys Leu Asp Ser Ser Phe His Leu Gly Met Asp Ile Ile

1235 1240 1245

Glu Ile Pro His Glu Met Ile Cys Asn Gly Ser Ile Ile Glu Ala

1250 1255 1260

Thr Phe Gly Asn Ser Ile Ser Ile Asp Asn Ile Lys Asn Ile Ser

1265 1270 1275

Lys Arg Ala Ile Leu Cys Pro Lys Asn Glu His Val Gln Lys Leu

1280 1285 1290

Asn Glu Glu Ile Leu Asp Ile Leu Asp Gly Asp Phe His Thr Tyr

1295 1300 1305

Leu Ser Asp Asp Ser Ile Asp Ser Thr Asp Asp Ala Glu Lys Glu

1310 1315 1320

Asn Phe Pro Ile Glu Phe Leu Asn Ser Ile Thr Pro Ser Gly Met

1325 1330 1335

Pro Cys His Lys Leu Lys Leu Lys Val Gly Ala Ile Ile Met Leu

1340 1345 1350

Leu Arg Asn Leu Asn Ser Lys Trp Gly Leu Cys Asn Gly Thr Arg

1355 1360 1365

Phe Ile Ile Lys Arg Leu Arg Pro Asn Ile Ile Glu Ala Glu Val

1370 1375 1380

Leu Thr Gly Ser Ala Glu Gly Glu Val Val Leu Ile Pro Arg Ile

1385 1390 1395

Asp Leu Ser Pro Ser Asp Thr Gly Leu Pro Phe Lys Leu Ile Arg

1400 1405 1410

Arg Gln Phe Pro Val Met Pro Ala Phe Ala Met Thr Ile Asn Lys

1415 1420 1425

Ser Gln Gly Gln Thr Leu Asp Arg Val Gly Ile Phe Leu Pro Glu

1430 1435 1440

Pro Val Phe Ala His Gly Gln Leu Tyr Val Ala Phe Ser Arg Val

1445 1450 1455

Arg Arg Ala Cys Asp Val Lys Val Lys Val Val Asn Thr Ser Ser

1460 1465 1470

Gln Gly Lys Leu Val Lys His Ser Glu Ser Val Phe Thr Leu Asn

1475 1480 1485

Val Val Tyr Arg Glu Ile Leu Glu

1490 1495

<210> 69

<211> 4682

<212> PRT

<213> Artificial Sequence

<220><223> Tol2 transposon - 40-14491

<400> 69

Cys Ala Gly Ala Gly Gly Thr Gly Thr Ala Ala Ala Gly Thr Ala Cys

1 5 10 15

Thr Thr Gly Ala Gly Thr Ala Ala Thr Thr Thr Thr Ala Cys Thr Thr

20 25 30

Gly Ala Thr Thr Ala Cys Thr Gly Thr Ala Cys Thr Thr Ala Ala Gly

35 40 45

Thr Ala Thr Thr Ala Thr Thr Thr Thr Thr Gly Gly Gly Gly Ala Thr

50 55 60

Thr Thr Thr Thr Ala Cys Thr Thr Thr Ala Cys Thr Thr Gly Ala Gly

65 70 75 80

Thr Ala Cys Ala Ala Thr Thr Ala Ala Ala Ala Ala Thr Cys Ala Ala
85 90 95
Thr Ala Cys Thr Thr Thr Thr Ala Cys Thr Thr Thr Thr Ala Cys Thr
100 105 110
Thr Ala Ala Thr Thr Ala Cys Ala Thr Thr Thr Thr Thr Thr Thr Ala
115 120 125
Gly Ala Ala Ala Ala Ala Ala Ala Ala Gly Thr Ala Cys Thr Thr Thr
130 135 140

Thr Thr Ala Cys Thr Cys Cys Thr Thr Ala Cys Ala Ala Thr Thr Thr
145 150 155 160
Thr Ala Thr Thr Thr Ala Cys Ala Gly Thr Cys Ala Ala Ala Ala Ala
165 170 175
Gly Thr Ala Cys Thr Thr Ala Thr Thr Thr Thr Thr Thr Gly Gly Ala
180 185 190
Gly Ala Thr Cys Ala Cys Thr Thr Cys Ala Thr Thr Cys Thr Ala Thr
195 200 205

Thr Thr Thr Cys Cys Cys Thr Thr Gly Cys Thr Ala Thr Thr Ala Cys
210 215 220
Cys Ala Ala Ala Cys Cys Ala Ala Thr Thr Gly Ala Ala Thr Thr Gly
225 230 235 240
Cys Gly Cys Thr Gly Ala Thr Gly Cys Cys Cys Ala Gly Thr Thr Thr
245 250 255
Ala Ala Thr Thr Thr Ala Ala Ala Thr Gly Thr Thr Ala Thr Thr Thr
260 265 270

Ala Thr Thr Cys Thr Gly Cys Cys Thr Ala Thr Gly Ala Ala Ala Ala
275 280 285
Thr Cys Gly Thr Thr Thr Thr Cys Ala Cys Ala Thr Thr Ala Thr Ala
290 295 300
Thr Gly Ala Ala Ala Thr Thr Gly Gly Thr Cys Ala Gly Ala Cys Ala
305 310 315 320

Thr Gly Thr Thr Cys Ala Thr Thr Gly Gly Thr Cys Cys Thr Thr Thr
325 330 335

Gly Gly Ala Ala Gly Thr Gly Ala Cys Gly Thr Cys Ala Thr Gly Thr
340 345 350

Cys Ala Cys Ala Thr Cys Thr Ala Thr Thr Ala Cys Cys Ala Cys Ala
355 360 365

Ala Thr Gly Cys Ala Cys Ala Gly Cys Ala Cys Cys Thr Thr Gly Ala
370 375 380

Cys Cys Thr Gly Gly Ala Ala Ala Thr Thr Ala Gly Gly Gly Ala Ala
385 390 395 400

Ala Thr Thr Ala Thr Ala Ala Cys Ala Gly Thr Cys Ala Ala Thr Cys
405 410 415

Ala Gly Thr Gly Gly Ala Ala Gly Ala Ala Ala Ala Thr Gly Gly Ala
420 425 430

Gly Gly Ala Ala Gly Thr Ala Thr Gly Thr Gly Ala Thr Thr Cys Ala
435 440 445

Thr Cys Ala Gly Cys Ala Gly Cys Thr Gly Cys Gly Ala Gly Cys Ala
450 455 460

Gly Cys Ala Cys Ala Gly Thr Cys Cys Ala Ala Ala Ala Thr Cys Ala
465 470 475 480

Gly Cys Cys Ala Cys Ala Gly Gly Ala Thr Cys Ala Ala Gly Ala Gly
485 490 495

Cys Ala Cys Cys Cys Gly Thr Gly Gly Cys Cys Gly Thr Ala Thr Cys
500 505 510

Thr Thr Cys Gly Cys Gly Ala Ala Thr Thr Cys Thr Thr Thr Thr Cys
515 520 525

Thr Thr Thr Ala Ala Gly Thr Gly Gly Thr Gly Thr Ala Ala Ala Thr
530 535 540

Ala Ala Ala Gly Ala Thr Thr Cys Ala Thr Thr Cys Ala Ala Gly Ala
545 550 555 560

Thr Gly Ala Ala Ala Thr Gly Thr Gly Thr Cys Cys Thr Cys Thr Gly

565 570 575

Thr Cys Thr Cys Cys Cys Gly Cys Thr Thr Ala Ala Thr Ala Ala Ala

580 585 590

Gly Ala Ala Ala Thr Ala Thr Cys Gly Gly Cys Cys Thr Thr Cys Ala

595 600 605

Ala Ala Ala Gly Thr Thr Cys Gly Cys Cys Ala Thr Cys Ala Ala Ala

610 615 620

Cys Cys Thr Ala Ala Gly Gly Ala Ala Gly Cys Ala Thr Ala Thr Thr

625 630 635 640

Gly Ala Gly Gly Thr Ala Ala Gly Thr Ala Cys Ala Thr Thr Ala Ala

645 650 655

Gly Thr Ala Thr Thr Thr Thr Gly Thr Thr Thr Thr Ala Cys Thr Gly

660 665 670

Ala Thr Ala Gly Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr

675 680 685

Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr

690 695 700

Thr Gly Gly Gly Thr Gly Thr Gly Cys Ala Thr Gly Thr Thr Thr Thr

705 710 715 720

Gly Ala Cys Gly Thr Thr Gly Ala Thr Gly Gly Cys Gly Cys Gly Cys

725 730 735

Cys Thr Thr Thr Thr Ala Thr Ala Thr Gly Thr Gly Thr Ala Gly Thr

740 745 750

Ala Gly Gly Cys Cys Thr Ala Thr Thr Thr Thr Cys Ala Cys Thr Ala

755 760 765

Ala Thr Gly Cys Ala Thr Gly Cys Gly Ala Thr Thr Gly Ala Cys Ala

770 775 780

Ala Thr Ala Thr Ala Ala Gly Gly Cys Thr Cys Ala Cys Gly Thr Ala

785 790 795 800

Ala Thr Ala Ala Ala Ala Thr Gly Cys Thr Ala Ala Ala Ala Thr Gly

805 810 815

Cys Ala Thr Thr Thr Gly Thr Ala Ala Thr Thr Gly Gly Thr Ala Ala
820 825 830
Cys Gly Thr Thr Ala Gly Gly Thr Cys Cys Ala Cys Gly Gly Gly Ala
835 840 845

Ala Ala Thr Thr Thr Gly Gly Cys Gly Cys Cys Thr Ala Thr Thr Gly
850 855 860
Cys Ala Gly Cys Thr Thr Thr Gly Ala Ala Thr Ala Ala Thr Cys Ala
865 870 875 880
Thr Thr Ala Thr Cys Ala Thr Thr Cys Cys Gly Thr Gly Cys Thr Cys
885 890 895
Thr Cys Ala Thr Thr Gly Thr Gly Thr Thr Thr Gly Ala Ala Thr Thr
900 905 910

Cys Ala Thr Gly Cys Ala Ala Ala Ala Cys Ala Cys Ala Ala Gly Ala
915 920 925
Ala Ala Ala Cys Cys Ala Ala Gly Cys Gly Ala Gly Ala Ala Ala Thr
930 935 940
Thr Thr Thr Thr Thr Thr Cys Cys Ala Ala Ala Cys Ala Thr Gly Thr
945 950 955 960
Thr Gly Thr Ala Thr Thr Gly Thr Cys Ala Ala Ala Ala Cys Gly Gly
965 970 975

Thr Ala Ala Cys Ala Cys Thr Thr Thr Ala Cys Ala Ala Thr Gly Ala
980 985 990
Gly Gly Thr Thr Gly Ala Thr Thr Ala Gly Thr Thr Cys Ala Thr Gly
995 1000 1005
Thr Ala Thr Thr Ala Ala Cys Thr Ala Ala Cys Ala Thr Thr Ala
1010 1015 1020
Ala Ala Thr Ala Ala Cys Cys Ala Thr Gly Ala Gly Cys Ala Ala
1025 1030 1035

Thr Ala Cys Ala Thr Thr Thr Gly Thr Thr Ala Cys Thr Gly Thr
1040 1045 1050
Ala Thr Cys Thr Gly Thr Thr Ala Ala Thr Cys Thr Thr Thr Gly

1055 1060 1065

Thr Thr Ala Ala Cys Gly Thr Thr Ala Gly Thr Thr Ala Ala Thr

1070 1075 1080

Ala Gly Ala Ala Ala Thr Ala Cys Ala Gly Ala Thr Gly Thr Thr

1085 1090 1095

Cys Ala Thr Thr Gly Thr Thr Thr Gly Thr Thr Cys Ala Thr Gly

1100 1105 1110

Thr Thr Ala Gly Thr Thr Cys Ala Cys Ala Gly Thr Gly Cys Ala

1115 1120 1125

Thr Thr Ala Ala Cys Thr Ala Ala Thr Gly Thr Thr Ala Ala Cys

1130 1135 1140

Ala Ala Gly Ala Thr Ala Thr Ala Ala Ala Gly Thr Ala Thr Thr

1145 1150 1155

Ala Gly Thr Ala Ala Ala Thr Gly Thr Thr Gly Ala Ala Ala Thr

1160 1165 1170

Thr Ala Ala Cys Ala Thr Gly Thr Ala Thr Ala Cys Gly Thr Gly

1175 1180 1185

Cys Ala Gly Thr Thr Cys Ala Thr Thr Ala Thr Thr Ala Gly Thr

1190 1195 1200

Thr Cys Ala Thr Gly Thr Thr Ala Ala Cys Thr Ala Ala Thr Gly

1205 1210 1215

Thr Ala Gly Thr Thr Ala Ala Cys Thr Ala Ala Cys Gly Ala Ala

1220 1225 1230

Cys Cys Thr Thr Ala Thr Thr Gly Thr Ala Ala Ala Ala Gly Thr

1235 1240 1245

Gly Thr Thr Ala Cys Cys Ala Thr Cys Ala Ala Ala Ala Cys Thr

1250 1255 1260

Ala Ala Thr Gly Thr Ala Ala Thr Gly Ala Ala Ala Thr Cys Ala

1265 1270 1275

Ala Thr Thr Cys Ala Cys Cys Cys Thr Gly Thr Cys Ala Thr Gly

1280 1285 1290

Thr Cys Ala Gly Cys Cys Thr Thr Ala Cys Ala Gly Thr Cys Cys
1295 1300 1305
Thr Gly Thr Gly Thr Thr Thr Thr Thr Gly Thr Cys Ala Ala Thr
1310 1315 1320
Ala Thr Ala Ala Thr Cys Ala Gly Ala Ala Ala Thr Ala Ala Ala
1325 1330 1335

Ala Thr Thr Ala Ala Thr Gly Thr Thr Thr Gly Ala Thr Thr Gly
1340 1345 1350
Thr Cys Ala Cys Thr Ala Ala Ala Thr Gly Cys Thr Ala Cys Thr
1355 1360 1365
Gly Thr Ala Thr Thr Thr Cys Thr Ala Ala Ala Ala Thr Cys Ala
1370 1375 1380
Ala Cys Ala Ala Gly Thr Ala Thr Thr Thr Ala Ala Cys Ala Thr
1385 1390 1395

Thr Ala Thr Ala Ala Ala Gly Thr Gly Thr Gly Cys Ala Ala Thr
1400 1405 1410
Thr Gly Gly Cys Thr Gly Cys Ala Ala Ala Thr Gly Thr Cys Ala
1415 1420 1425
Gly Thr Thr Thr Thr Ala Thr Thr Ala Ala Ala Gly Gly Gly Thr
1430 1435 1440
Thr Ala Gly Thr Thr Cys Ala Cys Cys Cys Ala Ala Ala Ala Ala
1445 1450 1455

Thr Gly Ala Ala Ala Ala Thr Ala Ala Thr Gly Thr Cys Ala Thr
1460 1465 1470
Thr Ala Ala Thr Gly Ala Cys Thr Cys Gly Cys Cys Cys Thr Cys
1475 1480 1485
Ala Thr Gly Thr Cys Gly Thr Thr Cys Cys Ala Ala Gly Cys Cys
1490 1495 1500
Cys Gly Thr Ala Ala Gly Ala Cys Cys Thr Cys Cys Gly Thr Thr
1505 1510 1515

Cys Ala Thr Cys Thr Thr Cys Ala Gly Ala Ala Cys Ala Cys Ala

1520 1525 1530

Gly Thr Thr Thr Ala Ala Gly Ala Thr Ala Thr Thr Thr Thr Ala

1535 1540 1545

Gly Ala Thr Thr Thr Ala Gly Thr Cys Cys Gly Ala Gly Ala Gly

1550 1555 1560

Cys Thr Thr Thr Cys Thr Gly Thr Gly Cys Cys Thr Cys Cys Ala

1565 1570 1575

Thr Thr Gly Ala Gly Ala Ala Thr Gly Thr Ala Thr Gly Thr Ala

1580 1585 1590

Cys Gly Gly Thr Ala Thr Ala Cys Thr Gly Thr Cys Cys Ala Thr

1595 1600 1605

Gly Thr Cys Cys Ala Gly Ala Ala Ala Gly Gly Thr Ala Ala Thr

1610 1615 1620

Ala Ala Ala Ala Ala Ala Cys Ala Thr Cys Ala Ala Ala Gly Thr Ala

1625 1630 1635

Gly Thr Cys Cys Ala Thr Gly Thr Gly Ala Cys Ala Thr Cys Ala

1640 1645 1650

Gly Thr Gly Gly Gly Thr Thr Ala Gly Thr Thr Ala Gly Ala Ala

1655 1660 1665

Thr Thr Thr Thr Thr Thr Gly Ala Ala Gly Cys Ala Thr Cys Gly

1670 1675 1680

Ala Ala Thr Ala Cys Ala Thr Thr Thr Thr Gly Gly Thr Cys Cys

1685 1690 1695

Ala Ala Ala Ala Ala Thr Ala Ala Cys Ala Ala Ala Ala Cys Cys

1700 1705 1710

Thr Ala Cys Gly Ala Cys Thr Thr Thr Ala Thr Thr Cys Gly Gly

1715 1720 1725

Cys Ala Thr Thr Gly Thr Ala Thr Thr Cys Thr Cys Thr Thr Cys

1730 1735 1740

Cys Gly Gly Gly Thr Cys Thr Gly Thr Thr Gly Thr Cys Ala Ala

1745 1750 1755

Thr Cys Cys Gly Cys Gly Thr Thr Cys Ala Cys Gly Ala Cys Thr
1760 1765 1770
Thr Cys Gly Cys Ala Gly Thr Gly Ala Cys Gly Cys Thr Ala Cys
1775 1780 1785
Ala Ala Thr Gly Cys Thr Gly Ala Ala Thr Ala Ala Ala Gly Thr
1790 1795 1800
Cys Gly Thr Ala Gly Gly Thr Thr Thr Thr Gly Thr Thr Ala Thr
1805 1810 1815

Thr Thr Thr Thr Gly Gly Ala Cys Cys Ala Ala Ala Ala Thr Gly
1820 1825 1830
Thr Ala Thr Thr Thr Thr Cys Gly Ala Thr Gly Cys Thr Thr Cys
1835 1840 1845
Ala Ala Ala Thr Ala Ala Thr Thr Cys Thr Ala Cys Cys Thr Ala
1850 1855 1860
Ala Cys Cys Cys Ala Cys Thr Gly Ala Thr Gly Thr Cys Ala Cys
1865 1870 1875

Ala Thr Gly Gly Ala Cys Thr Ala Cys Thr Thr Thr Gly Ala Thr
1880 1885 1890
Gly Thr Thr Thr Thr Thr Ala Thr Thr Ala Cys Cys Thr Thr Thr
1895 1900 1905
Cys Thr Gly Gly Ala Cys Ala Thr Gly Gly Ala Cys Ala Gly Thr
1910 1915 1920
Ala Thr Ala Cys Cys Gly Thr Ala Cys Ala Thr Ala Cys Ala Thr
1925 1930 1935

Thr Thr Thr Cys Ala Gly Thr Gly Gly Ala Gly Gly Gly Ala Cys
1940 1945 1950
Ala Gly Ala Ala Ala Gly Cys Thr Cys Thr Cys Gly Gly Ala Cys
1955 1960 1965
Thr Ala Ala Ala Thr Cys Thr Ala Ala Ala Ala Thr Ala Thr Cys
1970 1975 1980
Thr Thr Ala Ala Ala Cys Thr Gly Thr Gly Thr Thr Cys Cys Gly

1985 1990 1995

Ala Ala Gly Ala Thr Gly Ala Ala Cys Gly Gly Ala Gly Gly Thr
2000 2005 2010
Gly Thr Thr Ala Cys Gly Gly Gly Cys Thr Thr Gly Gly Ala Ala
2015 2020 2025
Cys Gly Ala Cys Ala Thr Gly Ala Gly Gly Gly Thr Gly Ala Gly
2030 2035 2040
Thr Cys Ala Thr Thr Ala Ala Thr Gly Ala Cys Ala Thr Cys Thr
2045 2050 2055

Thr Thr Thr Cys Ala Thr Thr Thr Thr Thr Gly Gly Gly Thr Gly
2060 2065 2070
Ala Ala Cys Thr Ala Ala Cys Cys Cys Thr Thr Thr Ala Ala Thr
2075 2080 2085
Gly Cys Thr Gly Thr Ala Ala Thr Cys Ala Gly Ala Gly Ala Gly
2090 2095 2100
Thr Gly Thr Ala Thr Gly Thr Gly Thr Ala Ala Thr Thr Gly Thr
2105 2110 2115

Thr Ala Cys Ala Thr Thr Thr Ala Thr Thr Gly Cys Ala Thr Ala
2120 2125 2130
Cys Ala Ala Thr Ala Thr Ala Ala Ala Thr Ala Thr Thr Thr Ala
2135 2140 2145
Thr Thr Thr Gly Thr Thr Gly Thr Thr Thr Thr Thr Ala Cys Ala
2150 2155 2160
Gly Ala Gly Ala Ala Thr Gly Cys Ala Cys Cys Cys Ala Ala Ala
2165 2170 2175

Thr Thr Ala Cys Cys Thr Cys Ala Ala Ala Ala Ala Cys Thr Ala
2180 2185 2190
Cys Thr Cys Thr Ala Ala Ala Thr Thr Gly Ala Cys Ala Gly Cys
2195 2200 2205
Ala Cys Ala Gly Ala Ala Gly Ala Gly Ala Ala Ala Gly Ala Thr
2210 2215 2220

Cys Gly Gly Gly Ala Cys Cys Thr Cys Cys Ala Cys Cys Cys Ala
2225 2230 2235

Thr Gly Cys Thr Thr Cys Cys Ala Gly Cys Ala Gly Thr Ala Ala
2240 2245 2250
Gly Cys Ala Ala Cys Thr Gly Ala Ala Ala Gly Thr Thr Gly Ala
2255 2260 2265
Cys Thr Cys Ala Gly Thr Thr Thr Thr Cys Cys Cys Ala Gly Thr
2270 2275 2280
Cys Ala Ala Ala Cys Ala Thr Gly Thr Gly Thr Cys Thr Cys Cys
2285 2290 2295

Ala Gly Thr Cys Ala Cys Thr Gly Thr Gly Ala Ala Cys Ala Ala
2300 2305 2310
Ala Gly Cys Thr Ala Thr Ala Thr Thr Ala Ala Gly Gly Thr Ala
2315 2320 2325
Cys Ala Thr Cys Ala Thr Thr Cys Ala Ala Gly Gly Ala Cys Thr
2330 2335 2340
Thr Cys Ala Thr Cys Cys Thr Thr Thr Cys Ala Gly Cys Ala Cys
2345 2350 2355

Thr Gly Thr Thr Gly Ala Thr Cys Thr Gly Cys Cys Ala Thr Cys
2360 2365 2370
Ala Thr Thr Thr Ala Ala Ala Gly Ala Gly Cys Thr Gly Ala Thr
2375 2380 2385
Thr Ala Gly Thr Ala Cys Ala Cys Thr Gly Cys Ala Gly Cys Cys
2390 2395 2400
Thr Gly Gly Cys Ala Thr Thr Thr Cys Thr Gly Thr Cys Ala Thr
2405 2410 2415

Thr Ala Cys Ala Ala Gly Gly Cys Cys Thr Ala Cys Thr Thr Thr
2420 2425 2430
Ala Cys Gly Cys Thr Cys Cys Ala Ala Gly Ala Thr Ala Gly Cys
2435 2440 2445
Thr Gly Ala Ala Gly Cys Thr Gly Cys Thr Cys Thr Gly Ala Thr

2450 2455 2460

Cys Ala Thr Gly Ala Ala Ala Cys Ala Gly Ala Ala Ala Gly Thr

2465 2470 2475

Gly Ala Cys Thr Gly Cys Thr Gly Cys Cys Ala Thr Gly Ala Gly

2480 2485 2490

Thr Gly Ala Ala Gly Thr Thr Gly Ala Ala Thr Gly Gly Ala Thr

2495 2500 2505

Thr Gly Cys Ala Ala Cys Cys Ala Cys Ala Ala Cys Gly Gly Ala

2510 2515 2520

Thr Thr Gly Thr Thr Gly Gly Ala Cys Thr Gly Cys Ala Cys Gly

2525 2530 2535

Thr Ala Gly Ala Ala Ala Gly Thr Cys Ala Thr Thr Cys Ala Thr

2540 2545 2550

Thr Gly Gly Thr Gly Thr Ala Ala Cys Thr Gly Cys Thr Cys Ala

2555 2560 2565

Cys Thr Gly Gly Ala Thr Cys Ala Ala Cys Cys Cys Thr Gly Gly

2570 2575 2580

Ala Ala Gly Thr Cys Thr Thr Gly Ala Ala Ala Gly Ala Cys Ala

2585 2590 2595

Thr Thr Cys Cys Gly Cys Thr Gly Cys Ala Cys Thr Thr Gly Cys

2600 2605 2610

Cys Thr Gly Cys Ala Ala Ala Ala Gly Ala Thr Thr Ala Ala Thr

2615 2620 2625

Gly Gly Gly Cys Thr Cys Thr Cys Ala Thr Ala Cys Thr Thr Thr

2630 2635 2640

Thr Gly Ala Gly Gly Thr Ala Cys Thr Gly Gly Cys Cys Ala Gly

2645 2650 2655

Thr Gly Cys Cys Ala Thr Gly Ala Ala Thr Gly Ala Thr Ala Thr

2660 2665 2670

Cys Cys Ala Cys Thr Cys Ala Gly Ala Gly Thr Ala Thr Gly Ala

2675 2680 2685

Ala Ala Thr Ala Cys Gly Thr Gly Ala Cys Ala Ala Gly Gly Thr
2690 2695 2700
Thr Gly Thr Thr Thr Gly Cys Ala Cys Ala Ala Cys Cys Ala Cys
2705 2710 2715

Ala Gly Ala Cys Ala Gly Thr Gly Gly Thr Thr Cys Cys Ala Ala
2720 2725 2730
Cys Thr Thr Thr Ala Thr Gly Ala Ala Gly Gly Cys Thr Thr Thr
2735 2740 2745
Cys Ala Gly Ala Gly Thr Thr Thr Thr Thr Gly Gly Thr Gly Thr
2750 2755 2760
Gly Gly Ala Ala Ala Ala Cys Ala Ala Thr Gly Ala Thr Ala Thr
2765 2770 2775

Cys Gly Ala Gly Ala Cys Thr Gly Ala Gly Gly Cys Ala Ala Gly
2780 2785 2790
Ala Ala Gly Gly Thr Gly Thr Gly Ala Ala Ala Gly Thr Gly Ala
2795 2800 2805
Thr Gly Ala Cys Ala Cys Thr Gly Ala Thr Thr Cys Thr Gly Ala
2810 2815 2820
Ala Gly Gly Cys Thr Gly Thr Gly Gly Thr Gly Ala Gly Gly Gly
2825 2830 2835

Ala Ala Gly Thr Gly Ala Thr Gly Gly Thr Gly Thr Gly Gly Ala
2840 2845 2850
Ala Thr Thr Cys Cys Ala Ala Gly Ala Thr Gly Cys Cys Thr Cys
2855 2860 2865
Ala Cys Gly Ala Gly Thr Cys Cys Thr Gly Gly Ala Cys Cys Ala
2870 2875 2880
Ala Gly Ala Cys Gly Ala Thr Gly Gly Cys Thr Thr Cys Gly Ala
2885 2890 2895

Ala Thr Thr Cys Cys Ala Gly Cys Thr Ala Cys Cys Ala Ala Ala
2900 2905 2910
Ala Cys Ala Thr Cys Ala Ala Ala Ala Gly Thr Gly Thr Gly Cys

2915 2920 2925

Cys Thr Gly Thr Cys Ala Cys Thr Thr Ala Cys Thr Thr Ala Ala

2930 2935 2940

Cys Cys Thr Ala Gly Thr Cys Thr Cys Ala Ala Gly Cys Gly Thr

2945 2950 2955

Thr Gly Ala Thr Gly Cys Cys Cys Ala Ala Ala Ala Ala Gly Cys

2960 2965 2970

Thr Cys Thr Cys Thr Cys Ala Ala Ala Thr Gly Ala Ala Cys Ala

2975 2980 2985

Cys Thr Ala Cys Ala Ala Gly Ala Ala Ala Cys Thr Cys Thr Ala

2990 2995 3000

Cys Ala Gly Ala Thr Cys Thr Gly Thr Cys Thr Thr Thr Gly Gly

3005 3010 3015

Cys Ala Ala Ala Thr Gly Cys Cys Ala Ala Gly Cys Thr Thr Thr

3020 3025 3030

Ala Thr Gly Gly Ala Ala Thr Ala Ala Ala Ala Gly Cys Ala Gly

3035 3040 3045

Cys Cys Gly Ala Thr Cys Gly Gly Cys Thr Cys Thr Ala Gly Cys

3050 3055 3060

Ala Gly Cys Thr Gly Ala Ala Gly Cys Thr Gly Thr Thr Gly Ala

3065 3070 3075

Ala Thr Cys Ala Gly Ala Ala Ala Gly Cys Cys Gly Gly Cys Thr

3080 3085 3090

Thr Cys Ala Gly Cys Thr Thr Thr Thr Ala Ala Gly Gly Cys Cys

3095 3100 3105

Ala Ala Ala Cys Cys Ala Ala Ala Cys Gly Cys Gly Gly Thr Gly

3110 3115 3120

Gly Ala Ala Thr Thr Cys Ala Ala Cys Thr Thr Thr Thr Ala Thr

3125 3130 3135

Gly Gly Cys Thr Gly Thr Thr Gly Ala Cys Ala Gly Ala Ala Thr

3140 3145 3150

Thr Cys Thr Thr Cys Ala Ala Ala Thr Thr Thr Gly Cys Ala Ala
3155 3160 3165
Ala Gly Ala Ala Gly Cys Ala Gly Gly Ala Gly Ala Ala Gly Gly
3170 3175 3180
Cys Gly Cys Ala Cys Thr Thr Cys Gly Gly Ala Ala Thr Ala Thr
3185 3190 3195

Ala Thr Gly Cys Ala Cys Cys Thr Cys Thr Cys Thr Thr Gly Ala
3200 3205 3210
Gly Gly Thr Thr Cys Cys Ala Ala Thr Gly Thr Ala Ala Gly Thr
3215 3220 3225
Gly Thr Thr Thr Thr Thr Cys Cys Cys Cys Thr Cys Thr Ala Thr
3230 3235 3240
Cys Gly Ala Thr Gly Thr Ala Ala Ala Cys Ala Ala Ala Thr Gly
3245 3250 3255

Thr Gly Gly Gly Thr Thr Gly Thr Thr Thr Thr Thr Gly Thr Thr
3260 3265 3270
Thr Ala Ala Thr Ala Cys Thr Cys Thr Thr Thr Gly Ala Thr Thr
3275 3280 3285
Ala Thr Gly Cys Thr Gly Ala Thr Thr Thr Cys Thr Cys Cys Thr
3290 3295 3300
Gly Thr Ala Gly Gly Thr Thr Thr Ala Ala Thr Cys Cys Ala Gly
3305 3310 3315

Cys Ala Gly Ala Ala Ala Thr Gly Cys Thr Gly Thr Thr Cys Thr
3320 3325 3330
Thr Gly Ala Cys Ala Gly Ala Gly Thr Gly Gly Gly Cys Cys Ala
3335 3340 3345
Ala Cys Ala Cys Ala Ala Thr Gly Cys Gly Thr Cys Cys Ala Gly
3350 3355 3360
Thr Thr Gly Cys Ala Ala Ala Ala Gly Thr Ala Cys Thr Cys Gly
3365 3370 3375

Ala Cys Ala Thr Cys Thr Thr Gly Cys Ala Ala Gly Cys Gly Gly

3380 3385 3390

Ala Ala Ala Cys Gly Ala Ala Thr Ala Cys Ala Cys Ala Gly Cys

3395 3400 3405

Thr Gly Gly Gly Gly Thr Gly Gly Cys Thr Gly Cys Thr Gly Cys

3410 3415 3420

Cys Thr Ala Gly Thr Gly Thr Cys Cys Ala Thr Cys Ala Gly Thr

3425 3430 3435

Thr Ala Ala Gly Cys Thr Thr Gly Ala Ala Ala Cys Thr Thr Cys

3440 3445 3450

Ala Gly Cys Gly Ala Cys Thr Cys Cys Ala Cys Cys Ala Thr Thr

3455 3460 3465

Cys Thr Cys Thr Cys Ala Gly Gly Thr Ala Cys Thr Gly Thr Gly

3470 3475 3480

Ala Cys Cys Cys Ala Cys Thr Thr Gly Thr Gly Gly Ala Thr Gly

3485 3490 3495

Cys Cys Cys Thr Ala Cys Ala Ala Cys Ala Ala Gly Gly Ala Ala

3500 3505 3510

Thr Cys Cys Ala Ala Ala Cys Ala Cys Gly Ala Thr Thr Cys Ala

3515 3520 3525

Ala Gly Cys Ala Thr Ala Thr Gly Thr Thr Thr Gly Ala Ala Gly

3530 3535 3540

Ala Thr Cys Cys Thr Gly Ala Gly Ala Thr Cys Ala Thr Ala Gly

3545 3550 3555

Cys Ala Gly Cys Thr Gly Cys Cys Ala Thr Cys Cys Thr Thr Cys

3560 3565 3570

Thr Cys Cys Cys Thr Ala Ala Ala Thr Thr Thr Cys Gly Gly Ala

3575 3580 3585

Cys Cys Thr Cys Thr Thr Gly Gly Ala Cys Ala Ala Ala Thr Gly

3590 3595 3600

Ala Thr Gly Ala Ala Ala Cys Cys Ala Thr Cys Ala Thr Ala Ala

3605 3610 3615

Ala Ala Cys Gly Ala Gly Gly Thr Ala Ala Ala Thr Gly Ala Ala
3620 3625 3630

Thr Gly Cys Ala Ala Gly Cys Ala Ala Cys Ala Thr Ala Cys Ala
3635 3640 3645

Cys Thr Thr Gly Ala Cys Gly Ala Ala Thr Thr Cys Thr Ala Ala
3650 3655 3660

Thr Cys Thr Gly Gly Gly Cys Ala Ala Cys Cys Thr Thr Thr Gly
3665 3670 3675

Ala Gly Cys Cys Ala Thr Ala Cys Cys Ala Ala Ala Ala Thr Thr
3680 3685 3690

Ala Thr Thr Cys Thr Thr Thr Thr Ala Thr Thr Thr Ala Thr Thr
3695 3700 3705

Thr Ala Thr Thr Thr Thr Thr Gly Cys Ala Cys Thr Thr Thr Thr
3710 3715 3720

Thr Ala Gly Gly Ala Ala Thr Gly Thr Thr Ala Thr Ala Thr Cys
3725 3730 3735

Cys Cys Ala Thr Cys Thr Thr Thr Gly Gly Cys Thr Gly Thr Gly
3740 3745 3750

Ala Thr Cys Thr Cys Ala Ala Thr Ala Thr Gly Ala Ala Thr Ala
3755 3760 3765

Thr Thr Gly Ala Thr Gly Thr Ala Ala Ala Gly Thr Ala Thr Thr
3770 3775 3780

Cys Thr Thr Gly Cys Ala Gly Cys Ala Gly Gly Thr Thr Gly Thr
3785 3790 3795

Ala Gly Thr Thr Ala Thr Cys Cys Cys Thr Cys Ala Gly Thr Gly
3800 3805 3810

Thr Thr Thr Cys Thr Thr Gly Ala Ala Ala Cys Cys Ala Ala Ala
3815 3820 3825

Cys Thr Cys Ala Thr Ala Thr Gly Thr Ala Thr Cys Ala Thr Ala
3830 3835 3840

Thr Gly Thr Gly Gly Thr Thr Thr Gly Gly Ala Ala Ala Thr Gly

3845 3850 3855

Cys Ala Gly Thr Thr Ala Gly Ala Thr Thr Thr Thr Ala Thr Gly
3860 3865 3870
Cys Thr Ala Ala Ala Ala Thr Ala Ala Gly Gly Gly Ala Thr Thr
3875 3880 3885
Thr Gly Cys Ala Thr Gly Ala Thr Thr Thr Thr Thr Ala Gly Ala Thr
3890 3895 3900
Gly Thr Ala Gly Ala Thr Gly Ala Cys Thr Gly Cys Ala Cys Gly
3905 3910 3915

Thr Ala Ala Ala Thr Gly Thr Ala Gly Thr Thr Ala Ala Thr Gly
3920 3925 3930
Ala Cys Ala Ala Ala Ala Thr Cys Cys Ala Thr Ala Ala Ala Ala
3935 3940 3945
Thr Thr Thr Gly Thr Thr Cys Cys Cys Ala Gly Thr Cys Ala Gly
3950 3955 3960
Ala Ala Gly Cys Cys Cys Cys Thr Cys Ala Ala Cys Cys Ala Ala
3965 3970 3975

Ala Cys Thr Thr Thr Thr Cys Thr Thr Thr Gly Thr Gly Thr Cys
3980 3985 3990
Thr Gly Cys Thr Cys Ala Cys Thr Gly Thr Gly Cys Thr Thr Gly
3995 4000 4005
Thr Ala Gly Gly Cys Ala Thr Gly Gly Ala Cys Thr Ala Cys Ala
4010 4015 4020
Thr Cys Ala Gly Ala Gly Thr Gly Cys Ala Thr Cys Thr Gly Gly
4025 4030 4035

Ala Gly Cys Cys Thr Thr Thr Gly Gly Ala Cys Cys Ala Cys Ala
4040 4045 4050
Ala Gly Ala Ala Gly Gly Ala Ala Thr Thr Gly Gly Cys Cys Ala
4055 4060 4065
Ala Cys Ala Gly Thr Thr Cys Ala Thr Cys Thr Gly Ala Thr Gly
4070 4075 4080

Ala Thr Gly Ala Ala Gly Ala Thr Thr Thr Thr Thr Thr Cys Gly
4085 4090 4095

Cys Thr Thr Cys Thr Thr Thr Gly Ala Ala Ala Cys Cys Gly Ala
4100 4105 4110
Cys Ala Ala Cys Ala Cys Ala Thr Gly Ala Ala Gly Cys Cys Ala
4115 4120 4125
Gly Cys Ala Ala Ala Gly Ala Gly Thr Thr Gly Gly Ala Thr Gly
4130 4135 4140
Gly Ala Thr Ala Thr Cys Thr Gly Gly Cys Cys Thr Gly Thr Gly
4145 4150 4155

Thr Thr Thr Cys Ala Gly Ala Cys Ala Cys Cys Ala Gly Gly Gly
4160 4165 4170
Ala Gly Thr Cys Thr Cys Thr Gly Cys Thr Cys Ala Cys Gly Thr
4175 4180 4185
Thr Thr Cys Cys Thr Gly Cys Thr Ala Thr Thr Thr Gly Cys Ala
4190 4195 4200
Gly Cys Cys Thr Cys Thr Cys Thr Ala Thr Cys Ala Ala Gly Ala
4205 4210 4215

Cys Thr Ala Ala Thr Ala Cys Ala Cys Cys Thr Cys Thr Thr Cys
4220 4225 4230
Cys Cys Gly Cys Ala Thr Cys Gly Gly Cys Thr Gly Cys Cys Thr
4235 4240 4245
Gly Thr Gly Ala Gly Ala Gly Gly Cys Thr Thr Thr Thr Cys Ala
4250 4255 4260
Gly Cys Ala Cys Thr Gly Cys Ala Gly Gly Ala Thr Thr Gly Cys
4265 4270 4275

Thr Thr Thr Thr Cys Ala Gly Cys Cys Cys Cys Ala Ala Ala Ala
4280 4285 4290
Gly Ala Gly Cys Thr Ala Gly Gly Cys Thr Thr Gly Ala Cys Ala
4295 4300 4305
Cys Thr Ala Ala Cys Ala Ala Thr Thr Thr Thr Gly Ala Gly Ala

4310 4315 4320

Ala Thr Cys Ala Gly Cys Thr Thr Cys Thr Ala Cys Thr Gly Ala

4325 4330 4335

Ala Gly Thr Thr Ala Ala Ala Thr Cys Thr Gly Ala Gly Gly Thr

4340 4345 4350

Thr Thr Thr Ala Cys Ala Ala Cys Thr Thr Thr Gly Ala Gly Thr

4355 4360 4365

Ala Gly Cys Gly Thr Gly Thr Ala Cys Thr Gly Gly Cys Ala Thr

4370 4375 4380

Thr Ala Gly Ala Thr Thr Gly Thr Cys Thr Gly Thr Cys Thr Thr

4385 4390 4395

Ala Thr Ala Gly Thr Thr Thr Gly Ala Thr Ala Ala Thr Thr Ala

4400 4405 4410

Ala Ala Thr Ala Cys Ala Ala Ala Cys Ala Gly Thr Thr Cys Thr

4415 4420 4425

Ala Ala Ala Gly Cys Ala Gly Gly Ala Thr Ala Ala Ala Ala Cys

4430 4435 4440

Cys Thr Thr Gly Thr Ala Thr Gly Cys Ala Thr Thr Thr Cys Ala

4445 4450 4455

Thr Thr Thr Ala Ala Thr Gly Thr Thr Thr Thr Thr Thr Gly Ala

4460 4465 4470

Gly Ala Thr Thr Ala Ala Ala Ala Gly Cys Thr Thr Ala Ala Ala

4475 4480 4485

Cys Ala Ala Gly Ala Ala Thr Cys Thr Cys Thr Ala Gly Thr Thr

4490 4495 4500

Thr Thr Cys Thr Thr Thr Cys Thr Thr Gly Cys Thr Thr Thr Thr

4505 4510 4515

Ala Cys Thr Thr Thr Thr Ala Cys Thr Thr Cys Cys Thr Thr Ala

4520 4525 4530

Ala Thr Ala Cys Thr Cys Ala Ala Gly Thr Ala Cys Ala Ala Thr

4535 4540 4545

Thr Thr Thr Ala Ala Thr Gly Gly Ala Gly Thr Ala Cys Thr Thr
4550 4555 4560
Thr Thr Thr Thr Ala Cys Thr Thr Thr Thr Ala Cys Thr Cys Ala
4565 4570 4575

Ala Gly Thr Ala Ala Gly Ala Thr Thr Cys Thr Ala Gly Cys Cys
4580 4585 4590
Ala Gly Ala Thr Ala Cys Thr Thr Thr Thr Ala Cys Thr Thr Thr
4595 4600 4605
Thr Ala Ala Thr Thr Gly Ala Gly Thr Ala Ala Ala Ala Thr Thr
4610 4615 4620
Thr Thr Cys Cys Cys Thr Ala Ala Gly Thr Ala Cys Thr Thr Gly
4625 4630 4635

Thr Ala Cys Thr Thr Thr Cys Ala Cys Thr Thr Gly Ala Gly Thr
4640 4645 4650
Ala Ala Ala Ala Thr Thr Thr Thr Thr Gly Ala Gly Thr Ala Cys
4655 4660 4665
Thr Thr Thr Thr Thr Ala Cys Ala Cys Cys Thr Cys Thr Gly
4670 4675 4680

<210> 70
<211> 649
<212> PRT
<213> Artificial Sequence
<220><223> Tol2 transposase - 40-14490
<400> 70

Met Glu Glu Val Cys Asp Ser Ser Ala Ala Ala Ser Ser Thr Val Gln
1 5 10 15
Asn Gln Pro Gln Asp Gln Glu His Pro Trp Pro Tyr Leu Arg Glu Phe
20 25 30
Phe Ser Leu Ser Gly Val Asn Lys Asp Ser Phe Lys Met Lys Cys Val
35 40 45
Leu Cys Leu Pro Leu Asn Lys Glu Ile Ser Ala Phe Lys Ser Ser Pro
50 55 60

Ser Asn Leu Arg Lys His Ile Glu Arg Met His Pro Asn Tyr Leu Lys

65 70 75 80

Asn Tyr Ser Lys Leu Thr Ala Gln Lys Arg Lys Ile Gly Thr Ser Thr

85 90 95

His Ala Ser Ser Ser Lys Gln Leu Lys Val Asp Ser Val Phe Pro Val

100 105 110

Lys His Val Ser Pro Val Thr Val Asn Lys Ala Ile Leu Arg Tyr Ile

115 120 125

Ile Gln Gly Leu His Pro Phe Ser Thr Val Asp Leu Pro Ser Phe Lys

130 135 140

Glu Leu Ile Ser Thr Leu Gln Pro Gly Ile Ser Val Ile Thr Arg Pro

145 150 155 160

Thr Leu Arg Ser Lys Ile Ala Glu Ala Ala Leu Ile Met Lys Gln Lys

165 170 175

Val Thr Ala Ala Met Ser Glu Val Glu Trp Ile Ala Thr Thr Thr Asp

180 185 190

Cys Trp Thr Ala Arg Arg Lys Ser Phe Ile Gly Val Thr Ala His Trp

195 200 205

Ile Asn Pro Gly Ser Leu Glu Arg His Ser Ala Ala Leu Ala Cys Lys

210 215 220

Arg Leu Met Gly Ser His Thr Phe Glu Val Leu Ala Ser Ala Met Asn

225 230 235 240

Asp Ile His Ser Glu Tyr Glu Ile Arg Asp Lys Val Val Cys Thr Thr

245 250 255

Thr Asp Ser Gly Ser Asn Phe Met Lys Ala Phe Arg Val Phe Gly Val

260 265 270

Glu Asn Asn Asp Ile Glu Thr Glu Ala Arg Arg Cys Glu Ser Asp Asp

275 280 285

Thr Asp Ser Glu Gly Cys Gly Glu Gly Ser Asp Gly Val Glu Phe Gln

290 295 300

Asp Ala Ser Arg Val Leu Asp Gln Asp Asp Gly Phe Glu Phe Gln Leu

305 310 315 320

Pro Lys His Gln Lys Cys Ala Cys His Leu Leu Asn Leu Val Ser Ser

325 330 335

Val Asp Ala Gln Lys Ala Leu Ser Asn Glu His Tyr Lys Lys Leu Tyr

340 345 350

Arg Ser Val Phe Gly Lys Cys Gln Ala Leu Trp Asn Lys Ser Ser Arg

355 360 365

Ser Ala Leu Ala Ala Glu Ala Val Glu Ser Glu Ser Arg Leu Gln Leu

370 375 380

Leu Arg Pro Asn Gln Thr Arg Trp Asn Ser Thr Phe Met Ala Val Asp

385 390 395 400

Arg Ile Leu Gln Ile Cys Lys Glu Ala Gly Glu Gly Ala Leu Arg Asn

405 410 415

Ile Cys Thr Ser Leu Glu Val Pro Met Phe Asn Pro Ala Glu Met Leu

420 425 430

Phe Leu Thr Glu Trp Ala Asn Thr Met Arg Pro Val Ala Lys Val Leu

435 440 445

Asp Ile Leu Gln Ala Glu Thr Asn Thr Gln Leu Gly Trp Leu Leu Pro

450 455 460

Ser Val His Gln Leu Ser Leu Lys Leu Gln Arg Leu His His Ser Leu

465 470 475 480

Arg Tyr Cys Asp Pro Leu Val Asp Ala Leu Gln Gln Gly Ile Gln Thr

485 490 495

Arg Phe Lys His Met Phe Glu Asp Pro Glu Ile Ile Ala Ala Ala Ile

500 505 510

Leu Leu Pro Lys Phe Arg Thr Ser Trp Thr Asn Asp Glu Thr Ile Ile

515 520 525

Lys Arg Gly Met Asp Tyr Ile Arg Val His Leu Glu Pro Leu Asp His

530 535 540

Lys Lys Glu Leu Ala Asn Ser Ser Ser Asp Asp Glu Asp Phe Phe Ala

545 550 555 560

Ser Leu Lys Pro Thr Thr His Glu Ala Ser Lys Glu Leu Asp Gly Tyr
565 570 575

Leu Ala Cys Val Ser Asp Thr Arg Glu Ser Leu Leu Thr Phe Pro Ala

580 585 590

Ile Cys Ser Leu Ser Ile Lys Thr Asn Thr Pro Leu Pro Ala Ser Ala
595 600 605

Ala Cys Glu Arg Leu Phe Ser Thr Ala Gly Leu Leu Phe Ser Pro Lys
610 615 620

Arg Ala Arg Leu Asp Thr Asn Asn Phe Glu Asn Gln Leu Leu Leu Lys
625 630 635 640

Leu Asn Leu Arg Phe Tyr Asn Phe Glu
645

<210> 71
<211> 636
<212> PRT
<213> Artificial Sequence
<220><223> TcBuster transposase - 40-18214
<400> 71

Met Met Leu Asn Trp Leu Lys Ser Gly Lys Leu Glu Ser Gln Ser Gln
1 5 10 15

Glu Gln Ser Ser Cys Tyr Leu Glu Asn Ser Asn Cys Leu Pro Pro Thr
20 25 30

Leu Asp Ser Thr Asp Ile Ile Gly Glu Glu Asn Lys Ala Gly Thr Thr
35 40 45

Ser Arg Lys Lys Arg Lys Tyr Asp Glu Asp Tyr Leu Asn Phe Gly Phe
50 55 60

Thr Trp Thr Gly Asp Lys Asp Glu Pro Asn Gly Leu Cys Val Ile Cys
65 70 75 80

Glu Gln Val Val Asn Asn Ser Ser Leu Asn Pro Ala Lys Leu Lys Arg
85 90 95

His Leu Asp Thr Lys His Pro Thr Leu Lys Gly Lys Ser Glu Tyr Phe

100 105 110

Lys Arg Lys Cys Asn Glu Leu Asn Gln Lys Lys His Thr Phe Glu Arg
115 120 125
Tyr Val Arg Asp Asp Asn Lys Asn Leu Leu Lys Ala Ser Tyr Leu Val
130 135 140
Ser Leu Arg Ile Ala Lys Gln Gly Glu Ala Tyr Thr Ile Ala Glu Lys
145 150 155 160
Leu Ile Lys Pro Cys Thr Lys Asp Leu Thr Thr Cys Val Phe Gly Glu
165 170 175

Lys Phe Ala Ser Lys Val Asp Leu Val Pro Leu Ser Asp Thr Thr Ile
180 185 190
Ser Arg Arg Ile Glu Asp Met Ser Tyr Phe Cys Glu Ala Val Leu Val
195 200 205
Asn Arg Leu Glu Asn Ala Lys Cys Gly Phe Thr Leu Gln Met Asp Glu
210 215 220
Ser Thr Asp Val Ala Gly Leu Ala Ile Leu Leu Val Phe Val Arg Tyr
225 230 235 240

Ile His Glu Ser Ser Phe Glu Glu Asp Met Leu Phe Cys Lys Ala Leu
245 250 255
Pro Thr Gln Thr Thr Gly Glu Glu Ile Phe Asn Leu Leu Asn Ala Tyr
260 265 270
Phe Glu Lys His Ser Ile Pro Trp Asn Leu Cys Tyr His Ile Cys Thr
275 280 285
Asp Gly Ala Lys Ala Met Val Gly Val Ile Lys Gly Val Ile Ala Arg
290 295 300

Ile Lys Lys Leu Val Pro Asp Ile Lys Ala Ser His Cys Cys Leu His
305 310 315 320
Arg His Ala Leu Ala Val Lys Arg Ile Pro Asn Ala Leu His Glu Val
325 330 335
Leu Asn Asp Ala Val Lys Met Ile Asn Phe Ile Lys Ser Arg Pro Leu
340 345 350

Asn Ala Arg Val Phe Ala Leu Leu Cys Asp Asp Leu Gly Ser Leu His
355 360 365

Lys Asn Leu Leu Leu His Thr Glu Val Arg Trp Leu Ser Arg Gly Lys
370 375 380
Val Leu Thr Arg Phe Trp Glu Leu Arg Asp Glu Ile Arg Ile Phe Phe
385 390 395 400
Asn Glu Arg Glu Phe Ala Gly Lys Leu Asn Asp Thr Ser Trp Leu Gln
405 410 415
Asn Leu Ala Tyr Ile Ala Asp Ile Phe Ser Tyr Leu Asn Glu Val Asn
420 425 430

Leu Ser Leu Gln Gly Pro Asn Ser Thr Ile Phe Lys Val Asn Ser Arg
435 440 445
Ile Asn Ser Ile Lys Ser Lys Leu Lys Leu Trp Glu Glu Cys Ile Thr
450 455 460
Lys Asn Asn Thr Glu Cys Phe Ala Asn Leu Asn Asp Phe Leu Glu Thr
465 470 475 480
Ser Asn Thr Ala Leu Asp Pro Asn Leu Lys Ser Asn Ile Leu Glu His
485 490 495

Leu Asn Gly Leu Lys Asn Thr Phe Leu Glu Tyr Phe Pro Pro Thr Cys
500 505 510
Asn Asn Ile Ser Trp Val Glu Asn Pro Phe Asn Glu Cys Gly Asn Val
515 520 525
Asp Thr Leu Pro Ile Lys Glu Arg Glu Gln Leu Ile Asp Ile Arg Thr
530 535 540
Asp Thr Thr Leu Lys Ser Ser Phe Val Pro Asp Gly Ile Gly Pro Phe
545 550 555 560

Trp Ile Lys Leu Met Asp Glu Phe Pro Glu Ile Ser Lys Arg Ala Val
565 570 575
Lys Glu Leu Met Pro Phe Val Thr Thr Tyr Leu Cys Glu Lys Ser Phe
580 585 590
Ser Val Tyr Val Ala Thr Lys Thr Lys Tyr Arg Asn Arg Leu Asp Ala

595 600 605

Glu Asp Asp Met Arg Leu Gln Leu Thr Thr Ile His Pro Asp Ile Asp

610 615 620

Asn Leu Cys Asn Asn Lys Gln Ala Gln Lys Ser His

625 630 635

<210> 72

<211> 1911

<212> PRT

<213> Artificial Sequence

<220><223> TcBuster transposase - 40-18215

<400> 72

Ala Thr Gly Ala Thr Gly Thr Thr Gly Ala Ala Thr Thr Gly Gly Cys

1 5 10 15

Thr Gly Ala Ala Ala Ala Gly Thr Gly Gly Ala Ala Ala Gly Cys Thr

20 25 30

Thr Gly Ala Ala Ala Gly Thr Cys Ala Ala Thr Cys Ala Cys Ala Gly

35 40 45

Gly Ala Ala Cys Ala Gly Ala Gly Thr Thr Cys Cys Thr Gly Cys Thr

50 55 60

Ala Cys Cys Thr Thr Gly Ala Gly Ala Ala Cys Thr Cys Thr Ala Ala

65 70 75 80

Cys Thr Gly Cys Cys Thr Gly Cys Cys Ala Cys Cys Ala Ala Cys Gly

85 90 95

Cys Thr Cys Gly Ala Thr Thr Cys Thr Ala Cys Ala Gly Ala Thr Ala

100 105 110

Thr Thr Ala Thr Cys Gly Gly Thr Gly Ala Ala Gly Ala Gly Ala Ala

115 120 125

Cys Ala Ala Ala Gly Cys Thr Gly Gly Thr Ala Cys Cys Ala Cys Cys

130 135 140

Thr Cys Thr Cys Gly Cys Ala Ala Gly Ala Ala Gly Cys Gly Gly Ala

145 150 155 160

Ala Ala Thr Ala Thr Gly Ala Cys Gly Ala Gly Gly Ala Cys Thr Ala

165 170 175

Thr Cys Thr Gly Ala Ala Cys Thr Thr Cys Gly Gly Thr Thr Thr Thr

180 185 190

Ala Cys Ala Thr Gly Gly Ala Cys Thr Gly Gly Cys Gly Ala Cys Ala

195 200 205

Ala Gly Gly Ala Thr Gly Ala Gly Cys Cys Cys Ala Ala Cys Gly Gly

210 215 220

Ala Cys Thr Thr Thr Gly Thr Gly Thr Gly Ala Thr Thr Thr Gly Cys

225 230 235 240

Gly Ala Gly Cys Ala Gly Gly Thr Ala Gly Thr Cys Ala Ala Cys Ala

245 250 255

Ala Thr Thr Cys Cys Thr Cys Ala Cys Thr Thr Ala Ala Cys Cys Cys

260 265 270

Gly Gly Cys Cys Ala Ala Ala Cys Thr Gly Ala Ala Ala Cys Gly Cys

275 280 285

Cys Ala Thr Thr Thr Gly Gly Ala Cys Ala Cys Ala Ala Ala Gly Cys

290 295 300

Ala Thr Cys Cys Gly Ala Cys Gly Cys Thr Thr Ala Ala Ala Gly Gly

305 310 315 320

Cys Ala Ala Gly Ala Gly Cys Gly Ala Ala Thr Ala Cys Thr Thr Cys

325 330 335

Ala Ala Ala Ala Gly Ala Ala Ala Ala Thr Gly Thr Ala Ala Cys Gly

340 345 350

Ala Gly Cys Thr Cys Ala Ala Thr Cys Ala Ala Ala Ala Gly Ala Ala

355 360 365

Gly Cys Ala Thr Ala Cys Thr Thr Thr Thr Gly Ala Gly Cys Gly Ala

370 375 380

Thr Ala Cys Gly Thr Ala Ala Gly Gly Gly Ala Cys Gly Ala Thr Ala

385 390 395 400

Ala Cys Ala Ala Gly Ala Ala Cys Cys Thr Cys Cys Thr Gly Ala Ala

405 410 415

Ala Gly Cys Thr Thr Cys Thr Thr Ala Thr Cys Thr Cys Gly Thr Cys

420 425 430

Ala Gly Thr Thr Thr Gly Ala Gly Ala Ala Thr Ala Gly Cys Thr Ala

435 440 445

Ala Ala Cys Ala Gly Gly Gly Cys Gly Ala Gly Gly Cys Ala Thr Ala

450 455 460

Thr Ala Cys Cys Ala Thr Ala Gly Cys Gly Gly Ala Gly Ala Ala Gly

465 470 475 480

Thr Thr Gly Ala Thr Cys Ala Ala Gly Cys Cys Thr Thr Gly Cys Ala

485 490 495

Cys Cys Ala Ala Gly Gly Ala Thr Cys Thr Gly Ala Cys Ala Ala Cys

500 505 510

Thr Thr Gly Cys Gly Thr Ala Thr Thr Thr Gly Gly Ala Gly Ala Ala

515 520 525

Ala Ala Ala Thr Thr Cys Gly Cys Gly Ala Gly Cys Ala Ala Ala Gly

530 535 540

Thr Thr Gly Ala Thr Cys Thr Cys Gly Thr Cys Cys Cys Cys Cys Thr

545 550 555 560

Gly Thr Cys Cys Gly Ala Cys Ala Cys Gly Ala Cys Thr Ala Thr Thr

565 570 575

Thr Cys Ala Ala Gly Gly Cys Gly Ala Ala Thr Cys Gly Ala Ala Gly

580 585 590

Ala Cys Ala Thr Gly Ala Gly Thr Thr Ala Cys Thr Thr Cys Thr Gly

595 600 605

Thr Gly Ala Ala Gly Cys Cys Gly Thr Gly Cys Thr Gly Gly Thr Gly

610 615 620

Ala Ala Cys Ala Gly Gly Thr Thr Gly Ala Ala Ala Ala Ala Thr Gly

625 630 635 640

Cys Thr Ala Ala Ala Thr Gly Thr Gly Gly Gly Thr Thr Thr Ala Cys

645 650 655

Gly Cys Thr Gly Cys Ala Gly Ala Thr Gly Gly Ala Cys Gly Ala Gly

660 665 670

Thr Cys Ala Ala Cys Ala Gly Ala Thr Gly Thr Thr Gly Cys Cys Gly

675 680 685

Gly Thr Cys Thr Thr Gly Cys Ala Ala Thr Cys Cys Thr Gly Cys Thr

690 695 700

Thr Gly Thr Gly Thr Thr Thr Gly Thr Thr Ala Gly Gly Thr Ala Cys

705 710 715 720

Ala Thr Ala Cys Ala Thr Gly Ala Ala Ala Gly Cys Thr Cys Thr Thr

725 730 735

Thr Thr Gly Ala Gly Gly Ala Gly Gly Ala Thr Ala Thr Gly Thr Thr

740 745 750

Gly Thr Thr Cys Thr Gly Cys Ala Ala Ala Gly Cys Ala Cys Thr Thr

755 760 765

Cys Cys Cys Ala Cys Thr Cys Ala Gly Ala Cys Gly Ala Cys Ala Gly

770 775 780

Gly Gly Gly Ala Gly Gly Ala Gly Ala Thr Thr Thr Thr Cys Ala Ala

785 790 795 800

Thr Cys Thr Thr Cys Thr Cys Ala Ala Thr Gly Cys Cys Thr Ala Thr

805 810 815

Thr Thr Cys Gly Ala Ala Ala Ala Gly Cys Ala Cys Thr Cys Cys Ala

820 825 830

Thr Cys Cys Cys Ala Thr Gly Gly Ala Ala Thr Cys Thr Gly Thr Gly

835 840 845

Thr Thr Ala Cys Cys Ala Cys Ala Thr Thr Thr Gly Cys Ala Cys Ala

850 855 860

Gly Ala Cys Gly Gly Thr Gly Cys Cys Ala Ala Gly Gly Cys Ala Ala

865 870 875 880

Thr Gly Gly Thr Ala Gly Gly Ala Gly Thr Thr Ala Thr Thr Ala Ala

885 890 895

Ala Gly Gly Ala Gly Thr Cys Ala Thr Ala Gly Cys Gly Ala Gly Ala

900 905 910

Ala Thr Ala Ala Ala Ala Ala Ala Ala Cys Thr Cys Gly Thr Cys Cys

915 920 925

Cys Thr Gly Ala Thr Ala Thr Ala Ala Ala Ala Gly Cys Thr Ala Gly

930 935 940

Cys Cys Ala Cys Thr Gly Thr Thr Gly Cys Cys Thr Gly Cys Ala Thr

945 950 955 960

Cys Gly Cys Cys Ala Cys Gly Cys Thr Thr Thr Gly Gly Cys Thr Gly

965 970 975

Thr Ala Ala Ala Gly Cys Gly Ala Ala Thr Ala Cys Cys Gly Ala Ala

980 985 990

Thr Gly Cys Ala Thr Thr Gly Cys Ala Cys Gly Ala Gly Gly Thr Gly

995 1000 1005

Cys Thr Cys Ala Ala Thr Gly Ala Cys Gly Cys Thr Gly Thr Thr

1010 1015 1020

Ala Ala Ala Ala Thr Gly Ala Thr Cys Ala Ala Cys Thr Thr Cys

1025 1030 1035

Ala Thr Cys Ala Ala Gly Thr Cys Thr Cys Gly Gly Cys Cys Gly

1040 1045 1050

Thr Thr Gly Ala Ala Thr Gly Cys Gly Cys Gly Cys Gly Thr Cys

1055 1060 1065

Thr Thr Cys Gly Cys Thr Thr Thr Gly Cys Thr Gly Thr Gly Thr

1070 1075 1080

Gly Ala Cys Gly Ala Thr Thr Thr Gly Gly Gly Gly Ala Gly Cys

1085 1090 1095

Cys Thr Gly Cys Ala Thr Ala Ala Ala Ala Ala Thr Cys Thr Thr

1100 1105 1110

Cys Thr Thr Cys Thr Thr Cys Ala Thr Ala Cys Cys Gly Ala Ala

1115 1120 1125

Gly Thr Gly Ala Gly Gly Thr Gly Gly Cys Thr Gly Thr Cys Thr

1130 1135 1140

Ala Gly Ala Gly Gly Ala Ala Ala Gly Gly Thr Gly Cys Thr Gly

1145 1150 1155

Ala Cys Cys Cys Gly Ala Thr Thr Thr Thr Gly Gly Gly Ala Ala

1160 1165 1170

Cys Thr Gly Ala Gly Ala Gly Ala Thr Gly Ala Ala Ala Thr Thr

1175 1180 1185

Ala Gly Ala Ala Thr Thr Thr Thr Cys Thr Thr Cys Ala Ala Cys

1190 1195 1200

Gly Ala Ala Ala Gly Gly Gly Ala Ala Thr Thr Thr Gly Cys Cys

1205 1210 1215

Gly Gly Gly Ala Ala Ala Thr Thr Gly Ala Ala Cys Gly Ala Cys

1220 1225 1230

Ala Cys Cys Ala Gly Thr Thr Gly Gly Thr Thr Gly Cys Ala Ala

1235 1240 1245

Ala Ala Thr Thr Thr Gly Gly Cys Ala Thr Ala Thr Ala Thr Ala

1250 1255 1260

Gly Cys Thr Gly Ala Cys Ala Thr Ala Thr Thr Cys Ala Gly Thr

1265 1270 1275

Thr Ala Thr Cys Thr Gly Ala Ala Thr Gly Ala Ala Gly Thr Thr

1280 1285 1290

Ala Ala Thr Cys Thr Thr Thr Cys Cys Cys Thr Gly Cys Ala Ala

1295 1300 1305

Gly Gly Gly Cys Cys Gly Ala Ala Thr Ala Gly Cys Ala Cys Ala

1310 1315 1320

Ala Thr Cys Thr Thr Cys Ala Ala Gly Gly Thr Ala Ala Ala Thr

1325 1330 1335

Ala Gly Cys Cys Gly Cys Ala Thr Thr Ala Ala Cys Ala Gly Thr

1340 1345 1350

Ala Thr Thr Ala Ala Ala Thr Cys Ala Ala Ala Gly Thr Thr Gly

1355 1360 1365

Ala Ala Gly Thr Thr Gly Thr Gly Gly Gly Ala Ala Gly Ala Gly

1370 1375 1380

Thr Gly Thr Ala Thr Ala Ala Cys Gly Ala Ala Ala Ala Ala Thr

1385 1390 1395

Ala Ala Cys Ala Cys Thr Gly Ala Gly Thr Gly Thr Thr Thr Thr

1400 1405 1410

Gly Cys Gly Ala Ala Cys Cys Thr Cys Ala Ala Cys Gly Ala Thr

1415 1420 1425

Thr Thr Thr Thr Thr Gly Gly Ala Ala Ala Cys Thr Thr Cys Ala

1430 1435 1440

Ala Ala Cys Ala Cys Thr Gly Cys Gly Thr Thr Gly Gly Ala Thr

1445 1450 1455

Cys Cys Ala Ala Ala Cys Cys Thr Gly Ala Ala Gly Thr Cys Thr

1460 1465 1470

Ala Ala Thr Ala Thr Thr Thr Thr Gly Gly Ala Ala Cys Ala Thr

1475 1480 1485

Cys Thr Cys Ala Ala Cys Gly Gly Thr Cys Thr Thr Ala Ala Gly

1490 1495 1500

Ala Ala Cys Ala Cys Cys Thr Thr Thr Cys Thr Gly Gly Ala Gly

1505 1510 1515

Thr Ala Thr Thr Thr Thr Cys Cys Ala Cys Cys Thr Ala Cys Gly

1520 1525 1530

Thr Gly Thr Ala Ala Thr Ala Ala Thr Ala Thr Cys Thr Cys Cys

1535 1540 1545

Thr Gly Gly Gly Thr Gly Gly Ala Gly Ala Ala Thr Cys Cys Thr

1550 1555 1560

Thr Thr Cys Ala Ala Thr Gly Ala Ala Thr Gly Cys Gly Gly Thr

1565 1570 1575

Ala Ala Cys Gly Thr Cys Gly Ala Thr Ala Cys Ala Cys Thr Cys

1580 1585 1590

Cys Cys Ala Ala Thr Ala Ala Ala Ala Gly Ala Gly Ala Gly Gly

1595 1600 1605

Gly Ala Ala Cys Ala Ala Thr Thr Gly Ala Thr Thr Gly Ala Cys

1610 1615 1620

Ala Thr Ala Cys Gly Gly Ala Cys Thr Gly Ala Thr Ala Cys Gly

1625 1630 1635

Ala Cys Ala Thr Thr Gly Ala Ala Ala Thr Cys Thr Thr Cys Ala

1640 1645 1650

Thr Thr Cys Gly Thr Gly Cys Cys Thr Gly Ala Thr Gly Gly Thr

1655 1660 1665

Ala Thr Ala Gly Gly Ala Cys Cys Ala Thr Thr Cys Thr Gly Gly

1670 1675 1680

Ala Thr Cys Ala Ala Ala Cys Thr Gly Ala Thr Gly Gly Ala Cys

1685 1690 1695

Gly Ala Ala Thr Thr Thr Cys Cys Ala Gly Ala Ala Ala Thr Thr

1700 1705 1710

Ala Gly Cys Ala Ala Ala Cys Gly Ala Gly Cys Thr Gly Thr Cys

1715 1720 1725

Ala Ala Ala Gly Ala Gly Cys Thr Cys Ala Thr Gly Cys Cys Ala

1730 1735 1740

Thr Thr Thr Gly Thr Ala Ala Cys Cys Ala Cys Thr Thr Ala Cys

1745 1750 1755

Cys Thr Cys Thr Gly Thr Gly Ala Gly Ala Ala Ala Thr Cys Ala

1760 1765 1770

Thr Thr Thr Thr Cys Cys Gly Thr Cys Thr Ala Thr Gly Thr Ala

1775 1780 1785

Gly Cys Cys Ala Cys Ala Ala Ala Ala Ala Cys Ala Ala Ala Ala

1790 1795 1800

Thr Ala Thr Cys Gly Ala Ala Ala Thr Ala Gly Ala Cys Thr Thr

1805 1810 1815

Gly Ala Thr Gly Cys Thr Gly Ala Ala Gly Ala Cys Gly Ala Thr

1820 1825 1830

Ala Thr Gly Cys Gly Ala Cys Thr Cys Cys Ala Ala Cys Thr Thr

1835 1840 1845

Ala Cys Thr Ala Cys Thr Ala Thr Cys Cys Ala Thr Cys Cys Ala

1850 1855 1860

Gly Ala Cys Ala Thr Thr Gly Ala Cys Ala Ala Cys Cys Thr Thr

1865 1870 1875

Thr Gly Thr Ala Ala Cys Ala Ala Cys Ala Ala Gly Cys Ala Gly

1880 1885 1890

Gly Cys Thr Cys Ala Gly Ala Ala Ala Thr Cys Cys Cys Ala Cys

1895 1900 1905

Thr Gly Ala

1910

<210> 73

<211> 107

<212> PRT

<213> Artificial Sequence

<220><223> FKBP12 -40-18091

<400> 73

Gly Val Gln Val Glu Thr Ile Ser Pro Gly Asp Gly Arg Thr Phe Pro

1 5 10 15

Lys Arg Gly Gln Thr Cys Val Val His Tyr Thr Gly Met Leu Glu Asp

20 25 30

Gly Lys Lys Val Asp Ser Ser Arg Asp Arg Asn Lys Pro Phe Lys Phe

35 40 45

Met Leu Gly Lys Gln Glu Val Ile Arg Gly Trp Glu Glu Gly Val Ala

50 55 60

Gln Met Ser Val Gly Gln Arg Ala Lys Leu Thr Ile Ser Pro Asp Tyr

65 70 75 80

Ala Tyr Gly Ala Thr Gly His Pro Gly Ile Ile Pro Pro His Ala Thr

85 90 95

Leu Val Phe Asp Val Glu Leu Leu Lys Leu Glu

100 105

<210> 74

<211> 321

<212> DNA

<213> Artificial Sequence

<220><223> FKBP12 -40-18092

<400> 74

ggggtccagg tcgagactat ttcaccaggg gatgggcgaa catttccaaa aaggggccag 60

acttgcgtcg tgcattacac cgggatgctg gaggacggga agaaagtgga cagctccagg 120

gatcgcaaca agcccttcaa gttcatgctg ggaaagcagg aagtgatccg aggatgggag 180

gaaggcgtgg cacagatgtc agtcggccag cgggccaaac tgaccattag ccctgactac 240

gcttatggag caacaggcca cccagggatc attccccctc atgccaccct ggtcttcgat 300

gtggaactgc tgaagctgga g 321

<210> 75

<211> 5

<212> PRT

<213> Artificial Sequence

<220><223> linker-40-18093

<400> 75

Gly Gly Gly Gly Ser

1 5

<210> 76

<211> 15

<212> PRT

<213> Artificial Sequence

<220><223> linker-40-18094

<400> 76

Gly Gly Ala Gly Gly Ala Gly Gly Ala Gly Gly Ala Thr Cys Cys

1 5 10 15

<210> 77

<211> 282

<212> PRT

<213> Artificial Sequence

<220><223> truncatedCas9-40-18095

<400> 77

Gly Phe Gly Asp Val Gly Ala Leu Glu Ser Leu Arg Gly Asn Ala Asp

1 5 10 15

Leu Ala Tyr Ile Leu Ser Met Glu Pro Cys Gly His Cys Leu Ile Ile

20 25 30

Asn Asn Val Asn Phe Cys Arg Glu Ser Gly Leu Arg Thr Arg Thr Gly

35 40 45

Ser Asn Ile Asp Cys Glu Lys Leu Arg Arg Arg Phe Ser Ser Leu His

50 55 60

Phe Met Val Glu Val Lys Gly Asp Leu Thr Ala Lys Lys Met Val Leu

65 70 75 80

Ala Leu Leu Glu Leu Ala Gln Gln Asp His Gly Ala Leu Asp Cys Cys

85 90 95

Val Val Val Ile Leu Ser His Gly Cys Gln Ala Ser His Leu Gln Phe

100 105 110

Pro Gly Ala Val Tyr Gly Thr Asp Gly Cys Pro Val Ser Val Glu Lys

115 120 125

Ile Val Asn Ile Phe Asn Gly Thr Ser Cys Pro Ser Leu Gly Gly Lys

130 135 140

Pro Lys Leu Phe Phe Ile Gln Ala Cys Gly Gly Glu Gln Lys Asp His

145 150 155 160

Gly Phe Glu Val Ala Ser Thr Ser Pro Glu Asp Glu Ser Pro Gly Ser

165 170 175

Asn Pro Glu Pro Asp Ala Thr Pro Phe Gln Glu Gly Leu Arg Thr Phe

180 185 190

Asp Gln Leu Asp Ala Ile Ser Ser Leu Pro Thr Pro Ser Asp Ile Phe

195 200 205

Val Ser Tyr Ser Thr Phe Pro Gly Phe Val Ser Trp Arg Asp Pro Lys

210 215 220

Ser Gly Ser Trp Tyr Val Glu Thr Leu Asp Asp Ile Phe Glu Gln Trp

225 230 235 240

Ala His Ser Glu Asp Leu Gln Ser Leu Leu Leu Arg Val Ala Asn Ala

245 250 255

Val Ser Val Lys Gly Ile Tyr Lys Gln Met Pro Gly Cys Phe Asn Phe

260 265 270

Leu Arg Lys Lys Leu Phe Phe Lys Thr Ser

275 280

<210> 78

<211> 843

<212> DNA

<213> Artificial Sequence

<220><223> truncatedCas9-40-18096

<400> 78

tttggggacg tgggggccct ggagtctctg cgaggaaatg ccgatctggc ttacatcctg 60

agcatggaac cctgcggcca ctgtctgatc attaacaatg tgaacttctg cagagaaagc 120

ggactgcgaa cacggactgg ctccaatatt gactgtgaga agctgcggag aaggttctct 180

agtctgcact ttatggtcga agtgaaaggg gatctgaccg ccaagaaaat ggtgctggcc 240

ctgctggagc tggctcagca ggaccatgga gctctggatt gctgcgtggt cgtgatcctg 300

tcccacgggt gccaggcttc tcatctgcag ttccccggag cagtgtacgg aacagacggc 360

tgtcctgtca gcgtggagaa gatcgtcaac atcttcaacg gcacttcttg ccctagtctg 420

gggggaaagc caaaactgtt ctttatccag gcctgtggcg gggaacagaa agatcacggc 480

ttcgaggtgg ccagcaccag ccctgaggac gaatcaccag ggagcaaccc tgaaccagat 540

gcaactccat tccaggaggg actgaggacc tttgaccagc tggatgctat ctcaagcctg 600

cccactccta gtgacatttt cgtgtcttac agtaccttcc caggctttgt ctcatggcgc 660

gatcccaagt cagggagctg gtacgtggag acactggacg acatctttga acagtgggcc 720

cattcagagg acctgcagag cctgctgctg cgagtggcaa acgctgtctc tgtgaagggc 780

atctacaaac agatgcccgg gtgcttcaat tttctgagaa agaaactgtt ctttaagact 840

tcc 843

<210> 79

<211> 394

<212> PRT

<213> Artificial Sequence

<220><223> Inducible proapoptotic polypeptides

<400> 79

Gly Val Gln Val Glu Thr Ile Ser Pro Gly Asp Gly Arg Thr Phe Pro
1 5 10 15

Lys Arg Gly Gln Thr Cys Val Val His Tyr Thr Gly Met Leu Glu Asp
20 25 30
Gly Lys Lys Val Asp Ser Ser Arg Asp Arg Asn Lys Pro Phe Lys Phe
35 40 45
Met Leu Gly Lys Gln Glu Val Ile Arg Gly Trp Glu Glu Gly Val Ala
50 55 60
Gln Met Ser Val Gly Gln Arg Ala Lys Leu Thr Ile Ser Pro Asp Tyr
65 70 75 80

Ala Tyr Gly Ala Thr Gly His Pro Gly Ile Ile Pro Pro His Ala Thr
85 90 95
Leu Val Phe Asp Val Glu Leu Leu Lys Leu Glu Gly Gly Gly Gly Ser
100 105 110
Gly Phe Gly Asp Val Gly Ala Leu Glu Ser Leu Arg Gly Asn Ala Asp
115 120 125
Leu Ala Tyr Ile Leu Ser Met Glu Pro Cys Gly His Cys Leu Ile Ile
130 135 140

Asn Asn Val Asn Phe Cys Arg Glu Ser Gly Leu Arg Thr Arg Thr Gly
145 150 155 160
Ser Asn Ile Asp Cys Glu Lys Leu Arg Arg Arg Phe Ser Ser Leu His
165 170 175
Phe Met Val Glu Val Lys Gly Asp Leu Thr Ala Lys Lys Met Val Leu
180 185 190
Ala Leu Leu Glu Leu Ala Gln Gln Asp His Gly Ala Leu Asp Cys Cys
195 200 205

Val Val Val Ile Leu Ser His Gly Cys Gln Ala Ser His Leu Gln Phe
210 215 220
Pro Gly Ala Val Tyr Gly Thr Asp Gly Cys Pro Val Ser Val Glu Lys
225 230 235 240
Ile Val Asn Ile Phe Asn Gly Thr Ser Cys Pro Ser Leu Gly Gly Lys

245 250 255

Pro Lys Leu Phe Phe Ile Gln Ala Cys Gly Gly Glu Gln Lys Asp His

260 265 270

Gly Phe Glu Val Ala Ser Thr Ser Pro Glu Asp Glu Ser Pro Gly Ser

275 280 285

Asn Pro Glu Pro Asp Ala Thr Pro Phe Gln Glu Gly Leu Arg Thr Phe

290 295 300

Asp Gln Leu Asp Ala Ile Ser Ser Leu Pro Thr Pro Ser Asp Ile Phe

305 310 315 320

Val Ser Tyr Ser Thr Phe Pro Gly Phe Val Ser Trp Arg Asp Pro Lys

325 330 335

Ser Gly Ser Trp Tyr Val Glu Thr Leu Asp Asp Ile Phe Glu Gln Trp

340 345 350

Ala His Ser Glu Asp Leu Gln Ser Leu Leu Leu Arg Val Ala Asn Ala

355 360 365

Val Ser Val Lys Gly Ile Tyr Lys Gln Met Pro Gly Cys Phe Asn Phe

370 375 380

Leu Arg Lys Lys Leu Phe Phe Lys Thr Ser

385 390

<210> 80

<211> 1182

<212

> DNA

<213> Artificial Sequence

<220><223> Inducible proapoptotic polypeptides

<400> 80

ggggtccagg tcgagactat ttcaccaggg gatgggcgaa catttccaaa aaggggccag 60

acttgcgtcg tgcattacac cgggatgctg gaggacggga agaaagtgga cagctccagg 120

gatcgcaaca agcccttcaa gttcatgctg ggaaagcagg aagtgatccg aggatgggag 180

gaaggcgtgg cacagatgtc agtcggccag cgggccaaac tgaccattag ccctgactac 240

gcttatggag caacaggcca cccagggatc attccccctc atgccaccct ggtcttcgat 300

gtggaactgc tgaagctgga gggaggagga ggatccggat ttggggacgt gggggccctg 360

gagtctctgc gaggaaatgc cgatctggct tacatcctga gcatggaacc ctgcggccac 420

tgtctgatca ttaacaatgt gaacttctgc agagaaagcg gactgcgaac acggactggc 480

tccaatattg actgtgagaa gctgcggaga aggttctcta gtctgcactt tatggtcgaa 540

gtgaaagggg atctgaccgc caagaaaatg gtgctggccc tgctggagct ggctcagcag 600

gaccatggag ctctggattg ctgcgtggtc gtgatcctgt cccacgggtg ccaggcttct 660

catctgcagt tccccggagc agtgtacgga acagacggct gtcctgtcag cgtggagaag 720

atcgtcaaca tcttcaacgg cacttcttgc cctagtctgg ggggaaagcc aaaactgttc 780

tttatccagg cctgtggcgg ggaacagaaa gatcacggct tcgaggtggc cagcaccagc 840

cctgaggacg aatcaccagg gagcaaccct gaaccagatg caactccatt ccaggaggga 900

ctgaggacct ttgaccagct ggatgctatc tcaagcctgc ccactcctag tgacattttc 960

gtgtcttaca gtaccttccc aggctttgtc tcatggcgcg atcccaagtc agggagctgg 1020

tacgtggaga cactggacga catctttgaa cagtgggccc attcagagga cctgcagagc 1080

ctgctgctgc gagtggcaaa cgctgtctct gtgaagggca tctacaaaca gatgcccggg 1140

tgcttcaatt ttctgagaaa gaaactgttc tttaagactt cc 1182

<210> 81

<211> 463

<212> PRT

<213> Artificial Sequence

<220><223> CSR-CD2z-46-17062

<400> 81

Met Ser Phe Pro Cys Lys Phe Val Ala Ser Phe Leu Leu Ile Phe Asn

1 5 10 15

Val Ser Ser Lys Gly Ala Val Ser Lys Glu Ile Thr Asn Ala Leu Glu

20 25 30

Thr Trp Gly Ala Leu Gly Gln Asp Ile Asn Leu Asp Ile Pro Ser Phe

35 40 45

Gln Met Ser Asp Asp Ile Asp Asp Ile Lys Trp Glu Lys Thr Ser Asp

50 55 60

Lys Lys Lys Ile Ala Gln Phe Arg Lys Glu Lys Glu Thr Phe Lys Glu

65 70 75 80

Lys Asp Thr Tyr Lys Leu Phe Lys Asn Gly Thr Leu Lys Ile Lys His

85 90 95

Leu Lys Thr Asp Asp Gln Asp Ile Tyr Lys Val Ser Ile Tyr Asp Thr
100 105 110
Lys Gly Lys Asn Val Leu Glu Lys Ile Phe Asp Leu Lys Ile Gln Glu

115 120 125
Arg Val Ser Lys Pro Lys Ile Ser Trp Thr Cys Ile Asn Thr Thr Leu
130 135 140
Thr Cys Glu Val Met Asn Gly Thr Asp Pro Glu Leu Asn Leu Tyr Gln
145 150 155 160
Asp Gly Lys His Leu Lys Leu Ser Gln Arg Val Ile Thr His Lys Trp
165 170 175
Thr Thr Ser Leu Ser Ala Lys Phe Lys Cys Thr Ala Gly Asn Lys Val

180 185 190
Ser Lys Glu Ser Ser Val Glu Pro Val Ser Cys Pro Glu Lys Gly Leu
195 200 205
Asp Ile Tyr Leu Ile Ile Gly Ile Cys Gly Gly Gly Ser Leu Leu Met
210 215 220
Val Phe Val Ala Leu Leu Val Phe Tyr Ile Thr Lys Arg Lys Lys Gln
225 230 235 240
Arg Ser Arg Arg Asn Asp Glu Glu Leu Glu Thr Arg Ala His Arg Val

245 250 255
Ala Thr Glu Glu Arg Gly Arg Lys Pro His Gln Ile Pro Ala Ser Thr
260 265 270
Pro Gln Asn Pro Ala Thr Ser Gln His Pro Pro Pro Pro Pro Gly His
275 280 285
Arg Ser Gln Ala Pro Ser His Arg Pro Pro Pro Pro Gly His Arg Val
290 295 300
Gln His Gln Pro Gln Lys Arg Pro Pro Ala Pro Ser Gly Thr Gln Val

305 310 315 320
His Gln Gln Lys Gly Pro Pro Leu Pro Arg Pro Arg Val Gln Pro Lys
325 330 335
Pro Pro His Gly Ala Ala Glu Asn Ser Leu Ser Pro Ser Ser Asn Arg

340 345 350

Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly Gln

355 360 365

Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp

370 375 380

Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro

385 390 395 400

Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp

405 410 415

Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg

420 425 430

Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr

435 440 445

Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg

450 455 460

<210> 82

<211> 184

<212> PRT

<213> Artificial Sequence

<220><223> CD2 ECD with D11H-46-17119

<400> 82

Lys Glu Ile Thr Asn Ala Leu Glu Thr Trp Gly Ala Leu Gly Gln Asp

1 5 10 15

Ile Asn Leu Asp Ile Pro Ser Phe Gln Met Ser Asp Asp Ile Asp Asp

20 25 30

Ile Lys Trp Glu Lys Thr Ser Asp Lys Lys Lys Ile Ala Gln Phe Arg

35 40 45

Lys Glu Lys Glu Thr Phe Lys Glu Lys Asp Thr Tyr Lys Leu Phe Lys

50 55 60

Asn Gly Thr Leu Lys Ile Lys His Leu Lys Thr Asp Asp Gln Asp Ile

65 70 75 80

Tyr Lys Val Ser Ile Tyr His Thr Lys Gly Lys Asn Val Leu Glu Lys

85 90 95

Ile Phe Asp Leu Lys Ile Gln Glu Arg Val Ser Lys Pro Lys Ile Ser
100 105 110
Trp Thr Cys Ile Asn Thr Thr Leu Thr Cys Glu Val Met Asn Gly Thr
115 120 125
Asp Pro Glu Leu Asn Leu Tyr Gln Asp Gly Lys His Leu Lys Leu Ser
130 135 140
Gln Arg Val Ile Thr His Lys Trp Thr Thr Ser Leu Ser Ala Lys Phe
145 150 155 160

Lys Cys Thr Ala Gly Asn Lys Val Ser Lys Glu Ser Ser Val Glu Pro
165 170 175
Val Ser Cys Pro Glu Lys Gly Leu
180

<210> 83
<211> 463
<212> PRT
<213> Artificial Sequence
<220><223> CSR CD2z-D111H-46-17118
<400> 83
Met Ser Phe Pro Cys Lys Phe Val Ala Ser Phe Leu Leu Ile Phe Asn
1 5 10 15
Val Ser Ser Lys Gly Ala Val Ser Lys Glu Ile Thr Asn Ala Leu Glu

20 25 30
Thr Trp Gly Ala Leu Gly Gln Asp Ile Asn Leu Asp Ile Pro Ser Phe
35 40 45
Gln Met Ser Asp Asp Ile Asp Asp Ile Lys Trp Glu Lys Thr Ser Asp
50 55 60
Lys Lys Lys Ile Ala Gln Phe Arg Lys Glu Lys Glu Thr Phe Lys Glu
65 70 75 80
Lys Asp Thr Tyr Lys Leu Phe Lys Asn Gly Thr Leu Lys Ile Lys His

85 90 95

Leu Lys Thr Asp Asp Gln Asp Ile Tyr Lys Val Ser Ile Tyr His Thr
100 105 110
Lys Gly Lys Asn Val Leu Glu Lys Ile Phe Asp Leu Lys Ile Gln Glu
115 120 125
Arg Val Ser Lys Pro Lys Ile Ser Trp Thr Cys Ile Asn Thr Thr Leu
130 135 140
Thr Cys Glu Val Met Asn Gly Thr Asp Pro Glu Leu Asn Leu Tyr Gln

145 150 155 160
Asp Gly Lys His Leu Lys Leu Ser Gln Arg Val Ile Thr His Lys Trp
165 170 175
Thr Thr Ser Leu Ser Ala Lys Phe Lys Cys Thr Ala Gly Asn Lys Val
180 185 190
Ser Lys Glu Ser Ser Val Glu Pro Val Ser Cys Pro Glu Lys Gly Leu
195 200 205
Asp Ile Tyr Leu Ile Ile Gly Ile Cys Gly Gly Gly Ser Leu Leu Met

210 215 220
Val Phe Val Ala Leu Leu Val Phe Tyr Ile Thr Lys Arg Lys Lys Gln
225 230 235 240
Arg Ser Arg Arg Asn Asp Glu Glu Leu Glu Thr Arg Ala His Arg Val
245 250 255
Ala Thr Glu Glu Arg Gly Arg Lys Pro His Gln Ile Pro Ala Ser Thr
260 265 270
Pro Gln Asn Pro Ala Thr Ser Gln His Pro Pro Pro Pro Pro Gly His

275 280 285
Arg Ser Gln Ala Pro Ser His Arg Pro Pro Pro Pro Gly His Arg Val
290 295 300
Gln His Gln Pro Gln Lys Arg Pro Pro Ala Pro Ser Gly Thr Gln Val
305 310 315 320
His Gln Gln Lys Gly Pro Pro Leu Pro Arg Pro Arg Val Gln Pro Lys
325 330 335
Pro Pro His Gly Ala Ala Glu Asn Ser Leu Ser Pro Ser Ser Asn Arg

340 345 350

Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly Gln

355 360 365

Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp

370 375 380

Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro

385 390 395 400

Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp

405 410 415

Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg

420 425 430

Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr

435 440 445

Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg

450 455 460

<210> 84

<211> 89

<212> PRT

<213> Homo Sapiens

<400> 84

Leu Pro Ala Pro Lys Asn Leu Val Val Ser Glu Val Thr Glu Asp Ser

1 5 10 15

Leu Arg Leu Ser Trp Thr Ala Pro Asp Ala Ala Phe Asp Ser Phe Leu

20 25 30

Ile Gln Tyr Gln Glu Ser Glu Lys Val Gly Glu Ala Ile Asn Leu Thr

35 40 45

Val Pro Gly Ser Glu Arg Ser Tyr Asp Leu Thr Gly Leu Lys Pro Gly

50 55 60

Thr Glu Tyr Thr Val Ser Ile Tyr Gly Val Lys Gly Gly His Arg Ser

65 70 75 80

Asn Pro Leu Ser Ala Glu Phe Thr Thr

85

<210> 85

<211> 90

<212> PRT

<213> Homo Sapiens

<400> 85

Met Leu Pro Ala Pro Lys Asn Leu Val Val Ser Glu Val Thr Glu Asp
1 5 10 15

Ser Leu Arg Leu Ser Trp Thr Ala Pro Asp Ala Ala Phe Asp Ser Phe
20 25 30

Leu Ile Gln Tyr Gln Glu Ser Glu Lys Val Gly Glu Ala Ile Asn Leu
35 40 45

Thr Val Pro Gly Ser Glu Arg Ser Tyr Asp Leu Thr Gly Leu Lys Pro
50 55 60

Gly Thr Glu Tyr Thr Val Ser Ile Tyr Gly Val Lys Gly Gly His Arg
65 70 75 80

Ser Asn Pro Leu Ser Ala Glu Phe Thr Thr
85 90

<210> 86

<211> 270

<212

> DNA

<213> Homo Sapiens

<400> 86

atgctgcctg caccaaagaa cctggtggtg tctcatgtga cagaggatag tgccagactg 60

tcatggactg ctcccgacgc agccttcgat agttttatca tcgtgtaccg ggagaacatc 120

gaaaccggcg aggccattgt cctgacagtg ccagggtccg aacgctctta tgacctgaca 180

gatctgaagc ccggaactga gtactatgtg cagatcgccg gcgtcaaagg aggcaatatc 240

agcttccctc tgtccgcaat cttcaccaca 270

<210> 87

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> A-B Loop

<400> 87

Thr Glu Asp Ser

1

<210> 88

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> A-B Loop

<400> 88

Thr Ala Pro Asp Ala Ala Phe

1 5

<210> 89

<211> 6

<212> PRT

<213> Artificial Sequence

<220><223> A-B Loop

<400> 89

Ser Glu Lys Val Gly Glu

1 5

<210> 90

<211> 4

<212> PRT

<213> Artificial Sequence

<220><223> D-E Loop

<400> 90

Gly Ser Glu Arg

1

<210> 91

<

211> 5

<212> PRT

<213> Artificial Sequence

<220><223> E-F Loop

<400> 91

Gly Leu Lys Pro Gly

1 5

<210> 92

<211> 7

<212> PRT

<213> Artificial Sequence

<220><223> F-G Loop

<400> 92

Lys Gly Gly His Arg Ser Asn

1 5

<210> 93

<211> 411

<212> DNA

<213> Artificial Sequence

<220><223> PSMA centyrin

<400> 93

atgctgcctg caccaaagaa cctggtggtg tctcgggtga ccgaggactc tgccagactg 60

agctgggcca tcgacgagca gagggattgg ttcgagagct ttctgatcca gtatcaggag 120

tccgagaaag tgggcgaggc catcgtgctg acagtgcctg gcagcgagcg gtcctatgat 180

ctgaccggcc tgaagccagg cacagagtac accgtgtcca tctacggcgt gtatcacgtg 240

tacaggtcca atcctctgtc tgccatcttc accacaacca caacccctgc ccccagacct 300

cccacacccg cccctaccat cgcgagtcag cccctgagtc tgagacctga ggcctgcagg 360

ccagctgcag gaggagctgt gcacaccagg ggcctggact tcgcctgcga c 411

<210> 94

<211> 92

<212> PRT

<213> Artificial Sequence

<220><223> PSMA centyrin

<400> 94

Met Leu Pro Ala Pro Lys Asn Leu Val Val Ser Arg Val Thr Glu Asp

1 5 10 15

Ser Ala Arg Leu Ser Trp Ala Ile Asp Glu Gln Arg Asp Trp Phe Glu

20 25 30

Ser Phe Leu Ile Gln Tyr Gln Glu Ser Glu Lys Val Gly Glu Ala Ile

35 40 45

Val Leu Thr Val Pro Gly Ser Glu Arg Ser Tyr Asp Leu Thr Gly Leu

50 55 60

Lys Pro Gly Thr Glu Tyr Thr Val Ser Ile Tyr Gly Val Tyr His Val

65 70 75 80

Tyr Arg Ser Asn Pro Leu Ser Ala Ile Phe Thr Thr

85 90

<210> 95

<211> 336

<212> PRT

<213> Artificial Sequence

<220><223> PSMA CARTyrin

<400> 95

Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu

1 5 10 15

His Ala Ala Arg Pro Met Leu Pro Ala Pro Lys Asn Leu Val Val Ser

20 25 30

Arg Val Thr Glu Asp Ser Ala Arg Leu Ser Trp Ala Ile Asp Glu Gln

35 40 45

Arg Asp Trp Phe Glu Ser Phe Leu Ile Gln Tyr Gln Glu Ser Glu Lys

50 55 60

Val Gly Glu Ala Ile Val Leu Thr Val Pro Gly Ser Glu Arg Ser Tyr

65 70 75 80

Asp Leu Thr Gly Leu Lys Pro Gly Thr Glu Tyr Thr Val Ser Ile Tyr

85 90 95

Gly Val Tyr His Val Tyr Arg Ser Asn Pro Leu Ser Ala Ile Phe Thr

100 105 110

Thr Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile

115 120 125

Ala Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala

130 135 140

Gly Gly Ala Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr

145 150 155 160

Ile Trp Ala Pro Leu Ala Gly Thr Cys Gly Val Leu Leu Leu Ser Leu
165 170 175
Val Ile Thr Leu Tyr Cys Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile
180 185 190
Phe Lys Gln Pro Phe Met Arg Pro Val Gln Thr Thr Gln Glu Glu Asp
195 200 205
Gly Cys Ser Cys Arg Phe Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu
210 215 220

Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly
225 230 235 240
Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr
245 250 255
Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly Lys
260 265 270
Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys
275 280 285

Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg
290 295 300
Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala
305 310 315 320
Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro Arg
325 330 335

<210> 96
<211> 489
<212> DNA
<213> Artificial Sequence
<220><223> BCMA VH
<400> 96

gaagttcagc tgcttgaatc tggcggaggc ctggttcaac ctggcggatc tctgagactg 60

agctgtgccg ccagcggctt caccttcagc aattacgcca tgacctggat cagacaggcc 120

cctggcaaag gcctggaatg ggtgtccgga attacaggcg acggcggcag caccttttac 180

gccgattctg tgaagggcag attcaccatc agccgggaca acagcaagaa caccctgtac 240

ctgcagatga acagcctgag agccgaggac accgccgtgt actactgcgt gaaggactgg 300

aacaccacca tgatcaccga gagaggccag ggcacactgg tcaccgtgtc ctctacaaca 360

acaccggcgc ctagacctcc aacaccagct cctacaatcg cgagtcagcc cctgtctctc 420

agacccgaag cctgcaggcc agctgcagga ggagctgtgc acaccagggg cctggacttc 480

gcctgcgac 489

<210> 97

<211> 118

<212> PRT

<213> Artificial Sequence

<220><223> BCMA VH

<400> 97

Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Asn Tyr
20 25 30

Ala Met Thr Trp Ile Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
35 40 45

Ser Gly Ile Thr Gly Asp Gly Gly Ser Thr Phe Tyr Ala Asp Ser Val
50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65 70 75 80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Val Lys Asp Trp Asn Thr Thr Met Ile Thr Glu Arg Gly Gln Gly Thr
100 105 110

Leu Val Thr Val Ser Ser
115

<210> 98

<211> 362

<212> PRT

<213> Artificial Sequence

<220><223> BCMA VCAR

<400> 98

Met Ala Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu
1 5 10 15
His Ala Ala Arg Pro Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu
20 25 30
Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe
35 40 45

Thr Phe Ser Asn Tyr Ala Met Thr Trp Ile Arg Gln Ala Pro Gly Lys
50 55 60
Gly Leu Glu Trp Val Ser Gly Ile Thr Gly Asp Gly Gly Ser Thr Phe
65 70 75 80
Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser
85 90 95
Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr
100 105 110

Ala Val Tyr Tyr Cys Val Lys Asp Trp Asn Thr Thr Met Ile Thr Glu
115 120 125
Arg Gly Gln Gly Thr Leu Val Thr Val Ser Ser Thr Thr Thr Pro Ala
130 135 140
Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile Ala Ser Gln Pro Leu Ser
145 150 155 160
Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala Gly Gly Ala Val His Thr
165 170 175

Arg Gly Leu Asp Phe Ala Cys Asp Ile Tyr Ile Trp Ala Pro Leu Ala
180 185 190
Gly Thr Cys Gly Val Leu Leu Leu Ser Leu Val Ile Thr Leu Tyr Cys
195 200 205
Lys Arg Gly Arg Lys Lys Leu Leu Tyr Ile Phe Lys Gln Pro Phe Met
210 215 220
Arg Pro Val Gln Thr Thr Gln Glu Glu Asp Gly Cys Ser Cys Arg Phe

225 230 235 240

Pro Glu Glu Glu Glu Gly Gly Cys Glu Leu Arg Val Lys Phe Ser Arg
245 250 255
Ser Ala Asp Ala Pro Ala Tyr Lys Gln Gly Gln Asn Gln Leu Tyr Asn
260 265 270
Glu Leu Asn Leu Gly Arg Arg Glu Glu Tyr Asp Val Leu Asp Lys Arg
275 280 285
Arg Gly Arg Asp Pro Glu Met Gly Gly Lys Pro Arg Arg Lys Asn Pro
290 295 300

Gln Glu Gly Leu Tyr Asn Glu Leu Gln Lys Asp Lys Met Ala Glu Ala
305 310 315 320
Tyr Ser Glu Ile Gly Met Lys Gly Glu Arg Arg Arg Gly Lys Gly His
325 330 335
Asp Gly Leu Tyr Gln Gly Leu Ser Thr Ala Thr Lys Asp Thr Tyr Asp
340 345 350
Ala Leu His Met Gln Ala Leu Pro Pro Arg
355 360